Ducks, Rabbits, and Normal Science: Recasting the Kuhn's-Eye View of Popper's Demarcation of Science

Author(s): Deborah G. Mayo

Source: *The British Journal for the Philosophy of Science*, Jun., 1996, Vol. 47, No. 2 (Jun., 1996), pp. 271-290

Published by: Oxford University Press on behalf of The British Society for the Philosophy of Science

Stable URL: https://www.jstor.org/stable/687948

Brit. J. Phil. Sci. **47** (1996), 271–290

Ducks, Rabbits, and Normal Science: Recasting the Kuhn's-eye View of Popper's Demarcation of Science

Deborah G. Mayo

ABSTRACT

Kuhn maintains that what marks the transition to a science is the ability to carry out 'normal' science—a practice he characterizes as abandoning the kind of testing that Popper lauds as the hallmark of science. Examining Kuhn's own contrast with Popper, I propose to recast Kuhnian normal science. Thus recast, it is seen to consist of severe and reliable tests of low-level experimental hypotheses (normal tests) and is, indeed, the place to look to demarcate science. While thereby vindicating Kuhn on demarcation, my recasting of normal science is seen to tell against Kuhn's view of revolutionary science.

1 Introduction

Soon after the publication of his enormously influential book, *The Structure of Scientific Revolutions*, Thomas Kuhn offered 'a disciplined comparison' of his and Popper's views of science in a paper, 'Logic of Discovery or

Psychology of Research?' It begins with these lines:

> My object in these pages is to juxtapose the view of scientific development outlined in my book, [*Structure*] with the better known views of our chairman, Sir Karl Popper. Ordinarily I should decline such an undertaking, for I am not so sanguine as Sir Karl about the utility of confrontations ... Even before my book was published two and a half years ago, I had begun to discover special and often puzzling characteristics of the relation between my views and his. That relation and the divergent reactions I have encountered to it suggest that a disciplined comparison of the two may produce peculiar enlightenment (Kuhn [1970], p. 1).

'Peculiar enlightenment' is an apt description of what may be found in going back to Kuhn's early comparison with Popper and the responses it engendered. What makes my recasting of Kuhn peculiar is that while it justifies the very theses by which Kuhn effects the contrast with Popper, the picture that results is decidedly *un*Kuhnian. That being the case, I do not doubt that my recasting differs from the 'peculiar enlightenment' Kuhn intended, but my task, I should be clear at the start, is not a faithful explication of what Kuhn saw himself as doing. Rather, it is an attempt, at times deliberately unKuhnian, to see what philosophical mileage can be acquired from exploring the Kuhnian contrast with Popper. I shall also suggest how this leads to a more adequate image of experimental knowledge than we now possess.

Kuhn begins his comparison by listing the similarities that place both himself and Popper within the same minority of philosophers of science of the day. Both accept theory-ladenness of observation, hold some version of realism, and reject the view of 'progress by accretion', emphasizing instead 'the revolutionary process by which an older theory is rejected and replaced by an incompatible new one' (Kuhn [1970], p. 2). Despite these agreements Kuhn finds that he and Popper are separated by a 'gestalt switch'. Popper views the overthrowing and replacement of scientific theories as the main engine of scientific growth. 'Scientific knowledge', Popper declares, 'grows by a more revolutionary method than accumulation—by a method which destroys, changes, and alters the whole thing ...' (Popper [1962], p. 129). Kuhn views such revolutionary changes as extraordinary events radically different from the 'normal' scientific tasks of 'puzzle solving'—extending, applying, and articulating theories. While for Kuhn, 'normal science' constitutes the bulk of science, what has intrigued most philosophers of science is Kuhnian revolutionary science—with its big changes, gestalt switches, conversion experiences, incommensurabilities, and the challenges thereby posed to the rationality of theory change. Kuhn's description of normal science, when discussed at all, is generally dismissed as relegating day-to-day science to an unadventurous working out of 'solvable puzzles', and 'mopping-up' activities.

In this vein, Popper [1970] responds to Kuhn in 'Normal Science and Its Dangers'. He was aghast at Kuhnian normal science with its apparent call to 'abandon critical discourse' and embrace unquestioning allegiance to a single accepted paradigm, encompassing theories, as well as standards and values for their appraisal. Kuhnian normal science, were it actually to exist, Popper declares, would be pathetic or down right dangerous: 'In my view the 'normal' scientist, as Kuhn describes him, is a person one ought to be sorry for... He has been taught in a dogmatic spirit: he is a victim of indoctrination' (Popper [1970], pp. 52–3).

While finding there to be much to dislike about normal science, philosophers find little use for it when it comes to solving the challenges that Kuhn raises to the rationality of theory change. I find Kuhnian normal science to be considerably more fruitful. Far from being the uncritical affair Popper fears, normal science turns out to offer an effective basis for severe testing. This, in turn, provides the key to getting around the big problems alleged to arise in revolutionary science, or large-scale theory change—or so I argue. To begin, let us pursue a bit further the contrasts Kuhn draws between his philosophy of science and Popper's. Except where noted, all references will be to Kuhn [1970].

2 Turning Popper's duck into Kuhn's rabbit and vice versa

Kuhn asks:

> How am I to persuade Sir Karl, who knows everything I know about scientific development and who has somewhere or other said it... that *what he calls a duck can be seen as a rabbit*? How am I to show him what it would be like to wear my spectacles when he has already learned to look at everything I can point to through his own? (p. 3, emphasis added).

