

# Explanatory Coherence? Optimistic Induction at the Cost of Mysticism

Frank Zenker (frank.zenker@fil.lu.se)  
11 November 2009

*Abstract:* Paul Thagard (2007) has presented a truth-conducive account of explanatory coherence. It is naturalistically conceived and claims support from the historical development of empirical science. This account is here critically examined. It is attested to treat inadequately change-episodes in which a successor theory does not merely *extend*, but instead supersedes its predecessor(s), so-called ‘severe change’ (scientific-revolutions, theory dislodgement). It is argued that these historically regular, if infrequent, episodes demand a diachronic specification of (partial) theory-identity. In the absence thereof, they remain phenomena-to-be-explained in a complete account of truth-approximation (verisimilitude). Meanwhile, Thagard’s notions of ‘broadening’ and ‘deepening’ of empirical theories fail as inductively strong indicators of an epistemologically preferable, because presumably truth-approximating mode of theory-development. Therefore, his cautiously optimistic induction that science will continue to approximate truth must remain self-serving. At most, it is argued, our currently best empirical theories can lay a claim to differential empirical adequacy, but not to truth or approximation to it.

*Key words:* explanatory coherence, representation, truth-approximation, non-cumulative theory change, empirical theories

This paper argues for a more optimistic conclusion, that coherence of the right kind leads to approximate truth. (...) The right kind is explanatory coherence that involves theories that progressively broaden and deepen over time, where broadening is explanation of new facts and deepening is explanation of why the theory works. (...) This paper is not *about* naturalistic epistemology, but is an instance of it.

(Thagard 2007: 28f.)

## 1. Introduction

Largely following the most recent exposition of Thagard’s (2007) concept of explanatory coherence, this text develops the following three points of criticism:

- (1) Thagard understands diachronic relations between theories according to his synchronic view of reduction which systematically invokes explanation by lower level (“underlying”) mechanisms. Thereby, his use of the term ‘deepening’ also relates predecessor-successor theory pairs in so-called revolutionary change. However, it is at best unclear, for example, by what mechanism Newtonian gravitation is explained within General Relativity. In this very case, the successor theory offers neither a mechanism for, nor an explanation of gravitation.
- (2) What considerations of explanatory coherence may achieve, cognitive and sociological models of scientific development achieve already without invoking truth(-

approximation). Therefore, it remains unclear what the notion of ‘truth-approximation’ contributes above ‘comparative empirical adequacy’ which does not invoke truth.

- (3) Qualifying optimistic induction to a cautious level by weakening ‘truth’ to ‘approximate truth’ leaves room for expectable, but infrequent radical theory change. This room is a gap which is not motivated by the idea of explanatory coherence. Rather, the gap motivates the cautiousness of Thagard’s induction. If this gap can be filled, a naturalistic account of our ability to switch frames appears promising.

Section 2 presents the motivation for Thagard’s naturalized version of coherence and summarizes the main thrust of the criticism developed further below.

Section 3 reveals concessions with respect to abstract truth-bearers which Thagard incurs in preparation for the thesis that “(...) some kind of coherence theory of knowledge must be on the right track” (Thagard 2007: 31).

Section 4 follows Thagard into a balance between over-optimism and over-pessimism with respect to the outcome an inductive projection of past onto future scientific development, promising us “a more epistemologically satisfying induction that can tell us when we can take a coherent theory to be [approximately] true” (Thagard 2007: 34).

For purposes of contrast, section 5 introduces Lakatos’ (1978) approach to scientific change, to motivate why a proponent of explanatory coherence singles out “temporal properties of developing theories” (Thagard 2007: 35), which pertain also to large scale scientific changes (Kuhn’s revolutions). These properties then serve “(...) to identify a sense in which Newtonian gravitation is partly true (...)” (ibid.).

Section 6 shows candidacy for truth among empirical theories to be expressed in a mechanistically conceived deepening maxim, “(...) specified as the induction that theories can be judged to be [approximately] true if they have been deepened by having the mechanisms they describe decomposed into more fundamental mechanisms for which there is independent evidence” (Thagard 2007: 40). The critical reply is: To be useful, the mechanistic view lacks a definite stance on diachronic theory-reduction.

Section 7 points out a gap between Thagard’s identity criteria for theories which support the optimistic induction and criteria which hold across episodes of radical change. It is presupposed that, on pains of invoking mysticism, citing creativity or genius is unacceptable as an explanation of theory-dislodgement. This, however, appears to be the epistemological consequence of endorsing explanatory coherence *vis á vis* reconstructions that employ, for example, socio-psychological or cognitive models.

Section 8 offers a simple socio-economical explanation for the alleged inductive validity of broadening and deepening: Kuhn’s (1962) normal science, broadly construed. If this is accepted, then revolutionary science in its relation to truth (and any approximation to it) should factor with a greater weight into future inductive projections of current science than Thagard deems sufficient.

