



1997

Kuhn's Model of Revolutionary Science: Evidence for a Coherence Criteria of Truth

Angela Burnette '97

Illinois Wesleyan University

Recommended Citation

Burnette '97, Angela, "Kuhn's Model of Revolutionary Science: Evidence for a Coherence Criteria of Truth" (1997). *Honors Projects*. Paper 9.

http://digitalcommons.iwu.edu/phil_honproj/9

This Article is brought to you for free and open access by The Ames Library, the Andrew W. Mellon Center for Curricular and Faculty Development, the Office of the Provost and the Office of the President. It has been accepted for inclusion in Digital Commons @ IWU by the faculty at Illinois Wesleyan University. For more information, please contact digitalcommons@iwu.edu.

©Copyright is owned by the author of this document.

**Kuhn's Model of Revolutionary Science:
Evidence for a Coherence Criteria of Truth**

The traditional view in analytic philosophy concerning scientific progress and the justification of scientific theory is characterized by four central tenets:

- (i) Realism. There exists one mind-independent reality, and it is the scientist's goal to derive true theories which accurately describe that reality.
- (ii) Science is cumulative. Scientific endeavors are thought to produce a cumulative body of knowledge, where each new theoretical or methodological addition builds upon and subsumes what is already known.
- (iii) Theory-neutral observation. Scientific activity employs observations which are distinct and independent from theory; and such observations are the mediating basis upon which theories are upheld or rejected, and upon which competing theories are compared.
- (iv) Standards of justification. There exist underlying scientific standards which, together with observations, provide the rationale for the justification of scientific theories and propositions; and these standards are rigid and fixed, allowing for rational evaluations of all scientific theory.

These four points represent a characterization of science that enjoys widespread acceptance within the scientific and philosophical community. This view depicts science as an ever-advancing process of discovery guided by rational standards of justification and a rigorous methodology of theory-neutral observation and investigation. Alongside this view of science is a correspondence based picture of truth and knowledge. Deriving from the realism assumption,

truth is correspondence to reality. For some proposition to be true is just for its contents to accurately describe the relevant portion of independent reality. Further, in accordance with tenets (iii) and (iv), the methodology employed to check the truth of statements about reality requires a comparison of theory content and neutral observation in strict accordance with the appropriate rational standards of investigation and justification. Under this view, truth is correspondence, and knowledge of reality is accessible through observation. In addition, for those who subscribe to this traditional characterization, there is one correct description of the world, which scientists seek to discover. Although focused upon the sciences, this model provides the prototype for all epistemic inquiry for many in analytic philosophy.

One alternative view to this general scientific and epistemic picture is derivable from the work of T.S. Kuhn, who approaches science from an historical perspective and argues that the practice of science does not actually support the traditional analysis. He therefore forwards a view that subverts the traditional understanding of science and knowledge. Kuhn's model of revolutionary science provides strong evidence and convincing argumentation against the traditional doctrine that judgments of truth are based upon correspondence to reality, and for the alternative view that the proper criterion of truth is coherence.

Kuhn's revolutionary science

Kuhn's *The Structure of Scientific Revolutions* (1962) challenges the traditional view by denying three of the aforementioned core assumptions. Kuhn denies the existence of both theory-neutral observation and rational scientific standards; and he further denies that science is cumulative. Kuhn argues instead that science, which admits of several distinct phases, or periods, is better characterized as producing a body of knowledge which "shifts" in content according to the emergence and subsequent destruction of incompatible -- in fact incommensurable -- paradigms, which guide all of "normal" science. A paradigm is best understood as a holistic theoretical and practical matrix which dictates the accepted "law, theory, application, and instrumentation" for a specified scientific field, and also "provides models from which spring particular coherent traditions of scientific research" (Kuhn, 10). Examples of such traditions

include 'Ptolemaic astronomy', and 'Newtonian physics'. "Normal" science is the activity undertaken within the framework of such a tradition. According to Kuhn there are five cyclical phases of science: pre-paradigm science, normal science, crisis science, paradigm rejection and revolution, and finally new paradigm establishment, after which normal science resumes and the cycle repeats itself.

Prior to the establishment of a paradigm, scientific endeavor is characterized by Kuhn as being largely occupied with the activity of "random fact gathering," as there is no theoretical basis to guide investigation; also, what datum are collected within a given field are "likely to seem equally relevant" because there is no rational basis upon which to differentiate between important or promising observations and those which are not truly germane to the given field (Kuhn, 15). Moreover, before a paradigm is in place, there is little consensus between practitioners of a given discipline, which usually results in several factional theoretical schools, none of which can boast any explanatory achievements over competitors.

Kuhn argues that a paradigm generally emerges from among such competing schools as the result of a particularly attractive or powerful accomplishment that places one school in a better position than the others. With the establishment and common acceptance of one particular theoretical structure, researchers can direct their observations and experiments in accordance with the ontological and methodological landscape provided by the agreed upon paradigm. Under such guidance, scientists are in a position to judge the value of various observations, and delineate between important lines of investigation and those without promise. Thus normal science, in contrast to its "immature" precursor, is characterized as purposeful, directed, and capable of advancement just in light of the fact that a paradigm exists as a qualifying standard.

Kuhn argues that a paradigm directs normal science by setting out the parameters of the given field; the paradigm fixes the theoretical structure, methods, experiments, and the problems to be solved. Accordingly, normal science, which accepts a paradigm, is not concerned with theory verification or falsification, but rather with fitting nature to the paradigm through theory articulation and the production of more sophisticated experiments. In so much as science is

normalized within a paradigm, the respective theoretical structure simply is not open to debate, as this structure must be assumed as correct if normal science is to proceed.

