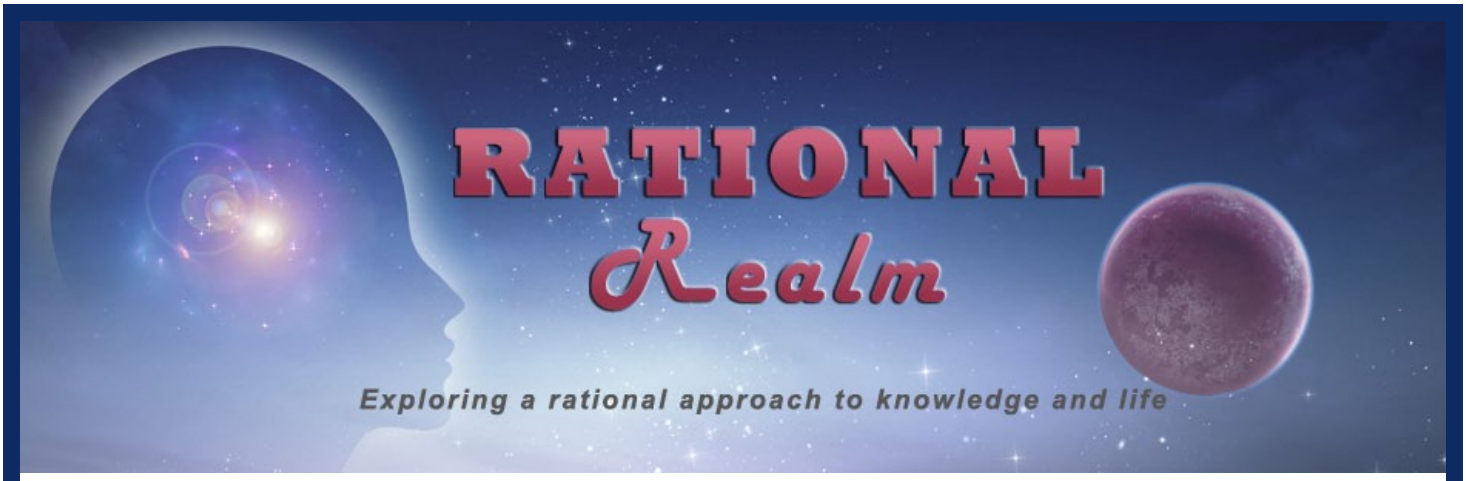

Towards an Objective Theory of Rationality

Leslie Allan

Published online: 18 January 2016

Drawing on insights from Imre Lakatos' seminal work on theories of rationality, Leslie Allan develops seven criteria for rational theory choice that surmount a major shortcoming of Lakatos' methodology. By articulating the various criteria of dependence and independence applying to evidence-statements, Allan shows how his axioms of rationality follow from the general demands of an objectivist epistemology. He thus avoids Lakatos' Achilles' heel; the presupposition that science is a rational enterprise. Allan concludes his essay with a consideration of two problems for his theory: the problem of logical incommensurability and of implicit rationality.

To cite this essay:

Leslie Allan 2016. Towards an Objective Theory of Rationality, URL = <http://www.RationalRealm.com/philosophy/epistemology/objective-theory-rationality.html>

To link to this essay:

www.RationalRealm.com/philosophy/epistemology/objective-theory-rationality.html

Follow this and additional essays at: www.RationalRealm.com

1. Introduction

Imre Lakatos may be characterized as having held two major theses in the philosophy of science. The first thesis was proposed in answer to the question, 'What is science and what are its methods?' Lakatos' answer to this problem was that science is a body of knowledge arrived at by a methodology of competing research programmes.[1] His second thesis was in answer to the problem of developing and testing theories of rationality. Lakatos proposed that a theory of rationality can be judged by how well it reconstructs the history of science as a rational enterprise. Not surprisingly, he offered his theory of rationality, based on his methodology of scientific research programmes, as the most adequate candidate for such a reconstruction.[2]

In this essay, I critically examine and further develop Lakatos' second thesis concerning theories of rationality. Lakatos' criteria for rational appraisal based on the successful prediction of novel facts are a significant advance on previously proposed criteria. Although, I shall argue, his attempted validation of his criteria by turning to the history of science is not altogether satisfactory. The major part of this essay will be devoted to remedying this deficiency by characterizing the demands of a general objectivist epistemology and demonstrating how Lakatos' criteria satisfy these demands.

I attempt this endeavour in two stages. In the first stage (§4), I introduce five criteria for objective theory choice. These five criteria explicate the necessary logical dependence relationships between evidence-statements and the theory under evaluation, on the one hand, and the necessary historical and psychological independence of evidence-statements from the theory on the other.

To complete the conditions for rational theory choice requires a commitment to either an idealist or realist metaphysical framework. In the second stage (§5 and §6), I make a case for realism and, on that basis, submit two further criteria for the rational acceptance of observation statements.

I conclude with a mapping of my seven criteria for objective theory choice to Lakatos' methodology and a review of two outstanding issues for objectivism. The first of these that I deal with is the problem of logical incommensurability between rival scientific theories. The second is the fact that different scientists claim to use different scientific methods. I show how neither of these is an impediment to the acceptance of my criteria for rational theory choice. At the end of the day, I will have achieved my objectives if I have gone some way in answering Feyerabend's question, 'What's so great about science?' [1979: 110]

[1] See Lakatos [1978a: ch. 1], also reprinted in Lakatos and Musgrave [1970: 91–196]. For an earlier formulation, see Lakatos [1978b: part 2, ch. 8, §6].

[2] See Lakatos [1978a: ch. 2].

2. The Methodology of Scientific Research Programmes

I accept Lakatos' description of science as a system of competing research programmes to be substantially correct. It does exhibit minor weaknesses and I offer some suggestions in remedying these in my Allan [2016b].

How did Lakatos see the nature and function of the research programme? Lakatos had considered that the unit of appraisal in science was not the isolated theory, but a research programme. To know whether a theory constitutes a part of science, according to Lakatos, it is necessary to know its history. If it had been arrived at by content reducing *ad hoc* modifications, in the face of anomalies, of earlier theories, it is not 'scientific'. It is a series of theories—a research programme—then, that is deemed 'scientific' or 'pseudo-scientific'.

A research programme, Lakatos explained, is composed of a 'negative heuristic' and a 'positive heuristic'. The 'negative heuristic' specifies the 'hard core' of the programme; its metaphysical foundations or conceptual framework. This 'hard core' is deemed irrefutable by the methodological fiat of the programme's proponents. Every worthwhile research programme develops in an ocean of anomalies. It is the function of the 'negative heuristic' of the programme to prevent such anomalies from refuting the 'hard core' by directing the scientists' attention to the revision of the 'protective belt' of auxiliary hypotheses and initial conditions. Just how the 'protective belt' is to be modified is specified by a partially articulated plan; the 'positive heuristic'.

A research programme was regarded by Lakatos as 'progressive' if the successive modifications of its protective belt satisfy the following two conditions. Firstly, each successive modification must be 'theoretically progressive', or have 'excess empirical content' in the sense that the new theory, which consists of laws of nature, auxiliary hypotheses and initial conditions, must predict some hitherto unexpected, novel fact. Secondly, the modifications must be 'empirically progressive' in the sense that the predicted novel facts must be at least occasionally corroborated. Conversely, a programme that is not 'progressive' is deemed 'degenerating'. Lakatos considered that for a research programme to be 'scientific', it must be at least theoretically progressive. For one research programme to supersede a rival, he added, it must be progressive while its rival is degenerating. Furthermore, it must satisfactorily explain the previous predictive successes of its rival. I now turn to his second thesis concerning theories of rationality.

3. Is Science a Rational Enterprise?

Lakatos had not only argued that his methodology of scientific research programmes (MSRP) is the best available characterization of scientific method, but that it is also the most adequate theory of rationality. According to his theory of rationality, it is rational to epistemically value progressive research programmes while epistemically devaluing degenerating programmes. Naturally, he considered it irrational to do the converse. In his [1978a: 119], Lakatos attempted to defend this thesis in a novel manner.

Lakatos' argument was this. All of the methodologies of science so far proposed have been offered as normative criteria for scientific rationality. Each methodology serves to provide a different rational reconstruction of the history of science. Furthermore, they differ on where to place the dividing line separating what is to be explained 'internally' as the idealized application of scientific method and what is to be explained 'externally', in socio-psychological terms, as the difference between the rationally reconstructed 'internal' history and actual history.

Conventionalist methodologies are notoriously difficult to criticize on logical and epistemological grounds. However, Lakatos proposed, they may be criticized for how inadequately they provide a rational reconstruction of the history of science. Judging how well a methodology rationally reconstructs the history of science by using that methodology itself, however, is inadequate as all methodologies fail their own standards. A more adequate normative historiographical meta-criterion is a meta-methodology of scientific research programmes or, what Lakatos termed, a 'methodology of historiographical research programmes' (MHRP). Different methodologies are now seen as the hard cores of normative historiographical research programmes. A progressive problem-shift in the research programme is marked by its ability to reconstruct more of the basic value judgements of scientists as rational, lead to an empirically progressive revision of some previously held basic value judgements, predict novel historical facts and anticipate further basic value judgements.

The advantage of the MHRP is that it allows the proponents of a historiographical research programme to ignore anomalies in the history of science as long as the programme is progressing. Secondly, the proponents of the programme need only take notice of criticism if it is constructive; that is, if the criticism will further our knowledge of method. On this meta-criterion, the MSRP is progressive since it reconstructs more of the history of science as rational, has led to the reversal of some historiographical appraisals and has successfully predicted novel historical facts.[3]

Lakatos [1978a: 132] was correct in 'maintaining that a theory of rationality has to try to organize basic value judgements in universal, coherent frameworks'. However, he failed to raise, let alone answer, the question of why Enlightenment and post-Enlightenment physics should be regarded as the paradigm of rationality. Lakatos had given us no argument as to why it is not more reasonable to accept, say, the basic value judgements of fundamentalist theologians as exemplars of rationality and test our rational reconstructions

[3] I don't think Lakatos' defence of his MHRP as a meta-criterion is wholly adequate. In my Allan [2016b], I propose what I think is a more successful vindication of Lakatos' criteria.

of the history of fundamentalist theology against these judgements. 'Rational reconstructions', therefore, cannot be methodologically instructive unless we have some defensible reason for our choice of historical subject matter.[4]

So, a methodologist, faced with a failure of his methodology *qua* rational reconstruction of science to stand up to historical tests, has the option of retaining his methodology while discarding the thesis that science is a rational enterprise. For Lakatos simply to charge such people with 'disrespect' [1978a: 127] and 'temerity' [1978a: 129], and to leave it at that, is just not good enough.

In spite of its plausibility in judging rival theories of rationality, the MHRP lacks persuasive force on its own. The arguments Lakatos advances in its favour are simply a meta-level application of the MSRP and so will not be persuasive to those not already partial to this particular methodology. For example, the supposed advantage that the MHRP can judge progress by the extent to which a methodology successfully postdicts novel facts will only have weight for those methodologists that already recognize the epistemic worth of corroborated novel facts. This is a serious weakness of Lakatos' meta-level defence of the MSRP.

I do think, though, that Lakatos' thesis is of some value if it is supplemented with an argument for the *prima facie* rationality of the scientific enterprise. Such an argument would refer to the way in which the spectacular predictive and technological success of the sciences was foreshadowed by deliberate theoretical developments, making it even more unlikely that this success was due to 'fantastic networks of coincidences'.[5] Any alternative theory of rationality must adequately explain this predictive and technological power as being the byproduct of irrational and non-rational forces. Seen in this light, the testing of methodologies against the history of science has *prima facie* plausibility.