## 2. Naturalized Coherence

Thagard's (2007) account of explanatory coherence attempts to secure the significance of coherence from the perspective of naturalized epistemology. While claiming truth-conduciveness, i.e., approximation without reaching truth, a proponent of explanatory coherence can deny that coherence is truth, that coherence is irrelevant to truth, and that coherence leads to probability which leads to truth (Thagard 2007: 44).<sup>1</sup> Thagard takes "(...) coherence to be a relation among mental representations, including sentence-like propositions and also word-like concepts and picture-like images" (Thagard 2007: 29) which either describe or fail to describe the world up to some order approximation. The thrust behind explanatory coherence is a cautiously optimistic induction "(...) from the history of science, under the assumption that natural science is the major source of human knowledge" (Thagard 2007: 29). The induced thesis is: Scientific theorizing is not only (largely) technologically effective, but ultimately successful. According to Thagard, science progresses towards truth, in the following sense: A comparative increase in explanatory coherence indicates a comparative approximation to truth.

It seems only reasonable to assume that, in order to qualify as naturalistic, any such account must

- (i) square with the historical development of empirical theories as man-made representations *cum* artifacts – after all, these serve as best-praxis examples –, without
- (ii) locating knowledge of any comparative approximation to truth beyond the cognitive-constraints of theory users.

As for (ii), to ensure the message remains receivable for a human agent, the concepts of broadening and deepening of theories and their explanations are introduced (roughly: subsumption of new facts under a theory,  $T$ , and explanation of  $T$  by underlying mechanism(s) postulated in  $T$ ). The terms 'broadening' and 'deepening' function as names for singularly and jointly less than fully reliable indicators which are capable of largely delivering positive results when applied to modern scientific development. With decent inductive validity – this is the suggestion –, they serve as "pretty good" descriptors of preferable theory-development.

As for (i), some historical data will not fit. Below, we will see Thagard struggle with historically infrequent cases of new empirical theories which eventually become accepted within a scientific community. Such transitions between theories remain exceptions to the trend which broadening and deepening suggest. The exceptions thus provide the reason for Thagard's optimism to remain cautious (though no less inductive) *vis à vis* the following two constellations:

- (i) A non-broadened and non-deepened superseding theory,  $T'$ , is a better approximation to truth than its predecessor,  $T$ , because  $T'$  yields more accurate predictions than  $T$ .

---

<sup>1</sup> *Vis à vis* recent negative results with probabilistic models of coherence (Bovens & Hartman 2003, Olsson 2005, Olsson 2007), this is not a weak claim. Add to this Thagard's earlier concession that it takes a "(...) much fuller account of the conditions under which coherent theories can be said to approximate the truth" (Thagard 2000: 280). For an early version of this claim, see Thagard (1989).

- (ii) A superseded theory, *T*, had been partly true (or: had true parts) because the successor theory *T'* retained some of the vocabulary or mathematical structure of *T*.

The criticism to be developed is as follows: Weakening the induction to a degree of reasonable caution leaves the genesis (discovery) of a new and possibly radically differing empirical theory entirely unaccounted for. Therefore, considerations of explanatory coherence cannot provide guidance with respect to appreciating a new theory at the moment of its genesis *vis à vis* then-current theories, because Thagard has tied the ability to appreciate a new theory its having been broadened and deepened, resulting in a *post hoc* quality criterion. This criterion is conservative in the sense of privileging the current *status quo* over a (radical) deviation from it.

However, because one cannot (without begging the question) locate the *status quo* on a truth distance measure, any improvement over the *status quo* (in the sense of an increase in empirical adequacy) must consequently also lack such distance information. Therefore, one might demand of a proponent of Thagard's position that the mood of induction better remain neutral rather than cautiously optimistic. This at least is reasonable if, as Thagard does, one's approach is claimed to be a version of *naturalized* epistemology.

### 3. The Link between Coherence and Truth

On Thagard's account, coherence is relevant to truth, unless this relation is explicated as a coherence theory of representations.

The most audacious of these [accounts about the relation of coherence to truth] is the coherence theory of truth, according to which the truth of a representation consists of its coherence with other representations, not of its correspondence to a nonmental world (Thagard 2007: 29).

Such accounts, Thagard argues, will not serve to explicate the relation between coherence and truth. Rather, Thagard's claim is: Coherence is not a relation merely between representations; coherence is not constituted merely intra-mentally – is “not a purely mental matter” (Thagard 2007: 30).

*Pace* the side-remark on truth being explicated (by some) as “coherence in the mind of God” (Thagard 2007: 30). *Pace* Thagard's contention that none of our successful empirical theories<sup>2</sup> “reduces to rationalist *or* empiricist foundations” (Thagard 2007: 31; *italics added*).<sup>3</sup> And *pace* the problems of interpretation, realization, and implementation<sup>4</sup> which Thagard can see to burden the operationalization of coherence within a mathematical theory of probability, because the works of Olsson (2002) and Shogenji (2002) have

---

<sup>2</sup> “Our greatest epistemic achievements are scientific theories such as relativity theory, quantum theory, the atomic theory of matter, evolution by natural selection, genetics, and the germ theory of disease” (Thagard 2007: 31).