Kuhn's tactic is to take the linchpins of Popper's philosophy and show how, wearing Kuhnian glasses, they appear topsy-turvy. While in Popper's view, what sets science apart from other practices is its willingness continually to subject its theories to severe and crucial tests, to the Kuhnian eye

> it is normal science, in which Sir Karl's sort of testing does not occur, rather than extraordinary science which most nearly distinguishes science from other enterprises. If a demarcation criterion exists (we must not, I think, seek a sharp or decisive one), it may lie just in that part of science which Sir Karl ignores (p. 6).

But because normal science, for Kuhn, does not involve Popperian-style testing, Kuhn provocatively declares, 'In a sense, to turn Sir Karl's view on its head, *it is precisely the abandonment of critical discourse that marks the transition to a science*' (Kuhn, 'Reflections on My Critics',[1] p. 273, emphasis added).

If only we would view the highlights of the Popperian landscape through his spectacles, Kuhn proposes, we would come to see how Popper's view gets turned on its head. Specifically, we would see why, where Popper sees a fundamental theory failing a severe test, Kuhn sees a paradigm failing in its 'puzzle-solving ability' (crisis), and why, where Popper sees a lack of testability, Kuhn sees a lack of puzzle-solving. In so doing, Kuhn assures us, we would begin to see the sense in which '*severity of test-criteria is simply one side of the coin whose other face is a puzzle-solving tradition*', and with this, Kuhn proclaims, 'Sir Karl's duck may at last become my rabbit' (p. 7).

I propose that we look at the high points of the Popperian landscape that the Kuhn's-eye view brings into focus. The main highlights that interest me underlie the following portions of the above passages:

(1) 'it is normal science, in which Sir Karl's sort of testing does not occur, rather than extraordinary science which most nearly distinguishes science from other enterprises.'
(2) 'it is precisely the abandonment of critical discourse that marks the transition to a science.'
(3) 'severity of test-criteria is simply one side of the coin whose other face is a puzzle-solving tradition' (p. 7).

In order to extract the epistemological lessons I am after, however, spectacles capable of seeing the normative dimension are required.

2.1 Turning Popper's duck into Kuhn's rabbit

Imre Lakatos set about providing a normative interpretation of Kuhn, yet what I am proposing should not be confused with the Lakatosian model. Lakatos's gambit is to reconstruct Kuhn's socio-psychological description of paradigm change by cloaking it in terms of rational changes of research programmes: 'I look at continuity in science through "Popperian spectacles". Where Kuhn sees "paradigms", I *also* see rational research programmes' (Lakatos [1970], p. 177)[2].

My task differs from that of Lakatos in three key ways. First, in his

[1] I will hereafter refer to this article as 'Reflections'.

[2] Lakatos provides this interesting way to view his reconstruction: *'my concept of a "research programme" may be construed as an objective, 'third world' reconstruction of Kuhn's socio-psychological concept of paradigm'*: thus the Kuhnian 'Gestalt-Switch' can be performed without removing one's Popperian spectacles' (Lakatos [1970], p. 179, n. 1).

rational reconstruction of Kuhn, Lakatos seeks not to prohibit so much as to give a rational spin to the very features to which Popper objects—in particular, allowing that researchers (if creative enough) can always rationally (progressively) defend a paradigm in the face of anomalies. My reworking of normal science, far from providing a rational gloss on Kuhn's picture of paradigm change, leads to a view that conflicts with what Kuhn himself says when it comes to revolutionary science. Despite having this implication for paradigm change, a second key point of contrast with Lakatos is that I move away from focusing on large-scale units of change, and focus instead on small-scale tests of local hypotheses in normal science (normal testing). Finally, my normative spectacles must reveal, not methods for after-the-fact reconstructions, but forward-looking methods of normal testing. They must enable us to answer epistemological questions: why does following normal scientific principles yield reliable knowledge? Why do enterprises not characterized by these principles turn out to be less successful sciences or not sciences at all?

2.2 Turning Kuhn's rabbit into Popper's duck

Turning Popper's duck into Kuhn's rabbit will not teach us the epistemological lessons we are after; but neither will it do to follow the path proposed by another Popperian, J.O. Wisdom, namely, to convert Kuhn's rabbit into a Popperian duck. Wisdom's [1974] idea is essentially to view Kuhn's normal science through Popper's testing spectacles. As Popper is well aware, in order for a theory genuinely to fail a test, it is not enough that one of its consequences turns out to be wrong, a whole crop of loopholes have to be investigated, and such investigations are the sorts of things that one can suppose normal science to accomplish. Normal science is called upon, in effect, to pinpoint blame in the face of an apparent anomaly.

Wisdom's analysis is insightful and is, in my view, completely correct, so far as it goes. But it goes only so far as to show how Kuhnian normal science can serve as handmaiden to Popperian testing. By starting from the Popperian point of view of what the task is, and utilizing Kuhnian ideas to fill out that task, it gets the order wrong and misses most of the uses of normal science. Normal science is needed to get extraordinary science off the ground all right, but starting with the latter as primary fails to do justice to Kuhn's main point of contrast with Popper.

Kuhn is quite clear as to the main point of contrast with Popper (e.g. in his 'Reflections'). Kuhn says, 'my single genuine disagreement with Sir Karl about normal science' is in holding that when a full-bodied theory is at hand 'the time for steady criticism and proliferation of theories has passed ... [Scientists] can instead apply their talents to the puzzles' of normal science ('Reflections', p. 246). This, Kuhn says, is a better 'research strategy'. Before we can evaluate Kuhn's research strategy, we need to get clearer on the aims of normal science and what is needed to achieve them. The result will be not to turn the Kuhnian rabbit into the Popperian duck nor vice versa. Instead it will be to extract from the very characteristics that Kuhn identifies in normal science, the basis for forward-looking procedures for acquiring reliable experimental knowledge.