In spite of this support, such dependence on a *prima facie* case for the rationality of science serves to weaken the force of tests of methodologies using the history of science. The arguments for Lakatos' meta-criterion partly presuppose the cogency of the MSRP as a theory of rationality. In this respect, it is of limited effect in independently supporting the MSRP. It is for these reasons that I have sought to provide independent arguments for the MSRP. My aim is to achieve this through demonstrating that its criteria are demanded by an analysis of the requirements of an objectivist epistemology. Before beginning this task, I want to look briefly at what Lakatos had to say about rationally justifying the MSRP *independently* of a prior acceptance of the rationality of science.

Lakatos' remarks on this are altogether uninspiring. (His brief discussions are contained in [1978a: 99–101, 154–67, 170–93; 1978b: 220–3]). For Lakatos, all we can do is 'hopefully guess' that the historical succession of scientific theories is leading us closer to the truth. [1970: 175; 1978a: 101 n.1; 1978b: 191, 223] However, there is even more depressing news. 'The body of science . . . cannot be the object of rational belief', Lakatos asserted, because this is a concession to psychologism in epistemology, and, anyway, the

[4]Feyerabend has forcefully argued this point in his [1975: 201–14] and again in his [1979: 109–20].

[5]In presenting one form of this argument, Shimony [1976: 474–8] refers to the difficulty in accounting for the success of science as the result of 'fantastic networks of coincidences'. Worrall [1976: 164] also briefly alludes to the *prima facie* plausibility of the rationality of science.

body of science is inconsistent [1978b: 176, 220]. Furthermore, the growth of knowledge is *via* research programmes, and one cannot 'rationally believe' a programme since a programme is more than a static set of propositions [1978b: 221]. Lastly, throughout the long history of human belief systems, there was no 'sudden change from animal belief to rational belief' [1978b: 221].

So, we are left with the paradoxical conclusion that the methods of science are rational, and yet all that we can do is base their epistemic value on 'hope'. There is also another paradox here. And that is that the epistemic value of the methods of science, and which deems such methods rational, is based on non-rational beliefs. Lakatos wrote:

> We may claim that progressive problemshifts do move us 'more likely than not' towards Truth rather than away from it. But this inductive principle which confers an epistemological status on our *convention* as to how to appraise problemshifts, is, in turn, backed by mere animal belief. Therefore problemshifts receive their epistemological rationality from animal belief (or, if you wish, from a bare postulate – an intellectual theft, as Russell used to characterize such 'posits').
>
> [Lakatos 1978b: 221]

And further on, in the same vein, he continued, 'There is no ultimate proof that, even where Elizabethan beliefs were replaced in the course of progressive problemshifts (like beliefs about heat, magnetism), we have been heading towards the Truth. We can only (non-rationally) believe, or, rather hope, that we have been. Unless hope is a "solution", there is no solution to Hume's problem.' [1978b: 223] In his 'Reply to Critics', he even more strongly attacked beliefs, regarding the belief that a hard core of a programme is true as one of the 'weaknesses of the human psyche', and even that it is 'naïve to believe either that one particular step [in the pursuit of Truth] is already part of the Truth or even that one is on the right path' [Lakatos 1970: 175].

So, according to Lakatos, the methods of science are especially rational, but the belief on which this rationality of science is based is naïve and non-rational, being 'animal belief'. To escape the charge that the ordaining of science as 'rational' is thus, at the very least, arbitrary (that is, why not, for example, believe that the methods of biblical fundamentalism are leading us closer to the Truth?), either the 'rationality' of science must be surrendered or we must admit that at least one belief is rational. But, if one belief is rational, why not more? In the next section, I will develop such a theory of 'rational belief'.

4. An Objectivist Epistemology

4.1 Preliminaries

A number of objectivist epistemologies have been proposed, and they have failed for various reasons. Contemporary epistemologies, such as conventionalism, pragmatism and falsificationism, in their attempts to overcome the problems of earlier foundationalist epistemologies, have led to a revision of some epistemological terms such as 'truth' and 'knowledge'. Such revisions are not undesirable in themselves, so long as it is understood that what is being offered is a revision of common-sense notions. If such a revision leads to greater clarity and a deeper understanding of the subject, this is well and good. However, some recent epistemologies introduce clearly subjective elements into rational appraisal, for example, considerations of simplicity (Poincaré's and Duhem's conventionalism) and psychological utility (James' pragmatism).

What I want to do here is to outline an objectivist epistemology that does justice to our commonly understood notions of 'truth', 'knowledge', 'rational', and so on, without sacrificing a thorough-going objectivism. I will develop seven criteria for assessing evidence-statements offered in support of a theory. The criteria can be divided into two types. The first type specifies the required logical relationships between the evidence-statement and the theory under evaluation. These criteria are relatively uncontroversial as they appear in other theories of rationality. I will call criteria of this type 'criteria of dependence'. The second type of criteria is what I refer to as 'criteria of independence'. They specify how the evidence-statement needs to be independent of the theory under evaluation and perceptual bias. It is this second set of criteria that marks off this theory of rationality from other more subjective schema.

How can we contrast an objectivist epistemology with its subjectivist rivals? I think we do that by considering what it means to be objective. For an entity to be objective, be it a person, social group, report or belief, it must of necessity be impartial, impersonal, disinterested and detached. Now contrast these requirements for objectivity with what it means to be subjective. To be subjective is of necessity to be prejudiced, biased or aligned. In a nutshell, to be objective is to be in a relation of independence from biasing factors.

Before discussing the criteria in some detail, I will deal with some preliminaries clarifying the meanings of terms and the nature of truth-bearers. I begin by giving an account of 'truth'. This term, as with most others found in natural languages, is vague. I think, though, that for our present purposes, its meaning may be adequately formalized as: 'what is actually the case'. I shall take the primitive truth-bearers to be interpreted sentences or propositions. It does not matter here which it is, but for convenience I shall assume it is interpreted sentences. So, 'x is true' means 'x states what is actually the case', where 'x' specifies an interpreted sentence.

This interpretation of 'true' is liberal enough to include logical truths, such as 'It is raining now or it is not raining now', in the class of true sentences. It is also broad enough to include subjunctive conditionals and statements about the past and future as contingent. I cannot expand on this anymore here, but must rely on the reader's intuitive understanding

of the concept to some extent. Needless to say, I think that coherence and pragmatist theories of 'truth' are explicating a notion that has little bearing on what we commonly understand by this term.

Next, I indicate the epistemic nature of truth-bearers in an objectivist epistemology. An objectivist epistemology stipulates that the epistemic status of a truth-bearer is a function of the properties of the truth-bearer itself and is independent of its relationship with subjective states, such as beliefs, preferences and attitudes. Alternatively, subjectivist and relativist epistemologies render the epistemic status of truth-bearers as the function of beliefs, preferences, attitudes, or some other subjective state of an individual or group of individuals. In the objectivist epistemology proposed here, then, whether a sentence, when interpreted, is true or false is dependent on the properties of the sentence, that is, whether it states what is actually the case. In contrast, its epistemic status is independent of its relations with subjective states. So, a sentence is true or false independently of whether some individual or group of individuals believe the sentence, or whether the sentence evokes praiseworthy thoughts or feelings of contentment, and so on.

I said that the *primitive* truth-bearers are sentences. I said this because we ordinarily also speak of beliefs as being true or false, and I think that we can give an account of this way of speaking by using the account of the epistemic status of sentences given above. A person's 'belief' can always be expanded into 'belief that x is true', where 'x' specifies an interpreted sentence. For example, 'Mary believes that New Delhi is the capital city of India' can be analysed as 'Mary believes that "New Delhi is the capital city of India" is true'. And Mary's belief is a true belief only if 'New Delhi is the capital city of India' is a true sentence. So, beliefs are also truth-bearers, but in a derivative sense.

I introduced the notion of 'true belief' in order to lead into the development of a theory of 'rational belief'. Not all contingent sentences of a language are indubitably true or false.[6] (This may also be the case for analytic sentences.)[7] An adequate objectivist epistemology, then, must specify acceptable criteria for rational belief in those cases in which the set of evidence statements do not lead indubitably to conclusive verification or falsification of a belief. This usage of the term 'rational', in which a belief may be correctly judged as 'rational' although further investigation may demonstrate its falsity, is given in ordinary language. We may therefore characterize the meaning of 'rational belief', sufficient for our purposes here, as 'epistemically justified belief', where 'epistemically justified' is further specified as 'epistemically justified relative to the acceptable set of available evidence-statements E'.

On this ordinary language account, for a person to assert genuinely that 'x is true' is to express a belief that 'x', where 'x' specifies an interpreted sentence. Furthermore, it is 'rational' for a person to assert genuinely that 'x is true', that is, to express a belief that 'x', iff 'x is true' is epistemically justified relative to the set of acceptable evidence-statements E available at the time. It may be 'rational' to believe that 'x' and to assert genuinely that 'x is true' even though 'x' may be false. This is in accord with our intuitive understanding of the meaning of 'rational'.

[6] I will give reasons below for thinking that no contingent sentences can be considered indubitably true or false.

[7] Here I use 'analytic' in the broad sense of 'analytically true' or 'analytically false'. Because of the fluidity of the analytic–synthetic distinction in natural languages, my comments here only apply to formalised, or roughly formalised, systems of sentences in which the meanings of sentences is conventionally fixed.

This leads us to the meaning of the term 'knowledge'. I shall accept here the conventional analysis of the meaning of 'knowledge' as 'justified true belief'. On this account, then, a person 'knows' that '*x*' iff that person's belief that '*x*' is a rational belief and '*x*' is true. One important consequence of this analysis is that it is logically impossible for a person to 'know *x*' and '*x*' be false. We can also conclude, then, that the set of our present rational beliefs constitute our sum of knowledge, for they are rational and true.

It may be asked: On what basis can we be assured that our current 'rational' beliefs are in fact true? My answer is that it is rational to assert them to be true since they are epistemically justified relative to the evidence. This usage is also in accord with our intuitive understanding of these terms, except in so far as that we may want to say that not all of our present rational beliefs constitute knowledge. We may want to say that a rational belief constitutes knowledge only if the evidence in support of the belief exceeds a minimum specified strength. We can only begin to solve this problem by first agreeing to a set of criteria for rationally appraising beliefs. I shall now attempt to formulate such a set.

The formulation of such principles is no easy matter, for what we appraise are beliefs about sentences on a number of different theoretical levels. We appraise observation sentences, such as 'You sped through that "Stop" sign and hit my car.' On the next level, we judge causal hypotheses, such as 'Smoking has a tendency to cause lung cancer.' More general hypotheses that explain a process in terms of some underlying microstructure or overlying macrostructure are also appraised. The operator gene theory of cancer and the atomic theory of chemical affinity are examples of the former, while Marxist analysis of revolutions in terms of social institutions and Einstein's explanation of the motions of celestial bodies as being a function of the topology of space are examples of the latter. In appraising these high level theories, it is clear that we are not judging isolated sentences, but a complex set of interconnected sentences, the members of which change with the development of the theory.

On an even higher level, we judge entire belief-systems that include a cosmology, ontology and epistemology in one package deal. In these high-level systems, these elements are interrelated to form one systematic whole and to replace one element necessarily involves modifying the others. Vedanta Hinduism, traditional Aboriginal religion and the Newtonian dualist–mechanistic empiricist world view are just such belief-systems.