In this sense, normal science is chiefly concerned with puzzle solving. The relevant paradigm dictates what has been “figured out” and what areas still prove problematic. Kuhn argues that these scientific “puzzles” are what drive the scientist. What is important to note is that implicit in the motivation to solve puzzles is the firm belief that the answer can be had. Scientists are driven to solve the problems that the paradigm demands because they believe that the accepted theories involved are essentially correct.

In addition to fixing the theoretical and methodological framework, paradigms consequently affect the scientist’s metaphysical picture by establishing the ontology of the scientist’s world; the contents of the world, the sorts of entities and phenomena that it contains, are paradigm-bound. Therefore, Kuhn argues that scientists working within different paradigms are in effect also occupying different worlds. He writes, “the historian of science may be tempted to exclaim that when paradigms change, the world changes with them...In so far as their only recourse to that world is through what they see and do, we may want to say that after a revolution scientists are responding to a different world” (111). For example, phlogiston, long since debunked, was once thought to be a legitimate ontological category. Scientists formulated an explanatory phlogiston theory, and conducted experiments and recorded observations accordingly. Those researchers found something in the world that scientists of today do not.

The idea of different worlds is suggestive of Kuhn’s main assertion regarding scientific paradigms. He insists that successive paradigms are logically incompatible, and therefore do not represent a cumulative enterprise. Kuhn employs the term “revolution” in description of the shift from one paradigm to the next in order to emphasize this point. When one paradigm ceases to be successful and faces an increasing number of anomalous phenomenon which cannot be explained or ignored, a paradigm is pushed into crisis science. At this point, the paradigm begins to breakdown, and a new theoretical system eventually moves in to take its place. Such has been the case throughout science history; the shift from Newtonian dynamics to Einsteinian is a prime

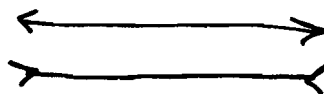
example. It has long been argued, and it is the majority view, that older theories, like Newton's, are derivable from their successors, and therefore are logically included in the newer, larger paradigm. But, if paradigms are correctly construed, this cannot be the case. A new paradigm is established to accommodate the anomalies of the older theory. Thus, the new paradigm must make predictions and offer explanation that the old theories did not. This being the case, the two paradigms cannot be logically compatible. Einstein's work could only be accepted alongside the admission that Newton was wrong.

If it is not the case that one paradigm is just an expansion of its predecessor, building upon the previous theoretical base, then it also cannot be the case that one paradigm is adopted over another according to rational considerations. Such criteria cannot be applied where paradigms themselves are in question; in the absence of an accepted paradigm, there are no rational standards by which one theory is chosen over any other. Rather, decisions between paradigms must make appeal to extra-scientific reasons and are thus in a sense revolutionary. Kuhn furthers this position by arguing that paradigms are incommensurable, and do not admit of discourse between them. There are two separate arguments for the incommensurability of paradigms.

The argument from theory-laden observation

Kuhn appeals to the doctrine of theory-laden observation advanced by N.R. Hanson (1961), arguing that there is no sense in which observation can be neutral with regard to the theoretical background of the observer. What the scientist will observe, or can observe, is fixed, or limited by what the paradigm has established as the legitimate entities in the universe. Further, scientists adhering to different paradigms "see" different things in the world, and make different observations and measurements, depending upon what each paradigm demands. Scientific observation is thus colored by the established paradigm to the extent that not only does the scientist see what he expects to find according to a paradigm, but such scientific observation cannot even exist without the guidance of a paradigm to make sense of raw perceptual input.

Kuhn appeals to the extensive psychological literature of the time to further this claim; he cites numerous observational and gestalt experiments suggesting that the perceived qualities of experimentally displayed objects is strongly correlated with the subjects previous experience. One type of experiment considered involves optical illusions such as the Muller-Lyer figures seen at right. The effect of the arrows is to make the line with the inward pointing arrows appear longer than the other line,



even though the lines are identical. An argument can be made that the theoretical background of the observer, involving three-dimensional projections of convex corners, etc., must account for the illusion, as the lines make equally long retinal imprints.

Kuhn asserts that this literature “suggests that something like a paradigm is prerequisite to perception itself...what a man sees depends both upon what he looks at and also upon what his previous visual-conceptual experience has taught him to see” (Kuhn, 113).

Thus, all observation is “theory laden” in as much as the scientist can only make observations which are allowed by the theoretical framework established by the paradigm to which the scientist is committed. Two scientists working within different paradigms will actually “see” different objects, or features of objects, when visually confronted with the same item. In other words, to the extent that raw perception involves the identification and categorization of what appears in the visual field, differing paradigms will result in different observations.

From this Kuhn argues for the incommensurability of paradigms. Kuhn thinks that adherents from different paradigms cannot communicate successfully because their theoretical backgrounds will be sufficiently different as to exclude a common theoretical ground by which observations can be mutually agreed upon and serve as the subject of a rational dialectic between paradigms. According to this argument, there is no theory-neutral set of observations upon which to base rational argumentation between paradigms.

But, there are problems with Kuhn’s advancement of incommensurability to the extent that theory-laden observation is a highly contentious thesis. J.A. Fodor employs the theory-laden

observation proponent's main evidential basis to argue against the notion that observation is theory-bound (Fodor, 19--). As mentioned above, proponents like Hanson often refer to the psychological experimentation on optical illusions as demonstrating the theory-ladenness of observation, in that optical illusions and other gestalt experiments seem to suggest that such illusions must be explained by appealing to their reliance upon the theoretical background of the observer. But, as Fodor points out, the knowledge that optical illusions are illusions should negate their effect, as such knowledge is then incorporated into the theoretical background of the observer. That such knowledge does not make optical illusions go away undermines the thesis that observation is dependent upon theory.

