Paul Thagard

WHY IS BEAUTY A ROAD TO THE TRUTH?

ABSTRACT. This paper discusses Theo Kuipers' account of beauty and truth. It challenges Kuipers' psychological account of how scientists come to appreciate beautiful theories, as well as his attempt to justify the use of aesthetic criteria on the basis of a "meta-induction." I propose an alternative psychological/philosophical account based on emotional coherence.

1. Introduction

In a recent article, Theo Kuipers (2002) offers an account of the relation between beauty, empirical success, and truth. Building on his impressive work on the nature of truth approximation (Kuipers 2000), he provides a "naturalistic-cum-formal" analysis that supports the contention of McAllister (1996) that aesthetic criteria are useful for scientific progress and truth approximation. I agree with this contention, but will challenge Kuipers' psychological account of how scientists come to appreciate beautiful theories, as well as his attempt to justify the use of aesthetic criteria on the basis of a "meta-induction." I propose an alternative psychological/philosophical account based on emotional coherence (Thagard 2000).

2. Kuipers on Beauty and Truth

According to Kuipers, the truth is beautiful in the sense that it has features that we have come to experience as emotionally positive due to the mere-exposure effect. This effect is a robust finding in experimental psychology that an increasing number of presentations of the same item tends to increase the affective appreciation of the item. Kuipers introduces the mere-exposure effect because it suggests that the human mind does a kind of affective induction in addition to the more familiar cognitive kind. Kuipers proposes that scientists do a kind of affective induction that leads them to react with positive emotions to recurring features of science that are not conceptually connected with empirical success, for example simplicity, symmetry, and visualizability.

In: R. Festa, A. Aliseda and J. Peijnenburg (eds.), *Cognitive Structures in Scientific Inquiry* (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, vol. 84), pp. 365-370. Amsterdam/New York, NY: Rodopi, 2005.

Assuming that there is indeed a correlation between such features and empirical success, the philosopher of science can then do a "cognitive meta-induction" that justifies scientists' affective inductions on the grounds that beauty really does correlate with truth. On this view, scientists acquire the tendency to find beautiful theories that possess features such as simplicity and symmetry on the basis of exposure to previous successful theories that had such features. Moreover, the acquisition is legitimate because, by the cognitive meta-induction, such features really do correlate with experimental success, which is an objective feature of theories. Kuipers not only tries to argue that the empirical success of theories signals their approximation to truth, but also that the correlating *non*-empirical features directly signal approximation to truth. Hence it is reasonable that scientists let themselves be guided by non-empirical features as well as empirical success.

I do not want to challenge Kuipers account of truth approximation, which strikes me as the most sophisticated currently available, but I see several problems with the way he connects beauty and truth. First, note that the mere-exposure effect is very different psychologically from affective induction. When mere exposure leads me to like something, the structure of the episode is: exposure to *X* → increased liking of *X*. In contrast, affective induction has a structure something like: *X* goes with *Y* and *Y* is liked → increased liking of *X*. Affective induction requires exposure to two features, e.g. simplicity and empirical success, whereas the mere-exposure effect does not require any such correlation. Hence the mere-exposure effect is logically and psychologically irrelevant to affective induction. I would not be surprised if human thinking does in fact use something like affective induction, but Kuipers needs to find empirical support for this kind of thinking from experiments other than those that support the existence of the mere-exposure effective.

Second, evidence is needed to support the claim that the positive emotional attitude toward simplicity and symmetry that many scientists exhibit is acquired by affective induction. Does scientific education really involve juxtaposition of aesthetic features and empirical success in ways that could lead budding scientists to acquire the emotional appreciation of simplicity and symmetry? In the first place, do scientists have an antecedent positive emotional attitude toward empirical success that would provide the basis of the affective induction that aesthetic features are good? I conjecture that science students acquire the tendency to find some theories beautiful through a partly innate and partly acquired ability to recognize coherence; the next section defends an emotional coherence account of aesthetic judgments in science. If this account is correct, then scientists acquire aesthetic attitudes by means different from affective induction.