This introduces a further complication. The details of an objectivist theory of rationality are a function of the ontology-cosmology that we choose to adopt. This is partly for the simple reason that our theories of observation will depend upon the way we think the world is structured and the way we think we are built. Conversely, details of our ontology-cosmology will be a function of how we think we rationally appraise theories.

The solution to these difficulties can be found if there are certain fundamental objective criteria for the appraisal of ontology-cosmologies independent of such theories. I think that there are such criteria that follow from a general objectivist epistemology and, furthermore, the application of such criteria justifies a realist framework within which it is possible to further develop an objectivist theory of rationality. In the remainder of this essay, I shall explicate such fundamental criteria, briefly argue that the application of such criteria favours a realist framework and then go on to develop further a theory of rationality within such a framework.

4.2 Criteria for Minimal Acceptance of Evidence

What are such fundamental criteria for rational theory appraisal? Every epistemology stipulates that the evidence *qua* evidence must be epistemically related to the theory under evaluation. The epistemic status of the evidence-statement must have some consequence for the epistemic status of the theory. In an objectivist theory of rationality, the evidence must partially determine whether the sentence or theory has the property of stating what is actually the case. Statements about the subjective states of people do not satisfy this requirement and so are not regarded as evidence. Debarred statements include statements that a person or a group of people:

- believe/disbelieve the sentence or theory,
- approve/disapprove of the political/ethical implications of the theory being believed by the general populace,
- prefer the theory to be true/false,
- consider acceptance of the theory more pragmatic,
- consider the theory more aesthetic or more simple.

The above condition is the minimal requirement of independence; that a statement be independent of subjective states for it to be regarded as evidence. It should be clear that this requirement follows from the general requirements of an objectivist epistemology outlined at the start of this section.

That a statement be minimally independent is a necessary but not a sufficient condition for it to be regarded as evidence. As I said above, the statement must necessarily be epistemically related to the sentence or theory. This is satisfied by a requirement of logical entailment. Take the case of an explanatory theory T and auxiliary hypotheses A. (I include here auxiliary hypotheses A because explanatory theories do not normally entail evidence-statements in isolation. In this case, the evidence-statement e cannot be logically entailed by the auxiliary hypotheses alone.) If a theory T and acceptable auxiliary hypotheses A logically entail evidence-statement e, and e is false, then e is evidence against T.[8] If e is true, then e is evidence for T.

We may now state the necessary and sufficient conditions for a statement to be evidence for a theory. A statement is evidence for a theory iff:

Criterion 1: the statement is true;
(*requirement of dependence on epistemic value*)

Criterion 2: the statement is logically entailed by the theory in conjunction with accepted auxiliary hypotheses, but not by the auxiliary hypotheses alone;
(*requirement of dependence on logical deducibility*)

Criterion 3: the statement is not a statement of the relationship between the theory and a subjective state or states.
(*requirement of independence from psychological states*)

[8] 'N' $=_{df}$ 'it is logically true that'. $N(N((T\&A) \rightarrow e)\&(-e\&A)) \rightarrow -T)$ in the limit where A and $-e$ are certain.

4.3 Criteria for Theory-Independence

There is an important caveat here. A theory may have evidence in its favour in the sense specified by the three minimal requirements introduced above. However, we should not accept it as true on that basis alone. This is for two reasons. Firstly, the argument for the theory will simply beg the question if the truth of the evidence-statement is determined by presupposing the truth of the theory under consideration. Consider, for example, the simple cosmology, 'Every entity is causally related to every other entity', and the evidence proposed in support of it, 'Percepts are causally related to every other entity'. If the truth of the evidence-statement was determined by deriving it from the statements, 'Every entity is causally related to every other entity' and 'Percepts are entities', then the argument is viciously circular. Cosmologies logically inconsistent with the above, such as, 'No entity is causally related to any other entity', can just as easily be supported by such circular arguments. What is required is a condition of theory appraisal that specifies that the epistemic status of the evidence-statement be determinable independently of the theory under consideration.

Secondly, for any given body of data, any number of mutually inconsistent explanatory theories, conjoined with auxiliary hypotheses and entailing the data, may be adduced. This is the well-known problem of the underdetermination of theory choice.[9] On a superficial analysis, it will appear that the data equally constitute evidence for all of these mutually inconsistent theories. In addition, the appearance of new data inconsistent with a theory and its auxiliary hypotheses can always be accommodated by *post hoc* revisions to the theory or to the auxiliary hypotheses. The anomalous data itself can be modified by means of a newly introduced auxiliary theory explaining how the data is to be revised.[10] (Ontology-cosmologies are so all-encompassing that I shall henceforth refer to the metaphysical core of the theory, the auxiliary hypotheses and any newly introduced auxiliary theories as all part of the one ontological-cosmological theory.)

In order to evaluate such a plethora of seemingly equally evidenced theories, an added criterion for objective theory appraisal is required. It must stipulate that the evidence be independent of that which equally supports all of these mutually inconsistent theories. That is, the evidence must be independent of the data used to construct the original theory and used in its later revisions.

Now, evidence-statements that are independent of theory construction and revision may be classified as belonging to either of two mutually exclusive categories. The evidence-statement is either a 'novel prediction' or a 'novelly derived fact', where a 'fact' is a true contingent sentence. I shall mean by 'novel prediction' (relative to a particular theoretical development) a statement that is derivable from the theory but was not known to be true by the constructors and revisers of the theory during construction and revision.[11]

[9]See, for example Quine and Ullian [1970] and Stanford [2013].

[10]For simple examples of *post hoc* revisions, see Lakatos [1978a: 16f, 97f].

[11]This is not exactly the same as Lakatos' use of the term 'novel prediction'. For Lakatos, a prediction is not novel if it is also predicted by a rival theory. See Lakatos [1978a: 32, 1978b: 170f]. Contrary to Lakatos, 'novelty' is relative to people and not to pairs of theories. For example, if a new theory predicts a novel fact of which we were hitherto unaware, and it is later discovered that the older rival theory had also, unbeknownst to us, predicted the same fact, it would be more natural to continue describing the newly discovered fact as

And by 'novelly derived fact' (relative to a particular theoretical development), I mean a statement that is derivable from the theory and was known to be true to the constructors or revisers of the theory during construction or revision, but was not used to construct or revise the theory.[12]

This composite class of evidence-statements that are independent of theory construction and revision (that is, the class of novel predictions and novelly derived facts) may be divided into three sub-classes:

i) those statements that are novel predictions or novelly derived facts for one or more rival theories;

ii) those statements that are derivable from one or more rival theories but are neither novel predictions nor novelly derived facts for such rivals;

iii) those statements that are either not derivable, or their negations derivable, from one or more rival theories.

If an evidence-statement for a theory is of type i), the statement is equally strong evidence for all rival theories for which it is also a novel prediction or a novelly derived fact. If it is of type ii), the statement is strong evidence for the theory and weak evidence for the rivals from which it can be derived. If it is of type iii), it is once again strong evidence for the theory, but is either no evidence or counterevidence for the rivals. Type iii) evidence-statements give the strongest *relative* weight to a theory because statements of this type can explain what the rivals cannot explain. Type ii) evidence-statements give a weaker, but still very significant, relative weight to the theory. This is because statements of this type explain independently that which the rivals can only explain dependently. Type i) evidence-statements give no relative weight to the theory in comparison with rivals that also give an independent explanation.

Recognising the crucial epistemic value of novel predictions and novelly derived facts helps to solve the problem of the underdetermination of theory choice. We are now in a position to supplement my three previously stated criteria for rational theory choice with two new conditions. We may roughly formalize these as follows.

An evidence-statement is *strong* evidence for a theory if:

Criterion 4: the epistemic status of the evidence-statement is determinable independently of the theory;
(*requirement of epistemic independence from theory*)

Criterion 5: the evidence-statement is independent of the data used to construct and modify the theory.
(*requirement of independence from theory development*)

'novel', even though it could also be deduced from the rival theory. Fortunately, in the end, this terminological difference makes no difference to the theory of theory appraisal advocated here.

[12]This is akin to Zahar's concept of 'novel fact'. See Zahar [1973] and Lakatos [1978a: ch. 4].

This completes my formulation of the fundamental criteria for rational theory choice. *Criterion 1*, *2* and *3* specify the necessary and sufficient conditions that must be satisfied for a statement to be regarded as evidence at all. That a theory be supported by evidence satisfying these three minimal conditions does not warrant automatic rational assent to its truth. As mentioned earlier, it is not too difficult to construct theories logically inconsistent with the one in mind that are equally supported by such evidence. A theory is rationally justified, then, only if it is supported by evidence of the type stipulated in *Criterion 4* and *5*. Even then, the theory in question is supported only if there is no rival theory supported to a greater extent, or nearly to the same extent, by evidence of this type. With these five criteria for objective theory choice, we have here the basic ingredients of Lakatos' MSRP (with Zahar's important modification), with its emphasis on the epistemic value of novel facts. Quite neatly, Lakatos' basic schema turns out to be a consequence of the demands of an objectivist epistemology.

5. Metaphysics and Epistemology

The specific details of an adequate theory of rationality are dependent on our metaphysical presuppositions; more particularly, on the ontological and cosmological commitments we make. How we infer truths about our internal states and the external world, for example, depends on our notions of phenomenal experience, mind and matter and how we think they interact. Our choice of metaphysics is, therefore, inextricably bound up with our theory of rationality, and vice versa. What ontology-cosmology should an objectivist assent to? The two great traditions, idealism and realism, have had a long and tortuous history dating back more than two thousand years. There have been many enlightening and interesting developments resulting from the interaction between the protagonists for each side. I will not discuss these here.[13] However, in this section, I do want to explore what I think is a new turn in the argument provided by the type of theory of rationality developed by Lakatos and his school and defended here.

I shall characterize idealism and realism as two broad streams of thought offering contrary theses; realism on the one hand asserting the existence of mind-independent entities, knowledge of whose properties is afforded access by sense-experience, and idealism on the other denying the existence of such mind-independent entities and the consummate role of sense-experience. Realist and idealist theories have had to contend with a number of serious problems, especially those of logical coherence and meaning. Whereas realism has progressed, though, idealism has faltered. Idealists had sought to provide an idealist analysis of natural languages that would eliminate realist assumptions. (Phenomenalists became engaged in their own programme of elimination.) They have been largely unsuccessful, for natural languages appear to be inherently realist.

There is a second serious impediment to accepting an idealist framework. Recent developments in neurophysiology and psychology indicate that the act of perceiving physical objects is not simply a matter of conveniently grouping discrete units in our visual field into particular packages. Our perception of solidity, depth and contour are not arrangements of the elements of our visual field, but the result of complex neural processing in the brain. In perceiving three-dimensional solid objects, there occur not two discrete acts, one of the awareness of the discrete elements in the visual field and the other of cognitive interpretation, but the one indissoluble act of seeing a three-dimensional object. An example of this is our perception of depth in the reversible perspective drawings of cubes, staircases and other objects given in textbooks on perception.[14]

This act of seeing solid objects in three-dimensional space is only partly learned. The form of neural processing of visual information that goes on in our brains is also partly genetically determined. Our brains are hard-wired for us to perceive depth and contour the way we do.[15] The Ames illusion[16] is a classic demonstration of how our brains are

[13] See my Allan [2016a] for a critical discussion of realism and idealism.

[14] For examples of reversible/bistable images, see Thomson [2009a, 2009b], Kogo [2012] and Nordhjem [2012].

[15] See Vernon [1971: ch. 8] for a general discussion. For details on the genetic component contributing to our perception of depth and contour, see the important report by Julesz [1964]. For experimental reports on neural processing in the visual cortex in higher vertebrates, see Hubel and Wiesel [1962, 1968].