Fodor offers an alternative explanation for the process of perception positing psychological modules which are like mediating mechanisms between retinal imprints and actual perception (Fodor,--). These mechanisms are uniquely both theory-laden and theory neutral, as they are severely and strictly delimited with regards to what segment of the theoretical background to which they have access. On this view, the perceptual mechanism has access to a small portion of the theoretical background which is set early in development, the mechanism then being in effect cut off from additional theoretical inputs. Here the segment of theory incorporated in the perceptual module would be common to all human observers, thus providing a basis for shared observation. Fodor's model of perception sharpens the objection made by Fodor against Hanson and Kuhn, and is in keeping with current psychological thought.

While this is a powerful argument against the notion of theory-laden observation as conceived of by Hanson and Kuhn, it does not necessarily refute Kuhn's theory of incommensurability. A second line of argumentation is still available to Kuhn.

The argument from theory-bound language and concepts

Even if raw observation is not necessarily theory-bound, there is still reason to think that observational language is still heavily influenced by experience and theoretical background. In so much as observations are made public through reports, raw perception may still be filtered

through theory in order to be operationalized into language, thus maintaining the weakened theory-laden thesis that observational reports and language are theory-bound. In adherence with this weaker version, paradigms still necessarily result in a unique theory-bound language, as well as conceptual schema. Thus scientists from different paradigms cannot effectively enter into dialogue because they do not share a common language of observation. Along these same lines, Fodor concedes that “belief fixation, unlike the fixation of appearances--what I’m calling observation--is a *conservative* process; to a first approximation, it uses everything you know” (Fodor, 249). In other words, one’s actual observations may be neutral with regard to a large portion of one’s theoretical background, but observations are then compared with the entire background theory in order to fix belief.

So, what scientists actually ‘see’ may be the same, but the beliefs which they form about what they see may still differ.

Moreover, these difficulties are compounded by the fact that no extra-paradigmatic conceptual standards exist to ground rational argumentation. Members of competing paradigms will differ in opinion concerning what entities exist in the world, what counts as legitimate scientific method and practice, and what problems should have priority, thus setting different paradigms at cross purposes, and destroying the possibility for effective argumentation regardless of the possibility of coherent communicative interaction. In this sense, when one scientist accuses another of being wrong, neither is in the position to argue their point, as there is no standard to which they both adhere; and neither can be right or wrong, as each need only answer to his own paradigm for justification. Kuhn writes, “...the men who called Copernicus mad because he proclaimed that the earth moved...were not either just wrong or quite wrong. Part of what they meant by ‘earth’ was fixed position...a law that cannot even be demonstrated to one group of scientists may...seem intuitively obvious to another” (150). Thus, scientists from different paradigms cannot provide rational argumentation against one another because they do not share the same conceptual language or framework.

This second argument in favor of incommensurability upon linguistic and conceptual grounds is more defensible than the theory-laden observation argument, while still effectively establishing incommensurability as a viable theory.

At this point, the objection might be raised that there will always be sufficient overlap between different paradigms, especially at the point of transition, to allow for dialogue between them. It simply seem eminently implausible that two scientists, working in the same field, would be unable to communicate about their shared discipline. The problem with this objection is that it fails to recognize a key feature of paradigms. The theoretical structure of beliefs and concepts which constitute a paradigm is such that one altered concept or rejected belief results in an alteration of all other concepts and beliefs. This is necessary if the paradigm as a whole is to remain coherent. Thus, when two scientists differ on one conceptual point, they necessarily differ on all points, although they may not explicitly realize it. Kuhn's paradigms, properly construed, admit of no overlap, let alone sufficient overlap to allow for rational argumentation and dialogue between them. Kuhn is less emphatic on this point, forwarding the idea of incommensurability, and yet saying only that paradigms are always "slightly at cross-purposes," hereby intimating that he perhaps doesn't quite buy his own theory. But, as Kuhn has defined them, I see no reason to deny paradigms full incommensurability by virtue of the feature described above. The possibility of overlap is simply eliminated.

Hence, if Kuhn's picture of science is taken to be essentially correct, then Kuhn has succeeded in undermining the traditional conception of scientific endeavor in crucial ways. First, there can be no theory-neutral observational base upon which to ground assessments of theories; second, the rational standards of justification posited by the traditional view disappear. According to Kuhn, such standards can only exist within a given paradigm, and cannot be appealed to when considering two competing paradigms. Moreover, these two considerations mean that science cannot be cumulative as new paradigms are incompatible with the ones they replace, and cannot therefore be built upon their predecessors.

It should now be clear that Kuhn rejects the possibility of theory neutral observation and of the sorts of extra-paradigmatic rational standards of justification forwarded by a traditional view. As a result, Kuhn also denies that science is cumulative. And, it follows from this that science is not strictly the pursuit of theories which correspond to reality. In fact, normal science merely accepts that such theories are already in place, and is more concerned with work that can only be done within an existing set a theories.

Kuhn and coherence

Theories of truth concern the truth of statements, or propositions; they are designed to explain first, what it is for a statement to be true, and second, how ascriptions of truth are to be applied. It is important to distinguish between these two types of explanation even though it may appear as though the methods by which truth is assigned must crucially depend upon how truth is defined. While such a link seems intuitively correct, it need not necessarily be the case that the definitional explanation of the concept of truth be of the same kind as an explanation of how to apply the concept. These two types of explanation need not necessarily be one and the same. For example, litmus paper tests for acidity, but it does not help define it.

Kuhn's characterization of paradigms and normal science has interesting epistemic applications that suggest an argument for a coherence theory of truth as the criterion for ascribing truth to statements. To the extent that theory determines how raw perception is organized into conscious observations, theory determines how we categorize the world. Consequently, the world is accessible as the subject of our thoughts, desires, beliefs and actions only as a world created by theory. This being the case, a notion of truth, which is a property of our statements about the world, is substantively employable only when characterized in terms of coherence to some paradigm, or world view. In so much as determining the truth or falsity of statements and beliefs is a substantive, meaningful activity employed daily in order that we might function effectively in the world, we employ truth as a tool for achieving our goals and avoiding error, and to be generally successful in our various pursuits.