3 Normal science as normal testing

Let us begin by asking what is involved when practitioners turn their attention and apply their talents to the tasks of normal science. Kuhn identifies three classes of problems:

(1) *Determination of significant fact*. This concerns 'that class of facts ... shown to be particularly revealing of the nature of things' ([1962] p. 25) such as stellar position and magnitude, the specific gravities of materials, wave lengths, electrical conductivities, boiling points, acidity of solutions.
(2) *Matching of facts with theory*. This concerns the methods and instruments to arrive at data that 'can be compared directly with predictions' (ibid., p. 26).
(3) *Articulation of the paradigm theory*. This has several parts including: a) determining physical constants (e.g. gravitational constant, Avogadro's number, Joules' coefficient, the electronic charge); b) determining quantitative laws; and c) conducting experiments to decide how to apply the paradigm to related areas.

Clearly, finding and evaluating hypotheses of these sorts is not a matter of uncreative hack science, nor does Kuhn suggest otherwise. What is required in each of these classes, Kuhn stresses, is shared criteria for determining if problems or puzzles are solved:

> No puzzle solving enterprise can exist unless its practitioners share criteria which, for that group and for that time, determine when a particular puzzle has been solved. The same criteria necessarily determine failure to achieve a solution, and anyone who chooses may view that failure as the failure of a theory to pass a test (p. 7).

I do. Perhaps it is largely to achieve a stark contrast with Popperian tests that Kuhn calls normal problems 'puzzles' to begin with. The idea is that the loser in case a conjectured solution fails is not the fundamental theory but the practitioner who was not brilliant enough. But since 'blaming the practitioner', even for Kuhn, just means that the practitioner's conjectured solution fails to hold up to testing, it is less misleading to talk in terms of testing conjectured solutions to normal problems. There is much in Kuhn to support this reading:

> There is one sort of 'statement' or 'hypothesis' that scientists do repeatedly subject to systematic test. I have in mind statements of an individual's best guesses about the proper way to connect his own research problem with the corpus of accepted scientific knowledge. He may, for example, conjecture that a given chemical unknown contains the salt of a rare earth, that the obesity of his experimental rats is due to a specified component in their diet, or that a newly discovered spectral pattern is to be understood as an effect of nuclear spin (p. 4).

But these are what we more usually regard as hypotheses—not mere puzzles. Moreover, the next steps, the test of experiment, have the flavour of hypothetico-deductive tests:

> the next steps ... are intended to try out or test the conjecture or hypothesis. If it passes enough or stringent enough tests, the scientist has made a discovery or has at least resolved the puzzle he had been set. If not, he must either abandon the puzzle entirely or attempt to solve it with the aid of some other hypothesis (p. 4).

So, as I read Kuhn, for a hypothesis to pass the test of experiment it must have passed 'enough or stringent enough tests', and to accept a normal hypothesis is to accept it as correctly solving the associated normal problem. If the hypothesis fails the test, it is concluded that it does not solve the puzzle (that it is incorrect or is false). Blaming the background theory is tantamount to *changing* the puzzle and is disallowed. Indeed, in my reading, the main purpose of calling a normal problem a 'puzzle' is to call attention to the fundamental restriction on what counts as an admissible solution: if a conjectured solution fails the test only the conjecture and 'not the corpus of current science is impugned' by the failure (p. 5).

Kuhn shows, as an example, how some eighteenth-century scientists, finding anomalies between the observed motions of the moon and Newton's laws, 'suggested replacing the inverse square law with a law that deviated from it at small distances. To do that, however, would have been to change the paradigm, to define a new puzzle, and not to solve the old one' (Kuhn [1962], p. 39). This was not an admissible solution. The normal scientist must face the music.

Underlying the stringency demand, I propose, is the implied requirement that before a hypothesis H is taken to solve a problem H must have stood up to scrutiny: if a hypothesis H is taken to solve the problem it must be very unlikely that it really does *not* solve the problem. This requirement, which we may call the *reliability* or *severity requirement*, is one that I develop elsewhere.[3] It requires that normal scientists declare a problem solved by a hypothesized solution *H* only if *H* has withstood a severe scrutiny—one that *H* would very probably have failed, were it not a correct solution. In testing solution *H*, normal scientists design tools to discriminate correct from erroneous solutions. It is a principal value of normal science to be able to put together so potent an arsenal for unearthing an erroneous solution to a puzzle that, when no error is found, there are excellent grounds for concluding that the error is absent.[4]

It is only by some such reliability or severity requirement, I maintain, that Kuhn is right to locate the growth of knowledge in normal science. It shows how Kuhn can be correct to regard it as a better research strategy to focus on normal testing rather than on Popperian testing (which he views as criticizing fundamental theories). The reason, I propose, is that one learns much more through normal testing.[5]

After all, why does Kuhn say that, in the fact of a rich enough theory to 'support a puzzle solving tradition', it is fruitful to concentrate on normal problems? The answer that my spectacles discern is this: if one has an interesting theory, one with predictions, suggestions for improvement, challenging puzzles, and so on, then taking up its challenges *will teach a great deal*, and *a portion of what is learned will remain despite changes in theory*. With respect to the solved problems in normal research, Kuhn says 'at least part of that achievement always proves to be permanent' ([1962], p. 25). To ignore its challenges is to forfeit this knowledge.