[16] For a visual demonstration, see Scientific American [2012].

hard-wired to see a trapezoidal-shaped room as symmetrical when the usual visual cues are hidden. Another well-known example is the rotating mask illusion.[17] A convex mask of a human face is rotated. As the inner-side of the mask comes into view, we see the mask as convex. This is because we have an inbuilt predisposition to believe that noses always point outwards on human faces. Because of this inextricable and unconscious fusion of our reception of sensory data and its interpretation, it is little wonder that idealists have been unable to construct a non-realist interpretation of natural languages.

This illuminates a third major obstacle for idealist theories. Realist assumptions are necessary for the living of our everyday lives. When a mother picks up her child, for example, she does not infer the expected position and look of her hands and her child's body as bundles of actual and possible sense-experiences. To calculate such inferences given the variability and uniqueness of much of our sensory world is a psychologically impossible task. Similarly, in everyday life, the idealist finds it impossible to regard other minds as simply bundles of actual and dispositional behaviours. Just as the idealists' gestalt sees physical bodies imbued with minds (that is, they see living human forms as persons and not as bundles of actual and dispositional behaviours), our gestalt sees external objects located and moving in three-dimensional space, with no conscious inference from discrete sense-impressions. Modern neurophysiology and psychology are revealing for us how this is so.

Now, an idealist may concede these points, claiming, however, that the presuppositions of a realistically interpreted natural language are no argument for the veracity of such unconscious assumptions. He may explain *post hoc* the latest developments in psychology and neurophysiology. In the final analysis, he may admit that it is useful to act as a realist would act, while consistently adhering to his idealism. I have little doubt that a coherent version of idealism can be developed and modified *post hoc* to account for any current and future developments in science and any possible configuration of our perceptual experiences. That such an idealist theory could explain our experiences is not strong evidence for its veracity. To earn our rational assent, according to *Criterion 5*, the idealist theory must have evidence in its favour independent of its construction and modification.

Realism, on the other hand, has overwhelming independent evidence in its support. The realist interpretation of physical and neurophysiological theories has led to the derivation and confirmation of novel phenomena; that is, phenomena that had played no part in the original construction and modification of the theory. Those theories that had independent confirmation, thereby satisfying my criteria for rational appraisal, confirm in a spectacular way the realist ontology-cosmology.

Consider, for example, Newton's Three Laws of Motion and his Universal Law of Gravitation. These two theories, interpreted realistically (in conjunction with a realist interpretation of the auxiliary hypotheses), were independently confirmed by the successful predictions of the existence and position of Neptune, the date of the return of Halley's comet and the novel derivation of the measure of the progress of the moon's apogee. These spectacular successes of Newtonian mechanics also simultaneously supported its

[17]For a visual demonstration, see eChalk Education [2012].

ontological and cosmological underpinnings; the idea that bodies possessing the properties related in Newton's axioms exist independently of minds, in absolute space and time.

The same can be said for the confirmations of Einstein's Special and General Theories of Relativity, the Rutherford-Bohr atomic theory, the kinetic theory of gases, the synthetic theory of evolution and the modern neurophysiological theory of perception. These examples are but a very small fraction of the empirical growth in our understanding of the universe prompted by advances in scientific research. My key point here is that each of these theories, realistically interpreted, made bold novel predictions that were subsequently stunningly confirmed by independent observers. Conversely, idealism has had little or no empirical confirmations that satisfy my criteria for objective theory appraisal. It is only able to explain the realist's anticipation of novel phenomena and his novel derivation of statements about well-known phenomena *post hoc*.

Taking a step back to fundamentals, we can say that the data for which realism and idealism seek an explanation are those private phenomenological experiences that we refer to as 'sense-experiences'.[18] These 'sense-experiences' for the idealist and the realist, equally, are epistemically primitive in the sense that our own private phenomenological experiences are what each of us have most directly. (Even though they are the most direct, they are not necessarily infallibly known.) The upshot here is that realistically interpreted ontology-cosmologies lead to the successful prediction of new and unexpected, and the novel derivation of previously well-known, sense-experiences. Idealism has been, in comparison, largely unsuccessful in the anticipation and novel derivation of such experiences.

The same type of argument may be successfully adduced against the instrumentalist interpretation of explanatory theories in science. It can also be used to elucidate how each of us as children comes to a realist interpretation of our phenomenal world.[19] The argument against idealism and in support of a belief in mind-independent entities is a substantial one. I shall now go on to further develop the theory of rationality, proposed above, within such a realist framework.

[18]The contents of our subjective experiences of observed objects have been variously termed 'sense-impressions', 'sense-data' and 'qualia' by philosophers.

[19]For an exploration of Piaget's and other developmental psychologists' research on how children develop the concept of an external physical world, see my Allan [2016a].

6. A Realist Theory of Knowledge

6.1 Theory-Ladenness of Observation Language

That we gain knowledge of an external mind-independent world *via* our senses is a fundamental axiom of realism. However, this view raises important questions about how our senses function in the acquisition of such knowledge and what methods we use to improve our knowledge. The early empiricist view that by an unbiased use of our sensory organs we may gain direct knowledge about the external world is no longer tenable for a number of reasons. These objections are mostly well covered in the literature, so I shall discuss briefly the two that are most pertinent to developing a theory of rationality. In the process, I will demonstrate the methodological implications for such a theory. The two complications to the empiricist's story of how we gain knowledge of the external world are the theory-ladenness of observation language and the perceptual effects from subjective influences. I will deal with theory-ladenness in this section and leave the discussion about perceptual bias to the following section.

To assert an observation statement about the external world is to assert much more than that the observer is having a particular private sensory-experience. It is to interpret the phenomenological experience in terms of a number of theories. To assert even simple sentences as 'The cat is on the mat' or 'I see a cat on the mat' is to presuppose physical theories concerning the properties of objects called 'cats' and 'mats' and the causal interaction between these physical properties. Whereas in Europe during the middle-ages, people saw some women as witches and some afflicted with a mental illness as victims of demon possession, today we see these same people as quite ordinary women and as people suffering from psychosis. As these examples illustrate, our observation statements are permeated with a host of theoretical presuppositions.

To accept an observation statement as true is also to presuppose some, perhaps relatively inarticulate, theory of observation of how we come to know about the existence of cats, mats, women and the mentally ill and of their properties. If we accept the modern, scientific account of how we visually perceive objects, we will tell the story about how our visual experience of a cat is the result of light waves of particular frequencies and intensities reflecting off the cat, those waves focusing in our retina, electrical signals travelling along our visual cortex to the visual processing areas in our brain and finally resulting in our experiencing a visual image of a cat.

It is in this sense, in which observation terms gain their meaning by being interpreted in terms of a physical theory and in which assertions of observation statements presuppose an observational theory, that we can say that observation terms and statements are 'theory-laden'. The once hoped for theory-neutral observation language on which we could provide a secure foundation for our theories is now seen as a chimera.

The upshot here is that our observation statements are only as precise as the physical and observation theories that they presuppose. In those cases where an observation statement made in support of a new theory presupposes an unsupported observation theory or some other unverified auxiliary hypothesis, prudence demands that we do not accept the observation statement as true. In the following, I will incorporate this lesson into my set of criteria for rational theory appraisal.

6.2 Perceptual Effects from Subjective Influences

Turning now to the second complication for an empiricist view of observation, scientific research reveals that what we perceive visually is not simply a two-dimensional retinal image. What we see is in fact the result of complex neurophysiological processing of the visual information that enters our eyes mediated by our current theoretical construct of the world. Cognitive science reveals that there are ubiquitous active top-down feedback networks that impact the bottom-up processing of our sensory inputs.[20] The nature of this processing means that our phenomenological experience is not only a function of the image cast on our retina, but also of our genetic constitution,[21] our prior beliefs and our expectations. Our expectations serve to influence our perception of colour, shape, speed of motion, brightness and distance, among other things.[22]

For example, in one study, Jones and Bruner [1954] showed experimental subjects a sequence of cards depicting a 'stick-man' running and a nonsense object moving. The observers reported perceiving the 'stick-man' moving faster and travelling a greater distance than the nonsense object, even though in actual fact they had both traversed the same distance in the same amount of time. The running 'stick-man' was seen to move faster and further because it was expected to move faster than an object that had no significance for the experimental observers.

Such expectations and, consequently, perceptions can also be influenced by suggestion. In another study [Hastorf 1950], experimental subjects perceived the same white rectangle to be at a further distance from them when they were informed that it was an envelope compared with when they were told that it was a calling card. Similarly, a white circle was perceived as being located further away when it was suggested that it was a billiard ball compared with when it was suggested that it was a ping pong ball.[23]

These experiments demonstrate that our expectations are dependent on our prior experiences and beliefs. However, our prior experiences and beliefs can also shape our perceptual experiences in the absence of accompanying expectations. For example, Steinfeld [1968] demonstrated how a group of experimental subjects having been first read a story about a ship more quickly and easily recognized a fragmented drawing of a ship compared with an otherwise similar group that had not.[24]

Research studies have also suggested that in some circumstances our perceptions are partly shaped by our needs, values and personalities.[25] Not only are our perceptual

[20]For an overview of research on top-down processing of stimuli, see Kveraga *et al* [2007] and Bubic *et al* [2010].

[21]See §5 above for the role played by the hard-wiring in our visual cortex in our perception of depth and contour.

[22]There is now a wealth of literature on this subject. For a reasonably comprehensive survey of the literature, see Vernon [1971]. For a selection of experimental reports, see the *Psychology of Perception* section under *References* at the end of this essay.

[23]For other studies on the influence of suggestion on perception, see, for example, Landauer and Rodger [1964] and McGee [1963].

[24]See also, for example, Borresen and Lichte [1962] and Proshansky and Murphy [1942].

[25]See Vernon [1971: chs 10, 11], Comalli Jr [1960], Levine, Chein and Murphy [1942], Smith, Parker and Robinson Jr [1951] and Tajfel [1957].

experiences influenced by our internal subjective states, but our recollections of our perceptual experiences can also be distorted by our previous beliefs and later experiences. For example, many studies have confirmed that our memories can be easily modified by later suggestions.[26] It is not only under abnormal conditions that our perceptual experiences and memories are shaped by our inner subjective states. Some of these experimental studies indicate that this happens under normal conditions as well.

The conclusion we must draw from these research studies is that we need to treat observational evidence presented in the form of personal anecdotes with due caution. We have already noted at the beginning of this section that observation statements about the external world are interpreted in the light of particular theories expressible in a public language about physical objects. We can also now add the insight that these same observation statements are based on phenomenological sense-experiences that can be easily distorted by the observer's psychological state at the time and on memories that can be shaped by later psychological states.

[26]See, for example, Loftus [1979].

6.3 Criteria for Accepting Observation Statements

Observation statements play a key role in building and testing our knowledge of the world and of ourselves. They often function as evidence for our explanatory theories. *Criterion 1* in my theory of knowledge already stipulates that a statement is acceptable as evidence for a theory only if it satisfies the condition that it is true. What are the epistemological implications for my theory of rationality given the theory-ladenness of observation statements on the one hand and the propensity for perceptual bias from subjective factors on the other?

Given the above considerations, I can now add two specific conditions for the acceptance of an observation statement as evidence for a theory. They are as follows.