Third, I am less confident than Kuipers about the connection between empirical success and truth. Even if there is a legitimate meta-induction connecting beauty and empirical success, it remains to be shown that there is a connection between empirical success and truth. On Kuipers view, the connection is direct, by virtue of the definition of approximate truth and the theorem that if *Y* is closer to the truth than *X*, then *Y* is at least as empirically successful as *X*. I agree that in general empirical success is a sign of truth, but it is hard to make the connection directly, since we have no independent way of establishing truth. This is concealed in Kuipers' framework because he identifies *the truth* as the strongest true theory rather than as how the world really is. In order to conclude that empirical success is a guide to how the world really is, we need to bring in other aspects of science such as its technological applicability, the substantial degree of agreement among scientists, and the largely cumulative nature of scientific development (Thagard 1988, ch. 8). In the past few hundred years, we have learned that empirical success is a much better guide to truth than other determinants of belief such as a priori reflection and divine inspiration, but it might have been otherwise. Hence the connection between empirical success and truth is just as much in need of argument as the connection between beauty and truth. The argument cannot be a cognitive meta-induction, because we have no way of identifying what is true. Rather, the form of argument is theoretical: we can infer that science acquires true theories because that is the best explanation of its technological success and largely cumulative development.

3. Beauty as Emotional Coherence

I will now sketch a different picture of the role of beauty in scientific inference. My most recent book develops a theory of emotional coherence that is used to explain how judgments of beauty arise (Thagard 2000, ch. 6). The theory extends a general theory of coherence as constraint satisfaction: when people make inferences, they do so in a way that maximizes coherence by maximizing the satisfaction of multiple positive and negative constraints among representations. The kind of inference most relevant to scientific thinking is explanatory coherence, in which the representations are of evidence and hypotheses, the positive constraints are based on explanation relations between hypotheses and evidence, and the negative constraints are based on relations of contradiction or competition between hypotheses. When scientists choose between competing theories, they do so by accepting those hypotheses that are part of the maximally coherent account. Various algorithms are

available for maximizing coherence, including psychologically plausible algorithms using artificial neural networks.

The theory of emotional coherence postulates that human thinking is a process that involves affective as well as cognitive constraints and that both kinds of constraint satisfaction are intimately related. Representations acquire *valences,* which constitute their emotional content, in addition to their degrees of acceptability. For example, your concept of *beer* involves in part a valence that represents whether or not you like beer. Propositional representations such as "Beer is good for you" also have a valence, as is evident in the different emotional reactions that might be given to this proposition from avid beer drinkers as opposed to those of teetotalers. From the perspective of emotional coherence theory, beauty is not a property of individual representations, but is a "metacoherence" property that arises as the result of a general assessment of coherence. A feeling of happiness emerges when most constraints are satisfied in a person's unconscious processing of cognitive and affective constraints, whereas feelings of sadness and anxiety can emerge when constraints are not satisfied. In particular, scientists find a theory beautiful when it is highly coherent with the evidence and with their other beliefs. Such coherence is largely a matter of empirical success, in that many of the constraints on a theory concern the data which it is intended to explain. But simplicity is intrinsically part of the coherence calculation, since the constraints that tie hypotheses with evidence are stronger if the explanations involve fewer hypotheses (see Thagard 1992, for a full exposition). Moreover, symmetry, which is another one of the aesthetic factors mentioned by Kuipers, is also a matter of coherence, of an analogical sort. Symmetry is a matter of having multiple parts of a theory or other set of representations that are analogous to each other (Thagard 2000, p. 203). For example, a face is symmetrical to the extent that the left side is analogous to the right side. Like explanatory inference, analogical thinking can be thought of in terms of satisfaction of multiple constraints (Holyoak and Thagard 1995).

In contrast to Kuipers, who views simplicity, symmetry, and analogy as problematic because they are nonempirical, I see them as an integral part of the coherence-based inferences about whether to accept or reject a theory. Beauty is the feeling that emerges to consciousness when a theory is very strongly coherent with respect to explaining the evidence *and* being consistent with other beliefs *and* possessing simplicity, symmetry, and other kinds of analogies. Psychologically, the beauty of a theory does not arise from affective inductions connecting aesthetic features with empirical success, but rather from the coherence of the theory that intrinsically includes those features.

4. Assessment

I have offered an alternative to Kuipers' psychological and philosophical explanations of why beauty is a road to the truth. Whose explanations are more plausible? First consider the competing psychological explanations of how scientists come to experience some theories as beautiful because of aesthetic features such as simplicity and symmetry.