One can go further. Pursuing normal problems is a good strategy because, if there are anomalies that call for changes in theory, doing so will reveal them as well as help point to the adjustments indicated:

> In the developed sciences ... it is technical puzzles that provide the usual occasion and often the concrete methods for revolution ... Because they can ordinarily take current theory for granted, exploiting rather than criticizing it, the practitioners of mature sciences are freed to explore nature to an esoteric depth and detail otherwise unimaginable. Because that exploration will ultimately isolate severe trouble spots, *they can be confident that the pursuit of normal science will inform them when and where they can most usefully become Popperian critics* ('Reflections', p. 247, emphasis added).

In short, the rationale for pursuing normal problems is that (if done right) some positive payoff is assured. If normal science yields problem

[3] I do so in Mayo [1991] and, much more fully, in Mayo [1996]. For a discussion of the difference between my notion of severity and that of Popper, see Mayo [1996].
[4] I call this an 'argument from error'. An analogous argument would be given for concluding that an error is present—that *H* is not a correct solution to the problem.
[5] In 'Reflections' (p. 243), Kuhn himself identifies the key difference between himself and Popper as one of research strategy.

solutions, then new knowledge would have been brought forth. If an anomalous result is determined by normal testing to be real—that 'it will not go away'—then there is knowledge of a real experimental effect. Further normal testing will serve to indicate if adjustments and revisions are called for. If, alternatively, the effect is a genuine anomaly for the underlying theory, normal science will allow finding this out by means of gross or repeated failures (crisis). Even such crises, Kuhn notices, serve a creative function in developing alternative theories. Finally, normal science is the source of the most effective and severe tests of fundamental or basic theory:

> [T]hough testing of basic commitments occurs only in extraordinary science, it is normal science that discloses both the points to test and the manner of testing (p. 6).

and

> because the [theory] test arose from a puzzle and thus carried settled criteria of solution, *it proves both more severe and harder to evade than the tests available within a tradition whose normal mode is critical discourse rather than puzzle solving* (p. 7, emphasis added).

This last sentence gets us to the heart of why, from the Kuhn's-eye point of view, 'severity is the flip side of puzzle solving', and why one finds the most severe tests of theories, just what Popper seeks, in practices that have been engaged in the puzzle-solving of normal science (normal testing).

4 The Kuhn's-eye view of demarcation

Indeed, as Kuhn remarks (p. 7), Popper's demarcation line and his own often coincide, despite the fact that they are identifying very different processes. They often agree on what counts as a science or not, despite the fact that, on the face of it, their two demarcation criteria are nearly opposite. For Popper, the hallmark of science is criticism, testability, and falsifiability, whereas Kuhn, in deliberate contrast, declares that 'it is precisely the abandonment of critical discourse that marks the transition to a science' (p. 6).

It is important to keep in mind, however, that the critical discourse Kuhn is disparaging is the special kind of criticism that he imagines Popper to be championing: a relentless attack on fundamentals. It helps, in grasping Kuhn here, if each time we read 'critical discourse' we tack on the phrase 'rather than puzzle solving'. For Kuhn, finding a practice whose normal mode is critical-discourse-rather-than-puzzle-solving is the surest tip-off that its scientific status is questionable. Hence Kuhn's provocative claim that a demarcation criterion may be found in the portion of science badly obscured by Popperian spectacles (normal science).

It is highly misleading to call what goes on in good normal testing an abandonment of critical discourse, as normal science is itself based on severe and critical tests. But, on my reading, what Kuhn takes good sciences to abandon is not normal testing—which is where all the fruitful learning really takes place—but rather, 'critical-discourse-rather-than-puzzle-solving'. Good sciences do not and should not do what Kuhn takes Popper to be championing: relentlessly attacking fundamental theories, looking always for rival theories, and doing so to the exclusion of the positive learning of normal science. Without endorsing this provocative–idiosyncratic usage of 'critical discourse', for the purposes of this paper it helps us to reach a plausible construal of Kuhn's demarcation criterion.

4.1 Astrology

In making out his contrast with Popper, Kuhn takes the example of astrology, wishing to avoid controversial areas like psychoanalysis (p. 7). Kuhn's focus, he says, is on the centuries during which astrology was intellectually respectable.[6] Astrology was unscientific, says Kuhn, not because it failed to be falsifiable nor even because of how practitioners of astrology explained failed predictions. The problem is that astrologers had no puzzles, they could not or did not engage in normal science. Let us try to unpack this.

Engaging in normal science requires a series of puzzles and strict criteria that all practitioners agree to use to tell if puzzles are solved. But a practice does not automatically *become* scientific by erecting a series of puzzles and rules to pronounce them solved or not. Becoming a genuine science is not something that can occur by community decree, nor does Kuhn think it is.[7] Kuhn balks at those who would find in him recipes for becoming scientific, apparently unaware of how he invites this reading by failing to articulate the kinds of tests needed to carry out normal science legitimately and why only these tests qualify. Still, there are several places where Kuhn hints at the criteria normal testing requires (namely, reliability or stringency). The most telling of all, I find, is Kuhn's critique of astrology.

With astrology, Kuhn observes, not only are the predictions statistical, there is a tremendous amount of 'noise' from background uncertainties.

> Astrologers pointed out, for example, that ... the forecast of an individual's future was an immensely complex task, demanding the

[6] The function of the example is not just to make out his demarcation but to show 'that of the two criteria, testing and puzzle solving, the latter is at once the less equivocal and the more fundamental' (p. 7).

[7] This is stressed by Hoyningen-Heune [1993] in his analysis of Kuhn's *Structure of Scientific Revolutions* (see, for example, p. 193).