The conditions for an observation statement to be evidence for a theory are:

Criterion 6: the statement is interpreted in terms of accepted physical and observational theories;
(*dependence on accepted presuppositions*)

Criterion 7: the statement is not the result of distorting subjective influences, such as prior beliefs and experiences, expectations, needs, values and personality.
(*requirement of independence from bias*)

These two new criteria serve to fill out the picture of my theory of knowledge within a realist framework. They further explicate my *Criterion 1;* the condition that a statement tendered as evidential support for a theory be true. *Criterion 6* ensures that the observation statement does not presuppose an unsupported observational theory or other such auxiliary hypothesis. *Criterion 7*, on the other hand, serves to eliminate perceptual biases from the observer's psychological state or viewing conditions.

How do these two criteria operate in practice? Two examples illustrating the application of *Criterion 6* follow. The first example concerns an observer who reports that the earth did not move as he jumped in the air. The observation report is proposed in support of the Ptolemaic-Aristotelian cosmology with a fixed earth at the centre of the cosmos. The application of *Criterion 6* disallows this observation report as the report presupposes the false theory of dynamics that entails that a person will not move unless acted upon by a force.

In the second example, an observer attends a public 'psychic' demonstration after which he reports that the 'psychic' on show had supernaturally bent his car key. Applying *Criterion 6*, this observation report is rejected as evidence for psychic abilities because it presupposes the false theory that untrained observers are able to follow the hand movements of people versed in the skills of misdirection.

The purpose of *Criterion 6* is not to prevent the acceptance of a new physical or observational theory. Our current theories are always open to revision in the light of further evidence. However, if an observation statement relies for its veracity on a new, untested theory (that is, on an unverified auxiliary hypothesis), then the observation statement in question cannot be accepted until the untested auxiliary hypothesis on which it relies has

undergone and passed independent tests. In this sense, *Criterion 6* formalises a degree of epistemic conservatism. We do not discard our current established theories on the basis of an anomalous observation statement until the rival theoretical assumptions grounding the observation statement are themselves tested and verified.

What follows next are two examples of the application of *Criterion 7*; the requirement of independence from subjective influences. For the first example, consider an observational report of a U.F.O. sighting from the night before. Subsequent investigation of the report reveals that the observer had an avid prior interest in U.F.Os. The application of *Criterion 7* methodologically excludes the acceptance of the observation statement as it cannot be ruled out that the observer was influenced in his perception by his prior belief in the existence of extra-terrestrial visitors.

In the second example, a report in a widely circulated newspaper of a sighting of the Yeti is followed by a large number of similar sightings. An analysis of the timing of the reports shows that there were no reported sightings between the time of the original sighting and its publication in the newspaper. Applying *Criterion 7* to this case again rules out the acceptance of the later observation reports as it is possible that these reports are the result of the increased expectation by the local inhabitants of seeing the Yeti generated by the newspaper report itself.

It is important to note that additional evidence may come to light that increases the veracity of the two reports in the examples above. For example, reports by independent observers with no propensity for belief in the paranormal may surface, along with physical evidence on the ground of extra-terrestrials or the Yeti. The point of *Criterion 7* is not to rule out all observational reports. Its purpose is to exclude only those for which there is a possibility that the report offered is tainted with one or more of the subjective biases discussed above.

Psychological considerations such as those that underpin *Criterion 7* above have led researchers to develop and refine sophisticated test protocols. The most rigorous of these, now commonplace in many research labs, is controlled double-blind testing. Through devising and using such procedures during the design and test phases, extraneous external causal influences and the subjective biases of the experimental subjects and researchers can be identified and eliminated.

In tests employing double-blind procedures, the value of one variable is under experimental control, with the experimental subjects being kept uninformed (blind) as to the value of that variable in order to eliminate or later isolate distorting psychological influences on the part of the subject. The judges of the value of a second variable are also kept uninformed (blind) as to the value of the first variable, once again, in order to eliminate their personal biases. How the control of the first variable is established will vary with the type of theory under test. Such details are not important here. However, what is important is that this type of procedure is a natural development not only from the psychological considerations discussed above, but is also necessitated by my general *Criterion 1* to *Criterion 5* if knowledge is to be systematically developed.

Controlled double-blind testing works to test theories using objective criteria not by testing a single theory in isolation, but by testing it against a second mutually inconsistent

theory. The first theory is of a type that postulates that a variable x has a statistically significant causal effect on another variable y. An example of this type of theory is one that states that there is a causal connection between the chemical properties of a particular drug and the disappearance of symptoms in a group of diseased patients. The second theory under test denies that there is such a causal relationship between the two variables.

During the test design phase, the logical relationships between the two theories and the probability distributions of test results are calculated (*Criterion 1* to *Criterion 4*). By holding all other known causal factors constant, or randomly distributing them throughout the samples, novel predictions (*Criterion 5*) concerning the statistical significance of the correlations between the two variables are made for each theory. The test is then performed. If a statistically significant result eventuates, this is counted as independent objective evidence for the first (causal) theory and counter-evidence for the second (non-causal) theory. If the ensuing result is not statistically significant, this is regarded as evidence against the first theory and evidence for the second.

So far, I have discussed the use of double-blind protocols only as they relate to observation statements made about the external world; that is, about the world of mind-independent physical objects and their properties. In this context, double-blind procedures are used extensively. A couple of further examples are their use in the analysis of data in particle physics experiments and in police photo lineups of crime suspects.

I want now to extend the application of *Criterion 7* to the domain of private, introspective reports. Just as observers report events in their external world, they also report their internal mental states, such as feelings of relaxation and sensations of pain. Unlike observations about external objects, investigators are not privy to checking directly the observations of an individual's private mental states. The challenge arises in that such private, introspective reports are used to test theories in a variety of domains. The testing of new medicines and studies on human perception, for example, regularly rely on such introspective reports about subjects' phenomenological experiences.

Here again, controlled double-blind test protocols serve to eliminate distorting influences from the test setup itself. Biases can be introduced by the subject's or the tester's belief or hope that the theory under test is true (placebo effect) or by the extra attention that the researchers pay the subject (Hawthorne effect), or both. As with observation statements about external physical entities, introspective reports distorted by environmental or psychological factors cannot be accepted as support for a theory undergoing critical appraisal. The upshot here is that our use of *Criterion 7* in barring observation statements that are the result of such distorting subjective influences applies to both types of observation statements. This criterion applies to observer reports about external physical entities as well as reports about the observer's internal mental states.

I have now discussed the criteria for the acceptance of observation statements tendered in support of explanatory theories. But what of evidence-statements that are not observation reports as such? These kinds of evidential support include statements about the past and future, such as the statement that an asteroid impacted the earth 65 million years ago and the statement that our galaxy will collide with the Andromeda Galaxy in another

four billion years. It also includes statements about objects and properties that cannot be observed, such as the statement that the temperature of the Earth's inner core is 5400 °C. How are evidence-statements such as these rationally acceptable when used in support of other theories? In a nutshell, they earn that honour because they are deducible from theories satisfying acceptance *Criterion 1* to *5* in conjunction with statements of initial conditions.

In fact, the complex of accepted highest-level theories (which some philosophers refer to as a 'paradigm' or 'world view') serves as the touchstone for determining what constitutes a 'rational reason' for believing a lower-level theory or a statement about the future or past or an unobserved event. It provides the 'rules of evidence', stipulating what type of statement can or cannot be regarded as evidence for such. Within the twenty-first century scientific paradigm, for example, the particular details of a newly-born child's horoscope is not allowable as evidence for statements about his future prospects. Similarly, that a person is left-handed is not admitted as the type of evidence required to demonstrate that he is possessed by a demon. Alternatively, within a medieval paradigm that incorporates an astrological cosmology and a demon possession theory of abnormal behaviour, such statements are accepted as rationally relevant to the appraisal of the statements in question.

However, we must not confuse these intra-paradigm standards of rationality with the extra-paradigm standards defended in this essay. We may, for example, criticise the New Age standard of rationality that stipulates that horoscope readings are rationally relevant to the determining of people's personalities by criticizing the astrological theories in which this standard is embedded. It is this confusion between intra-paradigm and extra-paradigm standards of rationality that has led Barnes and Bloor [1982] and Feyerabend [1978: 22f, 28–30, 32–4, 81–3] to lapse sometimes into a radical version of relativism that maintains that all standards of rationality are culture dependent.

7. Objective Criteria and MSRP

This completes my development from the general demands of an objectivist epistemology rational criteria for explanatory theory and sentence appraisal. These criteria may now be summarily presented.

The conditions for a statement to be evidence for a theory are:

Criterion 1: the statement is true;
(*requirement of dependence on epistemic value*)

Criterion 2: the statement is logically entailed by the theory in conjunction with accepted auxiliary hypotheses, but not by the auxiliary hypotheses alone;
(*requirement of dependence on logical deducibility*)

Criterion 3: the statement is not a statement of the relationship between the theory and a subjective state or states;
(*requirement of independence from psychological states*)

Criterion 4: the epistemic status of the evidence-statement is determinable independently of the theory;
(*requirement of epistemic independence from theory*)

Criterion 5: the evidence-statement is independent of the data used to construct and modify the theory;
(*requirement of independence from theory development*)

Criterion 6: the statement is interpreted in terms of accepted physical and observational theories;
(*dependence on accepted presuppositions*)

Criterion 7: the statement is not the result of distorting subjective influences, such as prior beliefs and experiences, expectations, needs, values and personality.
(*requirement of independence from bias*)

Criterion 1, *2* and *3* state the minimal conditions for a statement to be regarded as evidence for a theory. Statements satisfying *Criterion 4* and *5* are regarded as *strong* evidence for a theory. *Criterion 6* and *7* apply specifically to observation statements about the external world, with *Criterion 7* extending the scope to include introspective psychological reports.

As can be seen, I have tried to isolate and clarify both the various dependence and independence requirements necessary for an adequate objectivist epistemology. Some of these are common to other epistemologies. The dependence requirements specified in *Criterion 1* and *2* are seen in all sophisticated hypothetico-deductive systems. Less sophisticated systems assume, wrongly, that scientific theories entail predictions in isolation. The independence requirements stipulated in *Criterion 3* and *4* are usually taken

for granted, while that in *Criterion 7* is only recognized by epistemologies that incorporate a more sophisticated theory of observation than the camera lens analogy.

Lakatos had repeatedly emphasised that there is no theory-neutral observation base and that observational theories are revisable.[27] This is a fundamental tenet of his MSRP and is reflected in my *Criterion 7*. However, Lakatos had little cause to specify this condition separately and in the form that I have stated it here, for he had focused his attention almost entirely on the history of physics. The historical record shows physicists deciding the worth of rival theories, in large part, by their ability to predict novel phenomena within physical systems. Such phenomena were characterised by being open to scrutiny by the scientific community and interested public. Any theory that revealed phenomena that were only 'observable' to the protagonists of a particular theory, and not to the rivals and disinterested parties, was not taken seriously and was quickly forgotten.[28] The situation is more ambiguous in the case of the social sciences, such as psychology and sociology. In these disciplines, the primary evidence for rival theories is the psychological responses of human beings and these may be easily distorted by inadequate testing procedures, as I had explained in §6.3 above.

The 'requirement of independence from theory development', *Criterion 5*, is reflected in Lakatos' and Zahar's stipulation that for a series of theories (or, in my terminology, successive versions of the same theory) to be rationally justifiable, it must successfully predict novel facts or novelly derive known facts. This historical view of confirmation forming the basis of Lakatos' MSRP specifies that we cannot judge the epistemic relationship between a theory and evidence-statements in isolation, but must know the prior history of the theory. In other words, for an evidence-statement to support a theory, it must not have been used by its proponents in the construction and revision of the theory. The above considerations of the general demands of an objectivist theory of knowledge appear to vindicate Lakatos and Zahar's historical approach to theories of rationality.