Kuipers: Scientists come to like such aesthetic features because of a psychological mechanism of aesthetic induction akin to the mere exposure effect.

Thagard: Scientists find theories with such features beautiful because of their contribution to coherence which is inherently pleasurable.

There is currently little experimental evidence to enable us to discriminate directly between these two explanations; I have already argued that aesthetic induction is a very different process from the mere-exposure effect, so the considerable psychological evidence for the latter does not support the general plausibility of the former.

My main reason for preferring the emotional-coherence explanation of the pleasurable nature of simplicity and symmetry is that it derives scientific beauty from the same kind of psychological mechanism that produces intellectual pleasure in other domains, such as art, music, and mathematics. Aesthetic theorists such as Collingwood and Hutcheson, as well as mathematicians such as Hardy, have described beauty as deriving from unity, harmony, and coherence. Emotional coherence provides a unified (i.e. more beautiful!) explanation of scientific judgments of beauty, because it describes the same mechanism at work in science as in art and mathematics. Kuipers could well maintain that aesthetic induction on particular features operates in these other domains as well, which might serve to explain emotional preferences for particular kinds of art or mathematics. But aesthetic induction does not explain the general appreciation of beauty deriving from an overall appreciation of a work of art, a mathematical construction, or a scientific theory. In contrast, the theory of emotional coherence provides a specific computational mechanism by which positive feelings can emerge from global judgments of coherence, including ones that incorporate simplicity and symmetry.

I also think that the emotional-coherence account provides a better basis for the philosophical issue of justifying scientists' use of aesthetic judgments than Kuipers inductive account. Here are the two positions:

Kuipers: Scientists' use of aesthetic criteria such as simplicity and symmetry is justified by the cognitive meta-induction that these features correlate with empirical success and truth.

Thagard: Scientists' use of aesthetic criteria is justified more indirectly by the fact that they are integral to the coherence assessments that promote the largely cumulative development of theories, many of which are technologically successful.

I prefer the indirect strategy because it does not require the accumulation, by practicing scientists or by philosophers combing the history of science, of a large body of instances of correlations between aesthetic features and truth. It is also immune to the likely existence of counterexamples in the form of cases where theories that turned out to be false were initially adopted in part on the basis of aesthetic criteria. Judgments of scientific beauty, like all inductive reasoning, are highly fallible. My indirect method of justifying explanatory coherence assessment as scientific method does not assume that it always or even usually works, as meta-induction requires. Scientific reasoning, based on explanatory coherence and including judgments of beauty, is justified because it is sometimes successful and there is no other method that is anywhere near as successful in finding out how the world really is. Beauty is a road to truth, but the road can be a winding one.

In conclusion, I applaud Theo Kuipers for his development of elegant and plausible accounts of scientific reasoning and approximation to truth, and for his noble attempt to extend these accounts to explain the role of aesthetic judgments in science. But I have argued that the role of beauty in science is more fruitfully understood from the non-inductive perspective of emotional coherence.

University of Waterloo
Philosophy Department
Waterloo, Ontario ON N2L 3G1
Canada

REFERENCES

Holyoak, K.J. and P. Thagard (1995). *Mental Leaps: Analogy in Creative Thought*. Cambridge,MA: The MIT Press/Bradford Books.

Kuipers, T. (2000/ICR). *From Instrumentalism to Constructive Realism*. Dordrecht: Kluwer.

Kuipers, T. (2002). Beauty, a Road to the Truth. *Synthese* **131** (3), 291-328.

McAllister, J. W. (1996). *Beauty and Revolution in Science*. Ithaca, NY: Cornell University Press.

Thagard, P. (1988). *Computational Philosophy of Science*. Cambridge, MA: The MIT Press/BradfordBooks.

Thagard, P. (1992). *Conceptual Revolutions*. Princeton: Princeton University Press.

Thagard, P. (2000). *Coherence in Thought and Action*. Cambridge, MA: The MIT Press.

Theo A. F. Kuipers

AESTHETIC INDUCTION VERSUS COHERENCE

REPLY TO PAUL THAGARD

Paul Thagard's brief contribution deserves a long reply, but I confine myself here to some basic issues. I start with some concessions relative to SiS regarding simplicity and analogy, followed by rebutting Thagard's general and specific reserves about my recent naturalistic-cum-formal inductive account of the relation between beauty and truth. Finally, I raise some doubts about the exhaustiveness of his coherence account of that relation and its supposed incompatibility with my account.