> utmost skill, and extremely sensitive to minor errors in relevant data. The configuration of the stars and eight planets was constantly changing; the astronomical tables used to compute the configuration at an individual's birth were notoriously imperfect; few men knew the instant of their birth with the requisite precision. No wonder, then, that forecasts often failed (p. 8).

Kuhn's point seems to be this: astrology, during the centuries when it was reputable, did not fail to be scientific because it was not testable nor because practitioners did not take failures as grounds to overthrow astrology. There are plenty of perfectly good sciences that would act similarly. The reason the practice of astrology was unscientific is that *practitioners did not or could not learn from failed predictions.*[8] And they could not learn from them because there were too many justifiable ways to explain failure. They could not use failures or anomalies *constructively.*

> The occurrence of failures could be explained, but particular failures did not give rise to research puzzles, for no man, however skilled, could make use of them in a constructive attempt to revise the astrological tradition. There were too many sources of difficulty, most of them beyond the astrologer's knowledge, control, or responsibility. Individual failures were correspondingly uninformative ('Reflections', p. 276).

The above passage is most revealing. For failed predictions to 'give rise to research puzzles' a failure must give rise to a fairly well-defined problem—specifically, the problem of how to explain it. It must be possible, in other words, to set up a reliable enquiry to determine its cause and/or modifications called for. This is the kind of information normal tests can provide.

> Compare the situations of the astronomer and the astrologer. If an astronomer's prediction failed and his calculations checked, he could hope to set the situation right. Perhaps the data were are fault ... Or perhaps theory needed adjustment ... The astrologer, by contrast, had no such puzzles (p. 9).

In order to 'set the situation right' one needs to be able to discriminate between proposed explanations of the failure. Unless one can set up a stringent enough test of a hypothesized explanation (so that its passing can reliably be attributed to its being correct), then that failed prediction is unconstructive and uninformative.

By the same token, so long as there is no way to cut down these alternative explanations of failure, there are no grounds for arguing that the failures *should* have been attributed to the falsity of astrology as a

[8] It is unimportant for our point that Kuhn thinks it wrongheaded to speak of failed solution as mistakes—that he limits mistakes to erring in applying some rule.

whole. In other words, if failed predictions do not give rise to research puzzles (reliable enquiries into their cause), then one cannot come to learn whether and, if so, how they can be explained within the global background theory. Thus, they cannot warrant discrediting the whole theory; they cannot *warrant* (Popperian) critical discourse.

4.2 Unwarranted critical discourse

The practitioners of astrology, Kuhn notes, 'like practitioners of philosophy and of some social sciences ... belonged to a variety of different schools, and the interschool strife was sometimes bitter. But ... [f]ailures of individual predictions played very little role' (p. 9, n. 2). Practitioners were happy to criticize the basic commitments of competing astrological schools; rival schools were constantly having their basic presuppositions challenged. What they lacked was that very special kind of criticism that allows genuine learning—the kind where a failed prediction can be pinned on a specific hypothesis. Their criticism was not constructive: a failure did not *genuinely indicate* a specific improvement, adjustment, or falsification.

Thus I propose to construe the real force of Kuhn's disparaging practices 'whose normal mode is critical discourse' as disparaging those practices that engage in criticism even where the criticism fails to be driven by the constrained tests that exemplify good normal science. What is being disparaged, and rightly so, is unwarranted and unconstructive criticism. When the day-to-day practice is criticism that is not the result of the stringent constraints of normal testing, then that criticism is of the unwarranted or unconstructive variety. It is *mere* critical discourse. Non-sciences engage in mere critical discourse, not genuine criticism that allows learning from empirical tests.

The situation in astrology exemplified an extreme case of a situation in which severe tests are precluded. The situation might be described in modern statistical terms as having too much uncontrolled variability, or as lacking a way to distinguish the 'signal' from the noise. The situation is typical, Kuhn notes, of practices that one might call 'crafts', some of which eventually manage to make the transition to sciences (e.g. medicine). The transition from craft to science, Kuhn observes, correlates with supporting normal science or normal testing.

To see how, let us go to a practice that, unlike astrology, is sufficiently developed to support normal testing (puzzle-solving). If a hypothesized solution to a normal problem fails a test, it could, theoretically, be accounted for by alleging a fundamental flaw in the underlying theory—but such a criticism would very likely be unwarranted (at least not just from this one failure). Thus to proceed regularly this way would very often be in error, thereby violating the reliability requirement of normal testing. On these grounds, normal science calls for abandoning this type of criticism. For the same reason, it admonishes the practice of dealing with a failed solution (failed hypothesis) by changing the problem it was supposed to solve. An enterprise that regularly allowed such a cavalier attitude towards failure would often be misled.

Changing the problem, blaming one's testing tools or the background theory *where these are unwarranted*, is the kind of criticism that should be disallowed. Only then can the practice of hypothesis appraisal be sufficiently constrained so as to identify correctly genuine effects, gain experimental knowledge—more generally, accomplish the tasks of normal science *reliably*. Thus recast, Kuhn's demarcation criterion may be seen to pick out those practices that afford experimental learning. For my part, I suggest we view such a demarcation criterion as indicating when *particular enquiries*, rather than whole practices, are scientific. It becomes, roughly:

> *Demarcating scientific inquiry:* what makes an empirical enquiry scientific is that it can and does learn from normal tests, that it can accomplish one or more tasks of normal testing *reliably*.[9]

This criterion becomes more specific when particular types of normal test results are substituted.[10]

4.3 Summary of previous sections

Our analysis has so far brought us to the following recasting of the Kuhnian observations with which we began: to understand the nature of the growth of scientific knowledge one should look to tests of hypotheses about specific types of experiments (normal experimental testing). An adequate account of normal testing should be one that serves each of the functions Kuhn accords it, with the additional proviso that it do so reliably and with warrant. Seen through our spectacles, what distinguishes Kuhn's demarcation from Popper's is that for Kuhn the aim is not mere criticism but constructive criticism.