Moreover, in this active employed sense, truth is best characterized as coherence because, in order to be thus employed, truth must be situated as theory-bound, alongside the observations, beliefs and activities to which truth applies.

As for the commonly accepted alternative theory of truth as correspondence, there may be a sense in which statements about the world are true just when they accurately reflect how the world really is, but this is not the same concept employed to assign truth to statements. Such a criteria of truth is useless, as we have no direct access to the reality to which such truth would correspond, and could never fulfill the requirements of our criterion, and never be in a position to judge a statement to be true.

Also, even prior to Kuhnian considerations, correspondence provides a relatively poor account of how truth is to be assigned. According to the theory, deciding whether or not a given statement is true involves an investigation of the relevant facts in order to assess the level of agreement between statements and the reality they are supposed to represent. This methodology has many serious problems. It is not possible to check the facts of statements made about the past, or of statements that are universal in application. The facts of yesterday are simply not available, and it is impossible to check every instance of a universally applied proposition. Correspondence as a criteria for assigning truth has serious problems regardless of whether or not direct access to the world can be had.

These criticisms of the criteria of correspondence suggest a parallel argument in favor of coherence. Just as the actual facts of the past are inaccessible, our beliefs about these facts are readily available for comparison and evaluation. Likewise, a coherence methodology of truth ascription also eliminated the necessity of investigating every instance of a universally applied theory.

Kuhn's work also places coherence in a position to handle one stock objection to coherence theories in general. It has been objected that two individuals, with different set of coherent beliefs, or different paradigms, could consistently and coherently hold opposite beliefs. Person A might believe proposition P and person B might believe proposition $\sim P$. Both beliefs

would cohere with each persons set of beliefs, thus proposition P would be both true and not true, thereby constituting a contradiction.

A Kuhnian coherence model can respond to this objection in two ways. First, because of the incommensurability of the two sets of beliefs, the proposition in question could not be comparable. The content of person A's proposition P would be different from the content of person B's proposition P; there would be no contradiction because P would not be the same proposition for both A and B. This is part of what it means to say that they have different paradigms. Secondly, even if it were admitted that the exact same proposition could be judged true by one person and false by another, this is not a contradiction. Truth as coherence merely assigns truth to statements, it does not say anything about the way the world really is. Contradictory truth ascriptions certainly do not entail a contradiction in reality.

Hence, there are two characterizations of truth: truth as correspondence and truth as coherence. This dichotomous characterization is consistent with the dichotomous 'reality' in which we exist; one exists mind-independently, and one is 'created' by theory. The point is that truth as coherence most accurately describes the criterial concept of truth that we, as conscious, theory-bound agents necessarily employ.

This does not discount the validity of truth as correspondence as the proper definitional component of a comprehensive theory of truth; there may be some sense in which true statements are true just in light of the fact that they accurately describe the world. But, truth as such as no causal, or substantive efficacy; in accordance with Kuhn, truth as correspondence is in fact epistemically inert, and should therefore be abandoned in favor of the only conception of truth which we actively employ when assigning truth to statements. Truth as coherence to some paradigm is the only view consistent with our usage.

In terms of the traditional view of science and knowledge, Kuhn's arguments, if accepted, refute the tenets which constitute the traditional analysis, and also refute the epistemic model motivated by these tenets. Realism survives, although this too might fall under a more radical interpretation. Without theory-neutral observation and rational standards of justification that hold

across paradigms, the motivation for the resulting correspondence criterion of truth is destroyed. Kuhn's work also simultaneously creates a compelling picture of science that becomes the basis for a coherence theory of truth ascription, which is uniquely positioned to handle objections to coherence in a new and more convincing fashion.

Works Cited

Fodor, J.A. "Observation Reconsidered". *A Theory of Content and Other Essays*.

Cambridge: MIT Press, 1993.

Hanson, N.R. "Observation". *Philosophy of Science*. R. Hollinger and A Kline (ed).

Philosophy of Science. Prometheus, 1988.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. Chicago: University of Chicago

Press, 1962.

Plantinga and the Theory of Knowledge

10/20/18

Knowledge comes in many forms: the common knowledge that snow is white and grass green; slightly lesser known facts such as the atomic weight of silver; personal knowledge of headaches, appointments, and so on; even the time of day, which changes with every passing second is knowledge taken for granted by most. But exactly how it is that persons come to know the color of grass or snow, whether the light is red or green, how much money is left in the checking account and the like, is still the subject of much philosophic debate. When are we correctly judged to have knowledge? What peculiar feature is had by some of our beliefs such that these constitute knowledge while other beliefs do not?

Justified true belief has long been the basis for theories of knowledge and enjoys a central historical position. Moreover, truth, as it applies to the content of our beliefs, alongside the beliefs themselves, are often secondary, or incidental, in epistemic discussions. The meat of the matter traditionally comes down to the analysis of justification. Consequently, theories of knowledge have historically been concerned largely with deontological considerations such as epistemic duty and obligation. This view is called epistemic "internalism." On this account, one mustn't run around accepting beliefs willy-nilly. Beliefs must be carefully monitored, subjected to rigorous justificatory tests, and truth discerned and ascribed accordingly. While we perhaps cannot be blamed for whatever fleeting beliefs pass through the landscape of our consciousness, the internalist claims we must surely be held accountable for which of these we hold up as knowledge. Hence the ancient and all too common rebuke, "He should have known better."