Aesthetic Induction, Empirical Success, and Truth Approximation

Let me start by reporting some new considerations that are relevant to Thagard's contribution. In SiS I went as far as to claim that simplicity should only play a role in case of equal success (SiS, p. 238, and Section 11.2) and for analogy I saw no role at all (SiS, p. 297). Contrary to my previous beliefs, at the time of completion of SiS, very much stimulated by reading McAllister (1996), I was beginning to understand that there might be a relation between truth and simplicity, and, more recently, stimulated by a discussion with Thagard when he visited Groningen on the occasion of Alexander van den Bosch's promotion, even one between truth and analogy. Hence, in the light of my recent article on beauty and truth (Kuipers 2002), I have to qualify these claims in SiS.

Since "simplicity" figures, at least in certain periods of certain disciplines, in the prevailing aesthetic canon, to use McAllister's nice phrase, it has cognitive merits related to empirical success and even to truth approximation, which scientists favoring the dominant theory may value more than some empirical successes of a new theory that are failures of the old one. Repairs may well come to grips with these failures. Similarly, as McAllister (1996) also illustrates and my article implicitly justifies, "analogy" may also be seen

In: R. Festa, A. Aliseda and J. Peijnenburg (eds.), *Cognitive Structures in Scientific Inquiry* (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, vol. 84), pp. 371-374. Amsterdam/New York, NY: Rodopi, 2005.

as a nonempirical feature of certain theories that may play a cognitively justified role. Certainly, the relative weight assigned to such features should take into account that these features are based on "meta-induction," that is, induction of a recurring nonempirical feature correlating with empirical success, whereas general empirical successes are based on "object-induction," induction of a regularity about (the behavior of) a certain kind of objects. Although object-inductions are not very trustworthy, they are certainly more trustworthy than meta-inductions.

To be sure, the "uniform" notion of being "empirically more successful," as presented in ICR and SiS, leaving no room for empirical failures compensated by more impressive empirical successes, can be extended to the more general notion of "more successful," taking also "nonemprical" successes and failures uniformly into account. However, as explained in Section 6 of my article on beauty and truth, the interesting cases of nonempirical considerations come into the picture when they point in another direction than the empirical considerations. This would require a combined definition of 'more successfulness' taking relative weights of different kinds of considerations into account. Depending on one's weights, to use an example suggested to me by Thagard, one may then value the phlogiston theory or even the oxygen theory as less successful than the classical theory, according to which there are only four substances, viz., air, earth, fire, and water, because this theory is much simpler than the two famous competing theories.

I am happy to agree with Thagard's claim that my view of the relation between beauty and empirical success needs new experimental and historical evidence, although I would not say that the well established "mere-exposure effect" is irrelevant. In the article I argue that the aesthetic induction may be a variant of the mere-exposure effect, more precisely, a concretization, provisionally called a qualified-exposure effect. In line with its naturalized approach, I suggest at the end a number of experiments with normal and toy pieces of art and with scientific examples to establish the conditions and limitations of the effect. Moreover, further evidence for the varying character of the aesthetic canon when different phases or different research programs of the same discipline or of different disciplines are compared would strengthen the basic ideas around aesthetic induction as such and its diagnosis as a variant of the mere-exposure effect. Finally, as I also stress in my reply to Miller, in the companion volume, my refined claim about aesthetic induction can be falsified: determine a nonempirical feature which happens to accompany all increasingly successful theories in a certain area from a certain stage on and which is not generally considered beautiful by the relevant scientists. To be sure, the common interesting point of our diverging views is, of course, that both suggest (comparative) experiments and possible pieces of historical

evidence (see below), a rare but welcome aspect of primarily philosophical theories.