Our recasting of normal science, I believe, substantiates the three highlights of Kuhn's contrast with Popper with which we began. Let me repeat them here: (1) 'it is normal science, in which Sir Karl's sort of testing does

[9] The demarcation criterion that emerges should really be qualified to refer only to enterprises for predicting, controling, or understanding the world, in short, to *intended* sciences. It would not be a disparagement, say, of art, to violate this criterion.

[10] For example, an important type of normal test result is a failed prediction. The difference between a scientific and an unscientific treatment of a failed prediction is the extent to which it is used to learn (about its cause, about needed modifications, etc.).

not occur, rather than extraordinary science which most nearly distinguishes science from other enterprises'; (2) 'it is precisely the abandonment of critical discourse that marks the transition to a science'; (3) 'severity of test-criteria is simply one side of the coin whose other face is a puzzle-solving [i.e. a normal science] tradition' (p. 7).

Briefly, our gloss on them went as follows: the fundamental features of scientific enquiries are to be found in the criteria of normal testing, and these criteria demand stringent normal tests, not (uninformative) attacks on fundamental theory. Because anomalies that are reliably produced in normal tests indicate real effects that will not go away, they provide the most severe tests of theories—when these are warranted. This explains Kuhn's promise that scientists 'can be confident that the pursuit of normal science will inform them when and where they can most usefully become Popperian critics' ('Reflections', p. 247), that normal science will tell them when and where to find fault with the underlying theory. But Kuhn, we shall see, reneges on his promise. Once having brought normal scientists to the crisis point, Kuhn still will not let them be Popperian testers!

5 Theory choice, gestalt switches, and all that

Why, having brought normal scientists to the crisis point, to the point of a warranted criticism of theory, will Kuhn still not let them be Popperian testers? According to Kuhn, the products of normal science are never going to be decisive for falsifying or adjudicating between global theories. Testing and changing global theories or paradigms turns out not to be a matter of reasoning at all. To cite one of several colorful passages, Kuhn declares that

> the proponents of competing paradigms practice their trades in different worlds ... the two groups of scientists see different things when they look from the same point in the same direction ... before they can hope to communicate fully, one group or the other must experience the conversion that we have been calling a paradigm shift. Just because it is a transition between incommensurables, the transition between competing paradigms cannot be made a step at a time, forced by logic and neutral experience. Like the gestalt switch, it must occur all at once (though not necessarily in an instant) or not at all ([1962], p. 150).

This picture of revolutionary science has been convincingly criticized by numerous authors (e.g. Laudan [1984]; Shapere [1984]). But what I wish to consider, if only briefly, is how our recasting of normal science tells against Kuhn's view of global theory change.

Kuhn's notion of paradigm is notoriously equivocal.[11] We may agree to

[11] See, for example, Masterman [1970] and Shapere ([1984], Chs 3 and 4).

see a Kuhnian paradigm as including theories, specific hypotheses, an ontology as well as research aims and methods both for directing normal research and testing hypotheses. (In 'Reflections', p. 271, Kuhn says he would prefer to use the term 'disciplinary matrix'.) For Kuhn, sharing a paradigm or set of paradigms is what accounts for their 'relative unanimity in problem-choice and in the evaluation of problem-solutions' ('Reflections', p. 271). Viewing global theory change as switching all elements of the paradigm, Kuhn supposes there to be no place to stand and scrutinize two whole paradigms, as a genuine paradigm test would require. However, we must be very careful to distinguish what Kuhn runs together here. The ingredients of the problem choice task differs considerably from those of the task of criticizing proposed solutions. We need, in short, to distinguish the paradigm's role in providing (1) a *research programme*: a source of problems and guides for solving problems or puzzles, and (2) *normal testing models*—or what I call *experimental testing models*: tools for testing hypothesized solutions reliably or for normal hypothesis testing. These testing models need to be developed in their own right for a full experimental account. The main point of the distinction just now is to see why changing a research programme is not the same as changing experimental testing tools.

5.1 Experimental testing models (error paradigms)

Significantly, Kuhn remarks that he was originally led to the concept of a 'paradigm' in thinking of the concrete problem solutions or exemplars that practitioners share and which enable them to agree if a problem is solved ('Reflections', p. 272). This is the role I propose to give to certain *experimental testing models* or *testing exemplars*. Kuhn's own use of the example of astrology (as a classic non-science) is itself an example of what I have in mind here. Nevertheless, I depart from Kuhn in several important ways. The main difference is that, in my view, standard examples or normal testing exemplars are not a set of tools available only to those working within a given global theory or paradigm. Instead, they consist of any models and methods relevant for testing solutions of normal problems, and these come from various background theories, from mathematics, statistics, and from theories of instruments and experiments. While this and related departures result in a view of normal science very different from Kuhn's official position, it is quite in keeping with my normative recasting of Kuhn.

By Kuhn's own lights, before normal practitioners may take a puzzle as solved, the hypothesized solution must have passed stringent enough tests. The arsenal needed for normal testing, then, is a host of tools for detecting whether and how conjectured hypotheses (of a given type) can fail. They call for methods capable not just of determining whether a hypothesis correctly solves a problem, but of doing so reliably.