Here we see the core of epistemic internalism: knowledge is a matter of upholding epistemic duty, it is a matter of actively engaging our obligation to the truth by pursuing knowledge with a stringent adherence to justification. Hence, the attainment of knowledge is largely *internal* to us cognizers. If we view our beliefs with the proper critical eye, if we pursue the truth, then knowledge will be ours. The central claim that justification, and therefore knowledge, are intimately and crucially linked to aspects internal to the cognizer dominates the works of most philosophers from Plato to Descartes and Locke, continuing to the present day. This certainly sounds very attractive. Internalism provides for a brand of optimistic autonomy

that may help us sleep a little better at night. Under the umbrella of internalism, broadly construed, the power to attain knowledge is firmly within our grasp and, with diligence and care, we may avoid error and rest assured that knowledge can be had.

Despite this comforting, although laborious, conception of knowledge, recent years have seen a partial rejection of internalism and the traditional deontological notions that motivate it. All internalist analyses of knowledge have fallen prey to effective counterexamples, and philosophers have consequently turned to epistemic “externalism” as a possible alternative. Such theories are driven by the idea that there must be some causal or reliable link, *external* to the cognizer, between beliefs and their corresponding facts in the world if knowledge is to obtain. Externalists assert that mere adherence to epistemic duties is not enough to produce knowledge. Rather, on this analysis, what is most important is that connection, however characterized, that holds between agent and fact which must obtain to make our beliefs true. Instead of talk about duties and obligations, externalists concentrate on reliable mechanisms and nomological relations between beliefs and the world. For the classical externalist, we need know nothing about the justification of our beliefs for them to constitute knowledge.

This may sound like a mere academic shift in emphasis: remember, we did agree that, of course, our beliefs must be true in order to qualify as knowledge. But such an assessment is too hasty. The crux of the division between internalists and externalists is not simply which part of the justified-true-belief equation is most important. Epistemology is supposed to tell us more than what knowledge is; its greater purpose is prescriptive, not descriptive. When one realizes this, the debate becomes deeper, and the gulf between internalist and externalist widens. On the one hand, the internalist prescription is hard epistemic work and diligence: uphold the standards of justification and ye shall be rewarded. On the other hand...well, whatever the positive externalist program, how well we maintain our standards of justification has perhaps little or nothing to do with what knowledge we take home at the end of the day. Here the externalist position faces the risk of running our epistemic project aground in a fatalistic muck. If knowledge is more the result of an external relation than of the consequences of our internal conduct, then

what is left for us to do? The externalist program is in danger of stripping away the autonomy of rational agents with regard to knowledge, despite our powerful internalist intuitions.

The externalist's problem, then, should not only be investigating the nature of the supposed epistemic connection in the analysis of knowledge between belief and belief-producing reality, but it must also provide us with a program for actively improving and maintaining our epistemic lot in life. Without such a program, externalism necessarily subverts its own epistemic cause and leaks away into some other discipline, maybe metaphysics, but certainly not epistemology.

Internalism has routinely failed to produce a satisfactory account of knowledge, largely due to externalist criticisms. But these criticisms, which are also the motivation for positive externalist accounts, leave externalists with a fatalistic approach to knowledge that is just as unsatisfactory as the earlier tradition. It remains to be seen if an externalist theory can accommodate the autonomy we, as rational agents, must demand of a complete analysis of knowledge.

A promising recent externalist theory is offered by Alvin Plantinga, who adopts a naturalistic, externalist approach in response to the continued failure of both traditional internalist accounts and more recent strains of externalism. A careful assessment of Plantinga's account offers, I will argue, an analysis of knowledge that, with certain necessary adjustments, provides an account that is sufficiently externalist to avoid internalist problems, while steering clear of the fatalism inherent in other externalist accounts.

Plantinga and Proper Function

Plantinga's theory of knowledge is externalist because it relies upon access to the world. Plantinga's book, *Warrant: The Current Debate*, begins by establishing the concept of "warrant," defined as "that...which together with truth makes the difference between knowledge and mere true belief" (WCD, 3). Rather than endorse the conception of knowledge as justified-true-belief,

with the addition of a fourth condition to satisfy Gettier problems, Plantinga abandons internalism altogether in favor of an essentially externalist notion based upon what he considers to be the most important, and most neglected, “epistemic value:” properly functioning cognitive faculties (WPF, 3).

Plantinga finds that on each internalist theory, be it classical internalism, or some form of coherentism, the epistemic value that cannot be accommodated is the proper function of those faculties involved in belief formation, such as perception, memory, reasoning powers, and the like. Plantinga faults existing externalist accounts, such as reliabilism, for the same exclusion. According to Plantinga, a dutiful internalist and a reliable externalist can both satisfy all respective conditions for knowledge and still lack warrant. The reason, in Plantinga’s view, is the possibility of cognitive malfunction, which no current theory takes into consideration. If our mental faculties are not working as they ought, we’ll never be judged to have knowledge no matter how hard we work at it, nor will the proper relation between us and the world be enough to ensure epistemic success.

Once identified as the missing element, proper function of cognitive faculties is the core of Plantinga’s positive account of knowledge in *Warrant and Proper Function*. Of course, properly functioning faculties, although perhaps necessary, are far from sufficient to provide warrant, which, when added to true belief, yields knowledge. As Plantinga notes, we often consider our faculties to be in perfect working order, and yet many of our beliefs certainly do not count as knowledge. Plantinga thus derives three additional criteria that, taken together, are supposed to provide a complete set of necessary and sufficient conditions for warranted belief.

Plantinga’s fleshed out theory is as follows:

S is warranted in believing a proposition P iff P is formed by:

- (i) properly functioning cognitive faculties that are,
- (ii) functioning within a cognitive environment close to that for which they were designed,
- (iii) aimed at truth, and
- (iv) in some sense reliable.