Apparently I did not convince Thagard by arguing in ICR (p. 162) that there is a direct connection between empirical success and truth, and that we do not need his detour, as I explained in SiS (p. 298). The crucial point seems to be that I identify *the truth* as the strongest true theory (given a domain and a vocabulary) "rather than as how the world really is." Here Thagard is transgressing the boundaries of my kind of constructive realism and enters some kind of essentialist realism. In the introductory chapter to this volume I summarize my direct argument for a relation between truth and empirical success. In my reply to Hans Mooij in the other volume I try to specify my metaphysical position in some more detail. Since Thagard's truth does not exist in my view, his detour argument, that empirical success is a sign of truth, essentially pertains to my non-essentialist kind of truth(s), like my direct argument.

Emotional Coherence

Let me now turn to Thagard's theory of beauty as an aspect of emotional coherence. According to him, "scientists find a theory beautiful when it is highly coherent with the evidence and with their other beliefs," where simplicity, symmetry and analogy (of which symmetry is a special case) are intrinsically part of the coherence calculation. In SiS (Section 11.2), I argue in general against Thagard's "unstratified" theory of explanatory coherence (and its implementation in the ECHO program), in favor of the stratified priority of explanatory superiority (implemented by the evaluation matrix EM), by using a meta-application of simplicity considerations. I show that both are equally successful in accounting for all historical choices provided and "prepared" by Thagard himself, whereas ECHO is much more complicated than EM. (See my reply to Vreeswijk.) In other words, Thagard's coherence theory asks for historical cases in which explanatory superiority is sacrificed to simplicity, which would go against the stratified view.

Thagard associates the beauty of theories with all kinds of coherence. Hence, incoherent aspects of theories should be seen as ugly. Thagard (2000, pp. 199-200) argues in general that symmetry is aesthetically appreciated for its contribution to coherence, and asymmetry is ugly due to its incoherence. He mentions the symmetry of (most) human faces, as opposed to the asymmetry of a misshapen face. This type of example is interesting for two reasons. First, after habituation to a misshapen face, e.g. of a movie star, we may come to find it very beautiful. Second, we are used to pictures of the arrangement of

organs in the human body, including all kinds of asymmetries, and many of us will find the composition very beautiful, not least for these asymmetries. Hence, an overall coherence account of beauty is difficult to combine with the fact that at least certain people appreciate incoherencies, including scientists. The biologist Stephen Gould, for example, stresses in an interview (Kayzer 2000) that he, in contrast to the physicist Steven Weinberg, counts diversity, unrepeatable contingencies and irregularities among the sources of his ultimate aesthetic satisfaction. Gould mentions as examples of great aesthetic satisfaction the diversity of a certain species of land snails, called cerions (p. 32), and the incoherencies in the revolutions of earth and moon, which make it impossible to design a coherent calendar (p. 29). Ironically enough, Weinberg (Kayzer 2000, p. 78; see also Weinberg 1993, p. 119) mentions the gravedigger scene in Shakespeare's Hamlet as a surprising intermezzo in a logical sequence of events, which, according to Weinberg, illustrates the fact that in the arts there are even higher aesthetic phenomena than in science. Hence, Gould's claim and examples seem to be incompatible with an overall coherence view of beauty in science, and Weinberg's example at least suggests that coherence cannot be the only source of aesthetic appreciation in the arts, which makes it difficult to understand why there would be no experiences of beautiful incoherencies in science.

In the last part of his contribution Thagard gives a very clear statement of our diverging psychological and philosophical explanations of why beauty is a road to the truth. However, from the above it will be clear that I am not yet converted to his view. But I would also like to stress that they may be less incompatible than Thagard suggests. First, as to the psychological side, overall coherence might well be a feature that in certain disciplines and at certain stages can belong to the "aesthetic canon" as the result of aesthetic induction. Second, as to the philosophical side, I have already indicated that Thagard's supposed indirect connection between beauty and the essentialist truth, that is, the truth about how the world really is, boils down to a connection between beauty and constructive truths, for which connection there is a direct argument which, as a matter of fact, has not been disputed by Thagard.

REFERENCES

Kayzer, W. (2000). *Het Boek over de Schoonheid en de Troost*. Amsterdam: Contact.

Kuipers, T. (2002). Beauty, a Road to The Truth. *Synthese* **131** (3), 291-328.

McAllister, J. (1996). *Beauty and Revolution in Science*. Ithaca, NY: Cornell University Press.

Thagard, P. (2000). *Coherence in Thought and Action*. Cambridge, MA: The MIT press.

Weinberg, S. (1993). *Dreams of a Final Theory*. London: Vintage.