However, Kuhn fails to disentangle the experimental testing portion of the paradigm from immersion in a research programme. As such, it is not surprising that global theory change winds up appearing arational—quite like the (experimentally) unwarranted critical discourse Kuhn attributes to non-sciences. It is as if the very process that allows practices to become scientific starts to go in reverse, until we are back to 'mere' critical discourse:

> critical discourse recurs only at moments of crisis when the bases of the field are again in jeopardy. Only when they must choose between competing theories do scientists behave like philosophers (pp. 6–7).

5.2 The circularity thesis

Kuhn supposes that members of competing global theories necessarily subscribe to different values and standards in such a way as to make it inevitable that one's own global theory gets defended. This thesis, which we may call the *circularity thesis*, is most clearly stated in *Structure:*

> Like the choice between competing political institutions, that between competing paradigms proves to be a choice between incompatible modes of community life. When paradigms enter, as they must, into a debate about paradigm choice, their role is necessarily circular. Each group uses its own paradigm to argue in that paradigm's defense ([1962], p. 94).

> They will inevitably talk through each other when debating the relative merits of their respective paradigms. In the partially circular arguments that regularly result, each paradigm will be shown to satisfy more or less the criteria that it dictates for itself and to fall short of a few of those dictated by its opponent (ibid., pp. 109–10).

While such circular defences are possible, none of the requirements of paradigm theories, even as Kuhn conceives them, makes their role in theory appraisal necessarily circular. On the contrary, there is much in what normal science requires that militates against such circularity—even in times of crisis. After all 'the criteria that [the paradigm] dictates for itself' are those in the experimental testing part of the paradigm. Those standards, if appropriate for their *own* goals, must condemn such question-begging arguments as failing utterly to probe a theory severely. At some point (i.e. with regard to some normal hypothesis), defending a global theory no-matter-what clashes with the requirements of normal testing.

Such defences may and do occur, but they do not count as warranted—by the strictures of good normal science. Why? *Because they come down to a blanket refusal to acknowledge that a hypothesized solution to a normal problem fails, and that betrays an essential requirement of normal science.*[12]

Let us sketch what happens, according to Kuhn's circularity thesis, when a global theory, T_1, slips into a crisis. Within T_1, which we are assuming to be a genuine science with a normal tradition and so on, genuine anomalies have been identified. These anomalies identify genuine effects that need explaining. These give rise to normal puzzles, i.e. normal testing, to scrutinize attempted solutions, to 'set the situation right'. The criteria of T_1, by dint of its enjoying a normal science tradition, severely constrain attempts to deal with such anomalies. A genuine crisis is afoot when, after considerable effort, T_1 is unable to explain away the anomalies as due either to initial conditions or background hypotheses.

Notice, it follows that *the normal (experimental) testing criteria of T_1 themselves serve to warrant the existence of anomalies and crisis.* They, incorporating as they must, the general criteria of normal testing, indicate when an anomaly is real and (as Kuhn himself says) '*unevadable*'. That is, they indicate when to put blame elsewhere is tantamount to unwarranted criticism. Do they not, by the same token, indicate that any attempt to save a theory—if that defence depends upon evading the anomaly—violates the very norms upon which enjoying a normal 'puzzle-solving' tradition depends? The norms bar procedures of admitting hypotheses as solutions to puzzles, we saw, if they would often do so erroneously. Of course, it may take a while until attempted defences come up against the wall of normal testing strictures. But with a genuine crisis, it seems to me, that is exactly what happens. Moreover, from Kuhn's demarcation criterion, *it is possible to recognize* (even if not sharply) that a practice is losing its normal puzzle-solving ability. (Astrology serves as a kind of exemplar of a practice that falls over on to the non-science line.)

These remarks should not be misunderstood. What normal science must condemn is not saving a global theory in the face of severe anomaly—although that is what Popperian spectacles might have us see. What it must condemn (recalling Kuhn's demarcation) is being incapable of *learning* from normal testing. In any particular case, the obstacles to learning that are condemned are very specific: having to reject experimentally demonstrated effects, contradict known parameter values, change known error distributions of instruments or background factors, and so on.

Consider what Kuhn calls for when scientists, having split off from

[12] Of course, nothing guarantees that actual science obeys the constraints of Kuhnian normal science. In fact, however, Kuhn's account of normal science is descriptively accurate for the bulk of important scientific episodes.

global theory T_1 to develop some rival T_2, come knocking on the door of their less adventurous colleagues, who are still muddling through the crisis in T_1. Confronted with rival T_2, which, let us suppose, solves T_1's crisis-provoking problem, crisis scientists in T_1 necessarily defend T_1 circularly. This circularity thesis requires them to do a turn-about and maintain that T_1 will eventually solve this problem, or that the problem was not really very important after all. Once the members of rival T_2 go away (back to their own worlds, presumably), members of T_1 can resume their brooding about the crisis they have identified with their paradigm. Were they to do this, they would indeed be guilty of the unwarranted criticism and mere name-calling Kuhn finds typical of non-sciences. But Kuhn has given no argument to suppose that crisis scientists necessarily do this.

Nor will one find an argument as to why Kuhn takes away what I thought he had promised us—that a crisis compelled by good normal science lets us finally be *warranted* Popperian testers, and reject the theory (as having it wrong at least so far as its key hypotheses go)—quite apart from stopping work on it. Instead one finds that, when turning his gaze to the problem of large-scale theory appraisal, Kuhn is simply wearing spectacles that necessarily overlook the role of the shared strictures and arguments of normal testing.