The epistemology formulated here, as with Lakatos' parallel, is universally fallibilist about empirical knowledge. There is no immediately known observational knowledge and no indubitable synthetic *a priori* principles. Even realism is not on absolutely solid foundations, for it may turn out that those phenomenological experiences that we label 'sense-experiences' have been misdescribed, or even that they never existed. It may seem that this fallibilist epistemology can never be applied because we can never get to use it. Observation statements, to be acceptable in testing a theory, must be interpreted in terms of accepted physical and observational theories. However, such physical and observational theories require evidence in the form of observation statements to be deemed adequate. So, it may seem that the evidence-theory-evidence circle can never be broken long enough for us to enter into. I think there are two ways to answer this quandary; the philosophical and the psychobiological.

Dealing with the philosophical first, it seems epistemically prudent simply to jump into the circle at any point. Just as we tentatively accept the existence of those experiences that we label 'sense-experiences' in order to enter the debate between idealists and

[27] See, for example, Lakatos [1978a: 14–16, 23, 45f].

[28] For some amusing examples, see Martin Gardner [1957: ch. 10].

realists, we also tentatively accept some relatively inarticulate and undeveloped observational, psychological and physical theories to start the enterprise of theory improvement. We may then improve a theory by using our tentative theories as a support structure in order to make and test novel predictions. We only gain knowledge by initially pitching our tent somewhere and then investigating where we can cut a trail into the unknown using our epistemological exploration tools. In this way, we systematically improve our knowledge of the world, testing one domain at a time.

With this philosophical response to our predicament, we can either accept the invitation to set up a base camp or not. We come to the second approach to this puzzle by realising that we have no choice. Our biological evolution ensures that we do enter the evidence-theory-evidence circle at a predetermined point. It appears that we do not act, according to this evolutionary view of theory improvement, by a conscious volition. In §5 above, I outlined how research in neurophysiology and psychology show that our brains are hard-wired to recognise a world of external physical objects from a very early age. Our perceptual experiences and beliefs are shaped to see and accept mind-independent objects behaving in law-like ways. Realism most likely evolved as a privileged hypothesis during our evolutionary development because it offered survival advantage.[29] In a predatory environment, our ancestors were more likely to pass on their genes if they had the cognitive ability to quickly recognise a running tiger.

With the expectation of how objects move and behave hard-wired in our brains, current research also shows that we do not give up these expectations lightly in the face of seeming counterexamples. Studies in neurophysiology and psychology suggest that our propensity to revise our beliefs is tempered by a measure of epistemic conservatism. Situations where we are met with perceptual ambiguity are excellent illustrations of this kind of conservation of belief in the face of apparent refutations. For example, consider the illusions I raised in §5 above. When we encounter the trapezoidal room in the Ames illusion, we don't give up our belief that rooms are symmetrical. Similarly, when we see the inside of the convex face mask in the rotating mask illusion, we don't abandon our belief that noses point outwards on human faces.

Another example of perceptual incongruence is the McGurk effect,[30] where we hear the sound 'Fah' in synchronism with a person's lips appearing to make that sound, even though the actual sound emanating from the person is 'Bah'. One last example I'll draw your attention to is the vanishing ball illusion[31] in which a ball appears to leave a person's hand even though it was not actually thrown into the air. In these last two examples, we do not give up on the belief that sounds invariably match the speaker's lip movements and that balls continue to persist in space and time. What is instructive about these illusions is not that we consciously refuse to update our beliefs when faced with these types of counterexamples. It's that our brain conspires to hide the counterinstances from our conscious awareness.

No doubt, this epistemic conservatism also proved evolutionarily advantageous. Abandoning a well-established construct of the external world after a seeming

[29]See, for example, Kveraga *et al* [2007: 147].
[30]For a visual demonstration, see BBC [2010].
[31]For a visual demonstration, see Sunderland Echo [2008].

counterexample—to be left with no model of reality at all—would have cursed our ancestors to the next predator that happened to cross their path. There is a corollary here to naïve versions of Popper's falsificationism. Abandoning an established scientific theory at the first sign of an anomalous observation would similarly leave us without a model with which to view the world.

Of course, we do change our beliefs about the trapezoidal room, the hollow mask, the imaginary ball thrown in the air, and so on, when we are shown how the illusion works. That our illusory perceptions persist even when we know how the illusion works is testament to the conservative nature of our perceptual and cognitive apparatus. To bring this discussion back to Lakatos' theory of rationality, the philosophical and evolutionary psychological considerations discussed here demonstrate how the epistemically conservative approach elucidated by his methodology of scientific research programmes is well-grounded not only in theory, but also in practice.

The theory of rationality I have advocated in this essay also partakes in the dialectic between observation and theory. Our knowledge of how we know improves with the improvement in our observational, psychological and physical theories. And improvements in our theory of rationality will, likewise, require modifications in our appraisal of observational, psychological and physical theories. The general acceptance of double-blind procedures in recent years was motivated by the new developments in the psychology of perception. And this improvement in our methodology of theory appraisal has led to further improvements in our theory of perception. A universal fallibilism should cause us no concern, for it cannot undermine an objective search for truth.

Nonetheless, the theory of rationality that I have developed here is far from complete. The definitions set out in §4.1 were necessarily brief and require much more precision and expansion. Secondly, I had discussed briefly how to apply the criteria for strong evidence to the appraisal of competing theories in §4.3, but much more work needs to be done in developing a method of application in real historical situations. For example, I have not mentioned the problem of anomalies for theory appraisal, and Lakatos had deliberately ignored the problem completely. However, it does seem that the severity and number of anomalies for a theory is important for theory appraisal.

Furthermore, the situation is complicated by the fact that many of our theories are not strict deductive systems, as Lakatos had already pointed out. A complete theory of theory appraisal must incorporate some measure to gauge the strength of corroborated novel evidence, and this measure will have to be an adaption of Lakatos' measure of the severity of a test [Lakatos 1978b: 175]. So, on this account, the measure of the relative strength of corroborated novel evidence e for a theory T relative to a rival theory T', and sharing auxiliary theories A, will be the probability of the evidence e given T and A minus the probability of the evidence e given T' and A, or p(e,T&A) – p(e,T'&A).

8. Logical and Rational Incommensurability

A number of other outstanding problems require detailed solution. The two most important ones are the problem of logical incommensurability and the problem of implicit rationality. The problem of the logical incommensurability of at least some rival high-level theories had been pushed hard by Kuhn and Feyerabend. This problem is critical for two reasons. Firstly, if the epistemic status of an evidence-statement is not determinable independently of such theories, my *Criterion 4* above (§4.3) cannot be satisfied. Secondly, if there are no evidence-statements that simultaneously bear logical relationships of consistency with one theory (in conjunction with auxiliary hypotheses) and inconsistency with its rival, then it appears that such theories cannot be *directly* compared in the way that I have envisaged in my application of *Criterion 5* above (§4.3). It seems that on this account, crucial experiments used to decide between rival theories in the history of science are in name only.

Candidates often touted as a pair of logically incommensurable theories are Newtonian and Einsteinian mechanics. Both served to explain the motions of terrestrial and celestial bodies. The discrepancy between the theoretical and measured advance of the perihelion of the planet Mercury is widely taken as leading to the demise of Newton's theory and the confirmation of Einstein's. It is undeniable that the advance of the perihelion of Mercury cannot be described in an observation language that is interpretable independently of a theory of space and time. Even though the terms 'mass', 'space', 'time', and so on, are not invariant between Newton's and Einstein's theories, it is plausible to suggest that, at the very least, they are implicitly contradictory in their existential claims. Relativity theory, for example, implicitly denies that there are any entities that satisfy the postulates of Newtonian mechanics. It excludes the possibility of bodies that are mass-invariant with respect to velocity. Tied to an adequate theory of reference, such a realist interpretation may provide a comprehensive and compelling solution to the problem of logically incommensurable theories.

Notwithstanding a semantic solution to the problem of incommensurability, rival theories *can* be compared for adequacy. This is because the predicted and actual results of crucial experiments are described in a language that is neutral between the rival theories. The observation statements, for example, describing the advance of the perihelion of Mercury are the same from both the Newtonian and the Einsteinian point of view. The position measurements that are read off the astronomical instruments used in measuring the advance of Mercury's orbit are decipherable independently of the two theories. In consequence, we can say that the evidence-statements in support of Einsteinian Relativity and against Newtonian mechanics are 'independent' of the two theories. Thus, an analysis of the way scientists actually conduct experiments saves the application of my *Criterion 4* and *5*. And this is while conceding that the theories in contention do not stand exhaustively in the same logical and semantical relationships.

So, however an adequate theory of meaning change (sense and reference) may work out, the current lack of a solution does not impact the kind of theory of rationality I am advocating here. At the very least, even if the problem of logical incommensurability remains unsolvable, this does not detract from the rational commensurability of rival scientific theories. In the end, even for Kuhn [1977a: 185, 199], any purported logical

incommensurability does not prevent rational dialogue on the comparative merits and demerits of rival theories. Kuhn had explicitly conceded that his five criteria for rational theory choice apply across paradigms (or 'disciplinary matrix', as he later preferred to call them.)[32]

For Feyerabend also, during his rationalist phase before going over to radical epistemological relativism, logical incommensurability did not stand in the way of conducting crucial experiments to test rival theories. Feyerabend [1970: 226] argued that the truth of an evidence-statement resulting from a crucial experiment is not determinable by a rival theory's postulates alone, but by the interpretation of an experience by such postulates. When scientists conduct crucial experiments, he said, they are in fact conducting two experiments and not one. During a crucial experiment, each research programme is tested for its ability to anticipate and novelly derive phenomena interpreted in its own terms.[33] So, even without recourse to direct semantic comparison, Feyerabend conceded, Newton's program led to predictive failure while Einstein's led to predictive success.

The other problem that I had mentioned requiring detailed solution; that of implicit rationality, may be considerably easier to solve than the problem of logical incommensurability. It arises for those who wish to deem science a rational enterprise. The problem is that different scientists have, by and large, professed that their theories have been arrived at, developed or appraised by a number of different methodologies. These methodologies, opponents of Lakatos claim, bear little or no resemblance to the MSRP. Induction from 'facts', falsificationism and conventionalism are some of the more popular autobiographical offerings of scientists. If the MSRP is the most adequate methodology, then in what sense can scientists be said to have appraised theories rationally? The solution to this problem lies, I think, in making an important distinction between how an individual scientist or scientific community thinks it appraises theories and how in actual fact it evaluates them. It is then open for a Lakatosian to argue that scientists had appraised their theories rationally, although they had given the wrong reasons for thinking so.

This approach becomes credible when we appreciate the complex relations between theory and practice. Kuhn's elucidation of how scientists are inculcated into their scientific community's ways of doing things sheds light on scientists' thinking about their own practices. Coupled with the fact that most scientists working on the ground are unfamiliar with the discipline of epistemology and the various theories in the philosophy of science, it is understandable that scientists will adopt the particular philosophical outlooks for which they have had the most contact. Many successful scientists, for example, repeated the standard account they read in their text books; that theories are inductions from 'facts'.[34] Despite the popularity of Popperian falsificationism among scientists from the mid-nineteenth century, Kragh [2013] illustrates well how some well-respected cosmologists promulgated naïve falsificationism while paying it lip service in practice.

[32]For a discussion of Kuhn's weakening of his incommensurability thesis, see Musgrave [1971] and Sankey [1993].

[33]See Franklin [1984] for a practical example of this.

[34]This account originates with the early proponents of the Scientific Revolution, such as Francis Bacon and Isaac Newton, so they are in respectable company.