A belief is warranted for me if and only if, in addition to my cognitive faculties working as they ought (i.e. are functioning properly), I am within a cognitive environment that is appropriate for the design of those faculties. This condition is necessary to rule out those instances of belief formation during which my faculties, while internally functional, are nonetheless hindered by the environment in which I find myself. For example, although I have 20/20 vision, I cannot trust my eyes when I am in a completely dark room, or 30 meters underwater, or perhaps upon the surface of some foreign planet where only the invisible part of the light spectrum is available. In these instances, despite the fact that all my cognitive faculties might be performing perfectly, my beliefs would certainly not be warranted. Also, beliefs must be formed by those cognitive faculties which are aimed at truth. This is to exclude those belief forming mechanisms that, when functioning properly, are designed to benefit us in some other manner. For example, wishful thinking, which is arguably a faculty, might benefit us by creating false beliefs which keep us from becoming depressed over say, a terminal illness. This faculty is not eligible to generate warrant as it is not designed to produce true beliefs. Only faculties with this specific purpose, such as the ability to reason, can be the basis for warrant. Lastly, one's cognitive faculties must be in some sense reliable, which is to say that the design of one's faculties must be a good one.

All of the aforementioned conditions have been subsequent to Plantinga's primary criteria: properly functioning faculties. But, it is not yet clear what exactly is meant by 'proper function.' To explicate this concept, Plantinga introduces the idea of a design plan. He conceives of cognitive faculties as being like artifacts in so much as they must have a designer whose intentions determine what their function or purpose is, and what 'design plan' will result in the achievement of those ends. Plantinga writes that, essentially, "in the central and paradigm cases, design plans do indeed involve a thing's having been designed by one or more conscious designers who are aiming at an end of some sort..." (WPF, 21). In the case of humans, Plantinga posits God or, less attractive to himself, the evolutionary process of natural selection as possible designers of the

proper functions of human beings, (I will hereafter consider only natural selection as designer). On either view, Plantinga establishes the basis for the idea of a human design plan, which must be a good one if warrant is to exist.

This design plan, according to Plantinga, is a set of specifications that delineate how a thing will respond to certain circumstances so as to conform to the functions intended by the designer. For example, a properly functioning home furnace will kick on or off in order to maintain a preset temperature. When the temperature drops below a certain level, the furnace turns on. Hence, the design plan of an artifact can be described as a finite set of "circumstance, response, purpose or function triples" (WPF, 22) In the above example, the temperature is the circumstance, the turning on or off the response, and the maintenance of a preset temperature is the function. Plantinga holds that this notion of a design plan applies to human cognitive faculties as well, such that every natural cognitive mechanism can be said to have a function in accordance with the purposes of the evolutionary process (or God). And this purpose determines the design plan, which in turn delineates how our cognitive faculties should operate. Under Plantinga's account of proper function, then, cognitive faculties are indiscernible from other natural functions such as the pumping of blood, or the process of digestion.

At this point, we are in a position to assess Plantinga's account in terms of our primary prescriptive epistemic goals. What, according to Plantinga, are we to do in order to maximize the number of beliefs which count as knowledge? The answer is not explicitly clear, but it seems to be along the following lines. Decide which faculties meet Plantinga's conditions, and accept only those beliefs produced accordingly as warranted. Whatever remains, while perhaps useful in other circumstances, is not to count as knowledge. This seems straightforward enough, but there is still some fine tuning to be done before a final assessment can be made. First, we need to clarify Plantinga's notion of proper function. To do this, I will introduce the work of Larry Wright, who formalizes a definition of functions.

Wright identifies two types of functions, natural and conscious. Wright defines natural functions as "the common organismic ones such as the function of the heart" (Wright, 142). He

describes these by stating that “the natural function of something--say, an organ in an organism--is the reason the organ is there by invoking natural selection. If an organ has been naturally selected-for by virtue of something it does, we can say that the reason the organ is there is that it does that something. Hence we can say animals have kidneys because they eliminate metabolic waste...” (Wright, 142). It is important to note here that natural selection is a slow process, occurring across generations, and therefore a rather unique designer, with survival as its intended purpose. These things take time. For example, way back when, the taller amongst the giraffe ancestors supposedly got more food, lived longer, and were able to pass on their tall genes...fast forward a few millennia...and today all giraffes are quite tall. What is crucial to understand is that what can only be called a biological idiosyncrasy back then, is properly called a natural function in modern giraffes by virtue of the extensive selection process that, over the years, turned a few long necked beasts into an evolved species of all long-necked giraffes. Natural functions require an historical precedent, and that precedent has its basis in survival, or the passing on of genes. Thus we can formulate an analysis of natural functions as follows:

N: For any organism S, with attribute X, X has the natural function Y iff X has been naturally selected-for across generations *because* X accomplishes Y.

Contrast this account of natural functions with Wright’s characterization of conscious functions: these “commonly (although not necessarily) involve artifacts, such as the telephone...” (Wright, 142). Conscious functions, like the sorts of things that usually possess them, are radically different from their natural counterparts. First, the purpose of a conscious function can be just about anything--there’s one for every human artifact. Telephone’s have the function of communication, cars have the function of transportation, picture frames of decoration, etc. These common artifacts represent one aspect or type of conscious function. We, human artisans, specifically craft bits of our world out of raw material for various express purposes. This is the “if you build it, they will come” approach to functions. Lets call it C:

C: a thing X, has the conscious function Y iff Y is the reason X is there, and Y is accomplished because of X.

But, as alluded to in Wright's earlier quote, this is not necessarily the only way conscious functions come about. Sometimes we stumble across things that already fit some purpose of ours without any modification whatsoever. Think of these as belonging to the "ready to wear" approach to functions. For example, rocks make great paper weights, and can function as such if we decide to employ as such, even though no one designed them that way. Or, a better example, drawn from Wright's paper, the damaged assembly line. Wright describes an assembly line which acquires, quite by accident, a ripple in its belt which somehow detects all and only those items which it carries along that are defective in some way. No one designed it, but certainly the prudent factory owner would not replace the belt, and would probably take pains to preserve it. We can now amend C so as to accommodate this type of function as well. Let's call this C*:

C*: X has the conscious function Y iff Y is accomplished because of X, and X is actively selected-for maintained by a designer for the purpose of accomplishing Y.