Let me be clear about what I am claiming. I do not assert that experimental arguments always exist to guide theory appraisal, but rather deny Kuhn's claim that they never do. In order for experimental arguments to ground theory appraisal, moreoever, the experimental testing frameworks of the rival large-scale theories need not be identical. It is sufficient for the needed arguments to be made out by appeal to the interparadigmatic normal testing tools. How can we suppose such a shared understanding? It follows from taking seriously the criteria for good normal scientific practice, criteria which, *for Kuhn*, must hold for any practice that enjoys a normal scientific tradition. Moreover, the historical record reveals case after case where even the most ardent proponents are forced to relent on the basis of very local, but very powerful experimental tests. The Kuhn of normal science can explain this consensus quite naturally; the Kuhn of revolutionary science cannot.

A Kuhnian may agree with my thesis about shared testing models, yet deny that the experimental arguments provided offer a basis for appraising global theories. Nevertheless, that is still *no argument* for Kuhn's thesis that global theory change cannot turn on experimental arguments, and, indeed, Kuhn fails to supply one. Rather, his thesis results from assumptions I have not addressed here—that theory change is a conversion experience, that it requires one to 'go native', and is complete only when the new theory establishes a grip on one's mind. They, like the circularity thesis, should be rejected.

6 Conclusion

We began by asking what philosophical mileage could be acquired from exploring Kuhn's contrast of his position with Popper's. The main thing we saw is that we can have a perfectly sensible construal of Kuhn's comparison with Popper—and Kuhn is correct—as far as his discussion of normal science. Normal scientists, in my rereading of Kuhn, have special requirements without which they could not learn from experimental tests. They insist on stringent tests in the sense I have set out (reliable or severe). They could not learn from failed solutions to normal problems if they could always change the question, make alterations, etc. That is what Kuhn says. That is what having a normal science tradition is all about. But then we have some curious consequences at the level of theory appraisal.

A normal science may be led, via the criteria of normal science, says Kuhn, to crisis. It is recognized as crisis because of the stringent rules of normal science. Suddenly, when confronted with a rival theory, normal scientists, says Kuhn, do an about-face. They start furiously defending their theory and denying it is in crisis. Kuhn gives no argument for supposing this always happens. And my point is that his own view of normal science militates against this supposition.

Answering Kuhn does not require showing that global theory testing is always a function of experimental knowledge, but merely denying the Kuhnian view that it cannot be. My solution is based on one thing that normal practitioners, even from rival paradigms, have in common (by dint of enjoying a normal testing tradition): they can and do perform the tasks of normal science reliably. That is the thrust of Kuhn's demarcation criterion, as I have recast it.

The distinction identified by the demarcation criterion (although it is not intended to be sharp and may well admit of degrees) is not between theories but between enquiries that are scientific or informative, and those that are not. Enquiries are informative to the extent that they enable experimental knowledge, that is, learning from normal science. Taking Popperian aim at global theories when this is not constrained by severe normal testing is a poor strategy for obtaining experimental knowledge. The constraints of normal testing provide the basis for severe tests and informative scientific enquiries. To understand the nature and growth of experimental knowledge, it is to normal testing that one must look.

For Kuhn, in a genuine science, anomalies give rise to research puzzles. In our recasting of Kuhn this becomes, in a genuinely scientific enquiry, anomalies afford opportunities for learning—opportunities for learning from error. The aim of science is not avoiding anomaly and error, in our view. The aim is being able to learn from anomaly and error.

Department of Philosophy
Virginia Polytechnic Institute and State University
Blacksburg
VA 24061-0126
USA

Acknowledgements

I am grateful to Larry Laudan for very valuable discussions on the ideas in this paper. I thank Paul Hoyningen-Huene and an anonymous referee of this journal for insightful comments on an earlier draft. Portions of this paper are adapted from Mayo 1996.

References

Hoyningen-Huene, P. [1993]: *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*, Chicago, University of Chicago Press.

Kuhn, T. [1962/1970]: *The Structure of Scientific Revolutions*, 2nd edition, Chicago, University of Chicago Press.

Kuhn, T. [1970]: 'Logic of Discovery or Psychology of Research?' and 'Reflections on My Critics', in I. Lakatos and A. Musgrave (*eds*) [1970], pp. 1–23, 231–277.

Lakatos, I. [1970]: 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (eds) [1970], pp. 91–196.

Lakatos, I. [1978]: *The Methodology of Scientific Research Programmes*, in J. Worrall and G. Currie (*eds*), *Philosophical Papers, Vol. I: Imre Lakatos*, Cambridge, Cambridge University Press.

Lakatos, I. and Musgrave, A. (*eds*) [1970], *Criticism and the Growth of Knowledge*, Cambridge, Cambridge University Press.

Laudan, L. [1984]: *Science and Values: The Aims of Science and Their Role in Scientific Debate*, Berkeley, University of California Press.

Masterman, M. [1970]: 'The Nature of a Paradigm', in I. Lakatos and A. Musgrave (*eds*) [1970], pp. 59–89.

Mayo, D. [1991]: 'Novel Evidence and Severe Tests', *Philosophy of Science*, **58**, pp. 523–52.

Mayo, D. [1996]: *Error and the Growth of Experimental Knowledge*, Chicago, University of Chicago Press.

Popper, K. [1962]: *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Basic Books.

Popper, K. [1970]: 'Normal Science and Its Dangers', in I. Lakatos and A. Musgrave (*eds*) [1970], pp. 51–8.

Shapere, D. [1984], *Reason and the Search for Knowledge: Investigations in the Philosophy of Science*, Boston Studies in the Philosophy of Science, Dordrecht, D. Reidel.

Wisdom, J. O. [1974]: 'The Nature of "Normal" Science', in P.A. Schilpp (*ed.*) [1974], *The Philosophy of Karl Popper*, La Salle, IL, Open Court, pp. 820–42.