9. Conclusion

In this essay, I have assumed that Lakatos' MSRP *qua* historico-sociological thesis is substantially correct. My main purpose here was to provide a more thorough vindication of his MSRP *qua* theory of rationality. I had argued that Lakatos' defence of his theory of rationality, by appealing to its success as a historiographical research programme rationally reconstructing the history of science, was inadequate. This was because Lakatos had provided us with no reason for selecting the history of science as the exemplar of rationality, and so his MHRP appears arbitrary. Once it is supplemented with a *prima facie* argument for the rationality of science, his MHRP does serve as a partial vindication of his MSRP.

Nonetheless, the MHRP partly presupposes the adequacy of the MSRP *qua* theory of rationality. This fact, coupled with the reliance on a *prima facie* argument for the rationality of science, and Lakatos' poor public relations job in defending this rationality independently of the history of science, made the defence of the MSRP *qua* theory of rationality incomplete. I had sought to remedy this deficiency by laying the groundwork for a defence of the MSRP that was independent of the history of science.

My method for achieving this was to stipulate what we ordinarily mean by epistemological terms, such as 'truth' and 'rational', and to decide on the bearers of truth. I then indicated the epistemic relationship of truth-bearers to evidence in an objectivist epistemology and argued that this relationship is satisfied by three necessary conditions. Because these three conditions in isolation underdetermine theory choice, I had argued for two additional fundamental conditions.

Further elaboration of these five conditions requires an ontological-cosmological framework for dealing with observation statements. Applying the five fundamental criteria for theory choice, I argued that realism provides the best metaphysical framework. Drawing on recent work on the theory-ladenness of our observation language and on the psychology of perception, I then specified two explicit conditions for the acceptance of observation statements. I further argued that the use of controlled double-blind procedures is necessitated by these conditions.

No new elements appear in the epistemology developed here, for they can all be found elsewhere. What I have tried to achieve is the systematisation of these elements into one coherent framework, with that framework provided by the requirements of an objectivist epistemology. This framework of seven conditions is characterised by two kinds of criteria. The 'criteria of dependence' define the necessary logical relationships between theory and evidence while the 'criteria of independence' focus on the required freedom of evidence from theory and bias. If, as I argue, Lakatos' MSRP is reflected in the conditions elucidated here, then his theory of rationality is on the way to becoming vindicated independently of the history of science. There is much more work to be done in developing the outline that I have given here. If Lakatos was correct in his view that the history of science is the history of the application of the MSRP, then, hopefully, I have gone some way in answering Feyerabend's question, 'What's so great about science?'

References

Epistemology

Allan, Leslie 2016a. The Existence of Mind-Independent Physical Objects, URL = <http://www.rationalrealm.com/philosophy/metaphysics/mind-independent-physical-objects.html>.

Allan, Leslie 2016b. Imre Lakatos: A Critical Appraisal, URL = <http://www.rationalrealm.com/philosophy/epistemology/imre-lakatos-critical-apraisal.html>.

Ariew, Roger 2014. Pierre Duhem, *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/fall2014/entries/duhem/>.

Aune, Bruce 1970. *Rationalism, Empiricism and Pragmatism: An Introduction*, New York: Random House.

Ayer, A. J. 1969. *The Problem of Knowledge*, Harmondsworth: Penguin.

Ayer, A. J. 1971. *Language, Truth and Logic*, Harmondsworth: Penguin.

Ayer, A. J. 1972. *Probability and Evidence*, London: Macmillan.

Ayer, A. J. 1974. *The Origins of Pragmatism*, London: Macmillan.

Ayer, A. J. 1979. *The Foundations of Empirical Knowledge*, London: Macmillan.

Barnes, Barry and David Bloor 1982. Relativism, Rationalism and the Sociology of Knowledge, in *Rationality and Relativism*, eds M. Hollis and S. Lukes, Oxford: Blackwell: 21–47.

Berkeley, George 1710 (1974). Principles of Human Knowledge, in *Berkeley's Philosophical Writings*, ed. D. M. Armstrong, New York: Macmillan: 41–128.

Berkeley, George 1713 (1974). Three Dialogues Between Hylas and Philonous, in *Berkeley's Philosophical Writings*, ed. D. M. Armstrong, New York: Macmillan: 129–225.

Bird, Alexander 2012. The Structure of Scientific Revolutions and its Significance: An Essay Review of the 50th Anniversary Edition, URL = <http://bjps.oxfordjournals.org/content/early/2012/10/30/bjps.axs031.full> *The British Journal for the Philosophy of Science* (Advance Access: October 30, 2012): 1–26.

Bird, Alexander 2013. Thomas Kuhn, *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/fall2013/entries/thomas-kuhn/>.

BonJour, Laurence 2013. Epistemological Problems of Perception, *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/spr2013/entries/perception-episprob/>.

Chakravartty, Anjan 2014. Scientific Realism, *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/spr2014/entries/scientific-realism/>.

Chalmers, A. F. 1979. Towards an Objectivist Account of Theory Change, *British Journal for the Philosophy of Science* 30: 227–33.

Chalmers, A. F. 1980. An Improvement and a Critique of Lakatos's Methodology of Scientific Research Programmes, *Methodology and Science* 13: 2–27.

Chalmers, A. F. 1982. *What Is This Thing Called Science?*, 2nd edn, Queensland: University of Queensland Press.

Cohen, I. B. 1976. William Whewell and the Concept of Scientific Revolution, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 55–63.

Craig, William 1953. On Axiomatizability within a System, *Journal of Symbolic Logic* 18/1, March: 30–32.

Craig, William 1956. Replacement of Auxiliary Expressions, *Philosophical Review* 65: 38–55.

Duhem, Pierre 1962. *The Aim and Structure of Physical Theory*, trans. Phillip P. Weiner, New York: Atheneum.

Ellis, Brian 1965a. The Origin and Nature of Newton's Laws of Motion, in *Beyond the Edge of Certainty*, ed. E. G. Colodny, Prentice Hall: 29–68.

Ellis, Brian 1965b. A Vindication of Scientific Inductive Practices, *American Philosophical Quarterly* 2/4, October.

Ellis, Brian 1979. *Rational Belief Systems*, Oxford: Basil Blackwell.

Ellis, Brian 1984. The Ontology, La Trobe University, unpublished.

Ellis, Brian 1984. Scientific Realism, La Trobe University, unpublished.

Ellis, Brian 1984. The Programme of Analysis, La Trobe University, unpublished.

Ellis, Brian 1984. Realism and Truth, La Trobe University, unpublished.

Ellis, Brian 1984. Solving the Problem of Induction, La Trobe University, unpublished.

Feigl, Herbert 1970. Research Programmes and Induction, in *PSA 1970, Boston Studies in the Philosophy of Science, Volume 8*, eds R. C. Buck and R. S. Cohen, Dordrecht: Reidel: 147–50.

Feyerabend, Paul K. 1967. Review of the Structure of Science, *British Journal for the Philosophy of Science* 18: 237–49.

Feyerabend, Paul K. 1970. Consolations for the Specialist, in *Criticism and the Growth of Knowledge*, eds I. Lakatos and A. Musgrave, London: Cambridge University Press: 197–230.

Feyerabend, Paul K. 1974. Zahar on Einstein, *British Journal for the Philosophy of Science* 25: 25–28.

Feyerabend, Paul K. 1975. *Against Method*, London: Verso.

Feyerabend, Paul K. 1978. *Science in a Free Society*, London: Verso.

Feyerabend, Paul K. 1979. On the Critique of Scientific Reason, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 109–43.

Fox, John F. 1981. Critical Notice: Appraising Lakatos, *Australasian Journal of Philosophy* 59/1: 92–103.

Franklin, Allan 1984. Are Paradigms Incommensurable?, URL = <http://www.oxfordjournals.org/page/4874/3> *The British Journal for the Philosophy of Science* 35/1: 57–60.

Gardner, Martin 1957. *Fads and Fallacies in the Name of Science*, New York: Dover.

Grandy, Richard E., ed., 1973. *Theories and Observation in Science*, Englewood Cliffs: Prentice-Hall.

Hall, Richard J. 1970. Can We Use the History of Science to Decide Between Competing Methodologies, in *PSA 1970, Boston Studies in the Philosophy of Science, Volume 8*, eds R. C. Buck and R. S. Cohen, Dordrecht: Reidel: 151–9.

Hanson, Norwood Russell 1970. Is There a Logic of Scientific Discovery? in *Readings in the Philosophy of Science*, ed. B. A. Brody, Prentice-Hall: 620–33.

Hempel, Carl G. 1966. *Philosophy of Natural Science*, Englewood Cliffs: Prentice-Hall.

Hilgard, Ernest R. and Richard C. Atkinson 1953. *Introduction to Psychology*, 4th edn, New York: Harcourt, Brace and World.

Howson, Colin, ed., 1976. *Method and Appraisal in the Physical Sciences*, Cambridge: Cambridge University Press.

Hoyningen-Huene, Paul 1992. The Interrelations between the Philosophy, History and Sociology of Science in Thomas Kuhn's Theory of Scientific Development, URL = <http://www.oxfordjournals.org/page/4874/1> *The British Journal for the Philosophy of Science* 43/4: 487–501.

Hurd, D. L. and J. J. Kipling, eds, 1964. *The Origins and Growth of Physical Science, Volume 1*, Harmondsworth: Penguin.

Hurd, D. L. and J. J. Kipling, eds, 1964. *The Origins and Growth of Physical Science, Volume 2*, Harmondsworth: Penguin.

Irzik, Gürol and Teo Grünberg 1995. Carnap and Kuhn: Arch-Enemies or Close Allies?, URL = <http://www.oxfordjournals.org/page/4874/5> *The British Journal for the Philosophy of Science* 46/3: 285–307.

Koertge, Noretta, 1970. Inter-theoretic Criticism and the Growth of Science, in *PSA 1970, Boston Studies in the Philosophy of Science, Volume 8*, eds R. C. Buck and R. S. Cohen, Dordrecht: Reidel: 160–73.

Koertge, Noretta 1976. Rational Reconstructions, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 359–69.

Kragh, Helge 2013. "The Most Philosophically Important of All the Sciences": Karl Popper and Physical Cosmology, *Perspectives on Science* 21/3: 325–357.

Kuhn, Thomas S. 1957. *The Copernican Revolution*, New York: Random House.

Kuhn, Thomas S. 1970a. *The Structure of Scientific Revolutions*, 2nd edn, Chicago: University of Chicago Press.

Kuhn, Thomas S. 1970b. Notes on Lakatos, in *PSA 1970, Boston Studies in the Philosophy of Science, Volume 8*, eds R. C. Buck and R. S. Cohen, Dordrecht: Reidel: 137–46.

Kuhn, Thomas S. 1980. The Halt and the Blind: Philosophy and the History of Science, URL = <http://www.oxfordjournals.org/page/4874/8> *The British Journal for the Philosophy of Science* 31/2: 181–92.

Lakatos, Imre and Alan Musgrave, eds, 1970. *Criticism and the Growth of Knowledge*, London: Cambridge University Press.

Lakatos, Imre 1970. Replies to Critics, in *PSA 1970, Boston Studies in the Philosophy of Science, Volume 8*, eds R. C. Buck and R. S. Cohen, Dordrecht: Reidel: 174–82.

Lakatos, Imre 1978a. *The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1*, eds J. Worrall and G. Currie, Cambridge: Cambridge University Press.

Lakatos, Imre 1978b. *Mathematics, Science and Epistemology: Philosophical Papers Volume 2*, eds J. Worrall and G. Currie, Cambridge: Cambridge University Press.