In these special cases, a thing's function is independent of its physical organization; in fact, such items might be completely "un-designed" in this respect, owing their physical organization to no design or pre-intention whatsoever, as with the completely non-artifactual yet invaluable rock/paper weight. What crucially separates C* from N is that conscious functions are non-historical. Things with conscious functions do not need ancestors, things with natural functions do.

Returning to Plantinga, we can situate his properly functioning cognitive faculties in Wright's more sophisticated framework. Plantinga tells us that our cognitive mechanisms get the functions relevant to warrant, i.e. *natural* functions, from natural selection. So these functions must be historical, having developed across generations. Further, if the functions associated with warrant are natural functions, then the same must be the case for the belief forming functions had by our various cognitive mechanisms. Thus, a cognitive mechanism produces warranted beliefs according to Plantinga's account only when such mechanisms have, as the relevant epistemic function, an historical, natural function.

Problems in Plantinga's Account

I will argue in this section that Plantinga mistakenly ascribes the natural functions of our cognitive mechanisms as those functions producing warrant. Instead, I contend that those warrant producing functions are properly conceived of as socially created conscious functions. But, before I argue for such an ascription, let us consider a well-worn counter-example to Plantinga's account that, with a little tinkering, illustrates why the warrant producing functions cannot be natural functions.

The original counter-example is accredited by Plantinga to William Hasker, and asks us to imagine a person who undergoes a random genetic mutation resulting in a new biological mechanism which happens to produce true beliefs in a reliable fashion. Hasker and others have argued that beliefs formed by such a mechanism would be judged by us to be warranted even though the mechanism in question is nowhere provided for in any original design plan, thus making Plantinga's condition of proper function in accordance with a design plan unnecessary. I wish to take a slightly different tack and examine this example in greater detail.

Imagine someone, let's call him Peter, who wakes up one morning and discovers that he is totally deaf. But, Peter also finds himself experiencing a new type of tactile sensation through his skin which is amazingly interpretable by Peter as sound. In effect, Pete's ears are now useless flaps, but his skin has become sensitive to sound vibrations to the extent that he is able to process inputs through the skin and form appropriate beliefs based upon these inputs. He 'knows', for example, when the phone rings and, upon answering, is able to correctly identify the caller and carry on a conversation.

Now, on that first morning, are the beliefs which Peter forms based upon his new faculty judged by us to be warranted? Clearly not. First, he would have no way of determining the reliability of the new mechanism; in fact he could not confirm that he had acquired a new mechanism at all. Pete's new tactile/auditory experiences would have to be categorized along with other unproved phenomena such as telepathy or ESP. Now suppose that these experiences

persist, and suppose further that Peter rushes himself to Johns-Hopkins for testing. After months of extensive investigation, a team of experts from around the world concludes that his new experiences are the result of a legitimate, reliable biological mechanism resulting from a random genetic mutation. Now do we judge that Pete's beliefs based upon this new faculty as warranted? At this point they clearly would be, having been sanctioned by the experts as reliable.

But, and this is the problem for Plantinga, in what sense does this mechanism have a natural function? What would it be for the new mechanism to function properly or improperly? It cannot be that the new mechanism has a natural function of the sort posited by Plantinga because it is nowhere specified in Peter's original design plan and cannot therefore have a natural function. Plantinga argues that Pete simply got a new design plan. Remembering Wright's point, however, this is not an option. Natural functions require mechanisms with ancestors, while Pete's mutation has none. This is where Plantinga's theory breaks down. According to Plantinga, who only considers natural functions, there is no way of assigning any function, much less a proper one, to Pete's ahistorical mechanism.

Yet it seems intuitively correct to say that his new faculty does have a function--it provides reliable sense data about sounds, producing true beliefs about the environment. It further seems quite plausible to think of this faculty as functioning properly or improperly. Maybe, for example, it is the case that Peter's new "hearing" is impeded if his skin gets too dry, distorting the sense data. On occasions like this, his faculty would be functioning improperly. But from where does this function ascription derive?

We need not abandon Peter or Plantinga entirely. Let's rejoin Pete as he leaves Johns-Hopkins with the knowledge that the new mechanism is reliable. I contend that at this point, the mechanism still has no function. It is simply that, a mechanism. But Pete, as a sentient human being, has many epistemic goals--to form true beliefs, to avoid false beliefs, etc. He is now in a position, thanks to the experts at Johns-Hopkins, to reflect upon these goals, calculate the effectiveness of his new mechanism at achieving these goals, and epistemically *recruit* his new mechanism for the purpose of fulfilling them. At this point, his new mechanism can properly be

said to have a function--a *conscious* function, of the C* variety mentioned above, assigned by him to a natural mechanism for the purpose of achieving his own self-determined epistemic goals.

I further assert that, with regard to producing warranted beliefs, all cognitive faculties should be characterized as involving conscious functions in much the same way. One might argue against this position as follows: "Hey, Pete's funky skin is one thing, but as for the rest of us, our cognitive mechanisms are *not* random mutations, and they *did* evolve." I will not disagree. I will only say that, whatever the *natural* functions of those processes or mechanisms which we might describe as mental, these are completely independent of any *conscious* functions we may choose to put upon them. Moreover, in so much as our mental lives involve functions of producing warranted true beliefs, such functions are conscious.

Returning once again to natural selection, which we have established as the only source of a natural function, we must remember that natural functions are simply not in the business of belief ascription. If a function is to be natural, it must also be purely survivalistic.