Locke, John (1690) 1976. *An Essay Concerning Human Understanding*, ed. A. D. Woozley, London: Collins.

Maxwell, A. E. 1970. *Basic Statistics in Behavioural Research*, Harmondsworth: Penguin.

Mason, Stephen F. 1962. *A History of the Sciences*, New York: Collier.

Mayo, Deborah G. 1996. Ducks, Rabbits, and Normal Science: Recasting the Kuhn's-Eye View of Popper's Demarcation of Science, URL = <http://www.oxfordjournals.org/page/4874/9> *The British Journal for the Philosophy of Science* 47/2: 271–90.

McGinn, Colin 1979. An *A Priori* Argument for Realism, *Journal of Philosophy* 76/3: 113–33.

McMullin, Ernan 1976. The Fertility of Theory and the Unit for Appraisal in Science, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 395–432.

Miller, Arthur I. 1974. On Lorentz's Methodology, *British Journal for the Philosophy of Science* 25: 29–45.

Musgrave, Alan 1971. Kuhn's Second Thoughts, URL = <http://www.oxfordjournals.org/page/4874/6> *The British Journal for the Philosophy of Science* 22/3: 287–97.

Musgrave, Alan 1974. Logical versus Historical Theories of Confirmation, *British Journal for the Philosophy of Science* 25: 1–23.

Musgrave, Alan 1976. Method or Madness?, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 457–91.

Nagel, Ernest 1961. *The Structure of Science*, London: Routledge and Kegan Paul.

Pap, Arthur 1963. *An Introduction to the Philosophy of Science*, London: Eyre and Spottiswoode.

Popper, Karl R. 1968. *The Logic of Scientific Discovery*, London: Hutchinson.

Post, H. R. 1976. Novel Predictions as a Criterion of Merit, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 493–95.

Prokhovnik, S. J. 1974. Did Einstein's Programme Supersede Lorentz's, *British Journal for the Philosophy of Science* 25: 336-9.

Quine, Willard Van Orman 1963. *From a Logical Point of View*, 2nd edn, London: Harper and Row.

Quine, W. V. and J. S. Ullian 1970. *The Web of Belief*, New York: Random House.

Radnitzky, Gerard 1976. Popperian Philosophy of Science as an Antidote Against Relativism, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 505–46.

Reichenbach, Hans 1980. *From Copernicus to Einstein*, trans. Ralph B. Winn, New York: Dover.

Rudner, Richard 1953. The Scientist *Qua* Scientist Makes Value Judgements, *Philosophy of Science* 20/1: 1–6.

Salmon, Wesley C. 1966. *The Foundations of Scientific Inference*, Pittsburgh: University of Pittsburgh Press.

Sankey, Howard 1993. Kuhn's Changing Concept of Incommensurability, URL = <http://www.oxfordjournals.org/page/4874/2> *The British Journal for the Philosophy of Science* 44/4: 759–74.

Schaffner, Kenneth F. 1974. Einstein versus Lorentz: Research Programmes and the Logic of Comparative Theory Evaluation, *British Journal for the Philosophy of Science* 25: 45–78.

Scheibe, Erhard 1976. Conditions of Progress and the Comparability of Theories, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 547–68.

Shimony, Abner 1976. Comments on Two Epistemological Theses of Thomas Kuhn, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 569–87.

Stanford, Kyle 2013. Underdetermination of Scientific Theory, *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/win2013/entries/scientific-underdetermination/>.

Toulmin, Stephen 1976. History, Praxis and the 'Third World'. Ambiguities in Lakatos' Theory of Methodology, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 655–75.

Swinburne, Richard, ed., 1974. *The Justification of Induction*, London: Oxford University Press.

Swoyer, Chris 2014. Relativism, *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/sum2014/entries/relativism/>.

Thornton, Stephen 2014. Karl Popper, *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/sum2014/entries/popper/>.

Wartofsky, Marx W. 1976. The Relation Between Philosophy of Science and History of Science, in *Essays in Memory of Imre Lakatos, Boston Studies in the Philosophy of Science, Volume 39*, eds R. S. Cohen, P. K. Feyerabend and M. W. Wartofsky, Dordrecht: Reidel: 717–37.

Worrall, John 1976 .Thomas Young and the 'Refutation' of Newtonian Optics: A Case-study in the Interaction of Philosophy of Science and History of Science, in *Method and Appraisal in the Physical Sciences*, ed. Colin Howson, Cambridge: Cambridge University Press: 107–79.

Zahar, Elie 1973. Why Did Einstein's Research Programme Supersede Lorentz's, in *Method and Appraisal in the Physical Sciences*, ed. Colin Howson, Cambridge: Cambridge University Press: 211–75.

Psychology of Perception

Allport, G. W. and T. F. Pettigrew 1957. Cultural Influence on the Perception of Movement: The Trapezoidal Illusion Among Zulus, *Journal of Abnormal and Social Psychology* 55/1: 104–8.

Anstis, S. M. and R. L. Gregory 1965. The After-effect of Seen Motion: The Role of Retinal Stimulation and of Eye Movements, *Quarterly Journal of Experimental Psychology* 17: 173–4.

BBC 2010. Try the McGurk Effect! - Horizon: Is Seeing Believing? - BBC Two, URL = <http://www.youtube.com/watch?v=G-lN8vWm3m0> (Published: 10 November, 2010).

Bubic, Andreja, D. Yves von Cramon and Ricarda I. Schubotz 2010. Prediction, Cognition and the Brain, *Frontiers in Human Neuroscience* 4/25, doi: 10.3389/fnhum.2010.00025, URL = <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2904053/pdf/fnhum-04-00025.pdf > (Published: 22 March, 2010).

eChalk Education 2012. The Rotating Mask Illusion, URL = <http://www.youtube.com/watch?v=sKa0eaKsdA0> (Published: 20 July, 2012).

Julesz, Bela 1964. Binocular Depth Perception without Familiarity Cues, *Science* 145: 356–62.

Borresen, C. R. and W. H Lichte 1962. Shape-constancy: Dependence upon Stimulus Familiarity, *Journal of Experimental Psychology* 63/1: 91–7.

Bruner, J. S. and L. Postman 1950. On the Perception of Incongruity: A Paradigm, *Journal of Personality* 18: 206–23.

Comalli Jr., P. E. 1960. Studies in Physiognomic Perception: VI. Differential Effects of Directional Dynamics of Pictured Objects on Real and Apparent Motion in Artists and Chemists, *Journal of Psychology* 49: 99–109.

Cross, John and Jane Cross 1969. The Misperception of Rotary Motion, *Perception and Psychophysics* 5/2: 94–6.

Day, R. H. and R. P. Power 1963. Frequency of Apparent Reversal of Rotary Motion in Depth as a Function of Shape and Pattern, *Australian Journal of Psychology* 15/3: 162–74.

Gibson, J. J. and E. J. Gibson 1957. Continuous Perspective Transformations and the Perception of Rigid Motion, *Journal of Experimental Psychology* 54/2: 129–38.

Hanson, Norwood R. 1973. Observation, in *Theories and Observation in Science*, ed. Richard E. Grandy, Englewood Cliffs: Prentice-Hall: 129–46.

Harding, G. F., S. M. Glassman and W. C. Helz 1957. Maturation and the Spiral Effect, *Journal of Abnormal and Social Psychology* 54: 276–7.

Hastorf, A. H. 1950. The Influence of Suggestion on the Relationship between Stimulus Size and Perceived Distance, *Journal of Psychology* 29: 195–217.

Hubel, D. H. and T. N. Wiesel 1962. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex, *Journal of Physiology* 160: 106–54.

Hubel, D. H. and T. N. Wiesel 1968. Receptive Fields and Functional Architecture of Monkey Striate Cortex, *Journal of Physiology* 195: 215–43.

Jones, E. E. and J. S. Bruner 1954. Expectancy in Apparent Visual Movement, *British Journal of Psychology* 45: 157–65.

Joynson, R. B. and L. J. Newson 1962. The Perception of Shape as a Function of Inclination, *British Journal of Psychology* 53: 1–15.

Julesz, Bela 1964. Binocular Depth Perception without Familiarity Cues, *Science* 145: 356–62.

Kogo, Naoki 2012. Bi-Stable Figure-Ground Perception: Underlying Mechanisms and the Memory Effect, URL = <http://www.gestaltrevision.be/en/what-we-do/individual-projects/52-postdocs/316-bi-stable-figure-ground-perception-underlying-mechanisms-and-the-memory-effect> (Accessed: 19 December, 2014).

Kveraga, K., Avniel S. Ghuman and Moshe Bar 2007. Top-Down Predictions in the Cognitive Brain, *Brain and Cognition* 65: 145–168, doi: 10.1016/j.bandc.2007.06.007, URL = <http://kveragalab.org/uploads/3/4/8/6/3486904/2007braincog.pdf> (Published: 7 June, 2007).

Landauer, A. A. and R. S. Rodger 1964. Effect of 'Apparent' Instructions on Brightness Judgements, *Journal of Experimental Psychology* 68/1: 80–4.

Levine, R., I. Chein and G. Murphy 1942. The Relation of the Intensity of a Need to the Amount of Perceptual Distortion: A Preliminary Report, *Journal of Psychology* 13: 283–93.

Loftus, E. F. 1979. The Malleability of Human Memory, *American Scientist* 67: 312–20.

McGee, J. M. 1963. The Effect of Group Verbal Suggestion and Age on the Perception of the Ames Trapezoidal Illusion, *Journal of Psychology* 56: 447–53.

Nordhjem, Barbara 2012. The "Where" and the "What" of Bistable Perception, URL = <http://nordhjem.net/?p=128> (Published: 30 June, 2012).

Proshansky, H. and G. Murphy 1942. The Effects of Reward and Punishment on Perception, *Journal of Psychology* 13: 295–305.

Scientific American 2012. What Is the Ames Illusion? - Instant Egghead #23, URL = <http://www.youtube.com/watch?v=gJhyu6nlGt8> (Published: 11 October, 2012).

Singer, B. and V. A. Benassi 1981. Fooling Some of the People All of the Time, *The Skeptical Inquirer* 5/2: 17–24.

Slack, C. W. 1956. Familiar Size as a Cue to Size in the Presence of Conflicting Cues, *Journal of Experimental Psychology* 52/3: 194–8.

Smith, K. R., G. B. Parker and G. A. Robinson Jr. 1951. An Exploratory Investigation of Autistic Perception, *Journal of Abnormal and Social Psychology* 46: 324–6.

Steinfeld, G. J. 1968. Effect of Verbal Information on Recognition of Fragmented Figures, *American Journal of Psychology* 81: 425–30.

Sunderland Echo 2008. Vanishing Ball Illusion, URL = <https://www.youtube.com/watch?v=goYPBRA0i-0> (Published: 15 August, 2008).

Tajfel, H. 1957. Value and the Perceptual Judgement of Magnitude, *Psychological Review* 64/3: 192–204.

Thomson, Eric 2009a. Consciousness (6): Reversible Figures, URL = <http://neurochannels.blogspot.com.au/2009/09/consciousness-6-reversible-figures.html> (Published: 5 September, 2009).

Thomson, Eric 2009b. Consciousness (7): More Ambiguous Figures, URL = <http://neurochannels.blogspot.com.au/2009/09/consciousness-7-more-ambiguous-figures.html > (Published: 10 September, 2009).

Vernon, M. D. 1971. *The Psychology of Perception*, 2nd edn, Harmondsworth: Penguin.

Wyatt, D. F. and D. T. Campbell 1951. On the Liability of Stereotype or Hypothesis, *Journal of Abnormal Psychology* 46: 496–9.