Certainly our mental faculties do have the purpose of aiding in survival by allowing us to act in our own best interests and avoid bodily danger. Our cognitive faculties help us avoid enemies, find food, procure shelter and locate mates; but these activities are not the *direct* purpose of cognitive faculties, in so far as those faculties are viewed as having warrant production as their goal. Thus, the germane purposes of cognitive faculties are epistemic, not survivalistic. Moreover, natural selection holds no epistemic purposes; these are the domain of human cognizers, who determine their own epistemic goals. This is why cognitive functions must be conscious, as their corresponding purposes are epistemic goals held by the cognizant community.

A few brief examples illustrate this point. Consider the faculty of perception. While the visual apparatus is a natural mechanism, the function of vision, in so far as that function is to form true beliefs, is not natural. Forming true beliefs about the world is a conscious human purpose which, while intimately connected to survival, is not synonymous with survival; in so far as perception has only the purpose of forming true beliefs, it cannot be the purpose of natural selection. Contrast perception with the function of the kidneys. This function is properly a

natural function because its direct and immediate purpose is survival, as opposed to the human purpose of forming true beliefs. Processing metabolic waste is something that bodies do, forming true beliefs is something that cognizers do; thus it follows that the functions assigned to these activities must conform to the author of their intended purposes, one natural and the other conscious.

I offer one final illustration of a biological mechanism with a clear conscious function for two reasons. The first, to drive home arguments already made, and the second, to remind us of why we embarked upon this epistemic odyssey in the first place. Due to an injury years ago, my left knee aches in the hours before it rains. While my aching knee is not a one hundred percent accurate indicator of rain--it often fails to ache when it rains and sometimes aches for no reason whatsoever--it could have been that it were completely accurate and reliable. Assuming this were the case, and I my knee ached if and only if it were going rain, would my knee have acquired an additional function? It certainly could not be a natural function; aching-before-rain is not a useful natural mechanism, but rather an unpleasant side effect of injury. Such a phenomenon serves absolutely no natural purpose, and would by no means qualify as a natural function. But, it is reasonable to assert that my knee could take on a conscious function. Maybe I am a meteorologist, or a storm-chaser, or maybe I tend the courts at Wimbledon Stadium. Surely, the fact that my knee is one hundred percent accurate at indicating when it will rain fulfills a primary personal epistemic goal. Under these circumstances, I would certainly consider my knee as having the function of rain prediction; and this function would be of conscious origins, not natural.

In so much as their direct and immediate function is the formation of true beliefs, cognitive faculties are conscious functions in exactly the same way. The fact that such faculties are not the result of random genetic mutations or old sports injuries is incidental. Human beings establish what their epistemic goals should be and employ cognition as a means to these ends, regardless of any natural origins which may in fact obtain.

Conclusions

That Plantinga incorrectly counts cognitive faculties among natural functions results in an inadequacy in his theory which explains why it cannot handle the type of counter- example described above. It need not be the case that proper function be described purely in terms of historical natural functions which are the work the evolutionary process (or God). Either of these might be considered as having provided the raw natural and cognitive mechanisms responsible for the existence of those cognitive faculties. These are then employed by us to assure that, to the extent possible, our beliefs are warranted; but the proper or improper function of such faculties results from a conscious design plan, namely that epistemic plan which every cognizer sets for herself concerning her own epistemic goals. We have now saved externalism from fatalism. Here is the prescriptive aspect externalism seemed to lack.

As for Plantinga, his theory of warrant can now be reformulated in terms of conscious functions, rather than natural functions.

- S is warranted in believing a proposition P iff P is formed by cognitive faculties that are:
- (i) functioning properly in accordance with *conscious* functional ascriptions,
 - (ii) functioning within a cognitive environment close to that for which they were designed,
 - (iii) aimed at truth, and
 - (iv) in some sense reliable.

The facelift is descriptively minor, but prescriptively huge. His design plan should be not so much a blueprint of human natural mechanisms--leave that to the scientific community, or at least to the metaphysicians. Rather, Plantinga should focus on what is truly the realm of the epistemologist: an epistemic inquiry into a set of normative, yet scientifically grounded guidelines delineating what epistemic purposes we ought to pursue, and what cognitive faculties we ought to employ towards

those ends. Now we can save Plantinga and the rest of the externalists from fatalism and still maintain their essentially non-internalist flavor.

The original dilemma was how to save an externalist account without making it a fatalistic one. Plantinga's account, reformulated, accomplishes this. Although our brains must be in the right relation to the world in order to produce true beliefs, it is up to us as individual cognizers to sort through our mental capabilities and delineate those faculties which fit our epistemic needs and decide how best to employ them.

This is the project of epistemology. Plantinga would seem to implicitly agree here, as his conditions for warrant aside from properly functioning cognitive faculties regard such instances when we ought not to trust our cognitive mechanisms based upon what we know of how they work.

Such an enterprise is most properly the project of the epistemologist, whose concerns should not only be a description of how we know what we know, but more importantly, what types of epistemic purposes we ought to have and what actions are most likely to achieve those ends. Both components are integral to a complete analysis of knowledge. I suggest that an account like Plantinga's, properly reformulated, can ease the tension between externalism and internalism by addressing the problems and preserving the positive theoretical aspects of both. Such an account yields an analysis of knowledge that acknowledges the crucial link between the cognizer and her relationship to the environment in which beliefs are formed, while lending equal credence to the basic intuition that rational agents can actively exert control regarding the pursuit and maintenance of knowledge.

Works Cited

- Plantinga, Alvin. Warrant: The Current Debate. New York: Oxford University Press, 1993.
- Plantinga, Alvin. Warrant and Proper Function. New York: Oxford University Press, 1993.
- Wright, Larry. Functions. *Philosophical Review* 82: 139-168, 1973.

Angela Bomette, Philosophy
"Plantinga and the Theory of Knowledge"
proj. adv. Carl Fillett, —