<sup>3</sup> An exclusive disjunction, presumably, because the thesis is either false otherwise, or trivially true in case no successful reduction can be achieved.

<sup>4</sup> The problems are (i) choosing the objectivist (frequentist) or the subjectivist (credence value/degree of belief) interpretation of probability, (ii) finding a reliable coherence measure, and (iii) implementing processes that halt after reasonable calculation time. See Thagard (2007: 31f.) and footnote 7, below.

shown as much.<sup>5</sup> The strongest opponent here is someone who prefers explicating the relation between coherence and truth such that truth-bearers are not conceived as mental representations, but as “eternal abstract entities” (Thagard 2007: 30), i.e., Platonic ideas. Along this route, empirical theories – naturalized to the man-made representations they are – remain non-starters. They compare to each other and, optionally, to an external world, but not to ideas.

At this point, Thagard offers too little to address his opponent in an acceptable manner.

The problem with this platonic reply is that *we have no evidence* that such abstract entities exist. In contrast, *there is ample evidence* from contemporary psychology and neuroscience that people employ mental representations, which therefore qualify as potential bearers of truth. (Thagard 2007: 30; *italics added*)

He argues from ignorance, only to concede: Such evidence-comparison does not rule out that the bearers of truth may very well be platonic entities, after all.

Of course, the fact that, as far as we know, there were no mental representations 10 billion years ago does not undermine the correspondence theory of truth, because we can consider the fit, or lack of fit, between current representations and the state of reality at that time. The key point against the coherence theory of truth is that coherence with currently available evidence supports the view that reality is independent of representation of it. (Thagard 2007: 30)

The phrasing of the key point, then, must remain unclear, because it relies on a use of the term ‘coherence’ that is undefined.<sup>6</sup> For the sake of argument, one may grant the thesis: The right explication of the relation between coherence and truth must be given in terms that not only intra-relate representations. Why? Because such an explication would, in a trivial sense, fail to describe the world – which is not a representation of itself – in case there is “a world independent of representations of it” (Thagard 2007: 29).

Furthermore, when construed naturalistically, but explicated in terms which reference only representations, then – insofar as humans shall be able to cognize it – , truth would find one condition of possibility in the emergence of cognizing beings (Thagard 2007: 29) – from “bio-soup”, one might add. So, shouldering this claim as a commitment in explicating the relation between coherence and truth comes with the burden of explicating emergence.

It appears, then, that the naturalist seeks to shift the onus of proof for emergence onto those who, in his view, make too much of representations or abstract ideas, in order to support the thesis: There is a true world independent of representations of it. Indeed: “My argument does not refute the [purely mental or abstract truth bearer-featuring] coherence theory of truth, but shows that it implausibly gives minds too large a place in constituting truth” (Thagard 2007: 30).

---

<sup>5</sup> Thagard’s characterization of this research is somewhat misleading. The referenced work made previously voiced assumptions precise, in order to answer open questions. This work has not been conducted in the spirit suggested when ascribing to their authors the view: “[I]t should be possible to establish a connection between coherence and truth by means of an intermediary connection between coherence and probability” (Thagard 2007: 31). This, rather, was the hypothesis, and it has been weakened.

<sup>6</sup> As a transcendental argument, i.e., an argument that seeks to secure its own presuppositions, the key point may be read as a key to the author’s main assumptions.

#### 4. Explanatory Coherence and not-so-right Induction

Principles E1-E7 (Symmetry, Explanation, Analogy, Data Priority, Contradiction, Competition, Acceptance) stated, rather loosely interpreted and their formal operationalization in computer models claimed<sup>7</sup> (Thagard 2007: 32f.), Thagard asks: “Do we have any reason to believe that a set of hypotheses that are accepted because they maximize explanatory coherence are at least approximately true?” (Thagard 2007: 33). His answer is: Rather not, at least not normally, not in every-day life, where some or most assess the explanatory coherence of a set of hypotheses in the improper manner, perhaps all the time.

True, “(...) if explanatory coherence had been assessed properly, the erroneous inference might have been avoided” (2007: 33). But this is statement is true primarily by presupposing an explication of explanatory coherence. On presumably the same (non-)explication, Thagard claims that: “[T]he history of science contains many cases in which theories high in explanatory coherence have turned out to be false” (2007: 33).

On this background, then, if one were to induce from the past to the future, then the mood could be deemed over-pessimistic (so called pessimistic meta-induction). After all, science provides decent empirical theories, although they may be called false in the senses of being anomaly-laden or of incurring contrary-to-fact conditions.<sup>8</sup> Moreover, contemporary theories, though anomalous, will expectably develop into better ones. Thagard respects this, but insists:

[W]e should not associate the maximization of explanatory coherence with truth. The history of science thus suggests that coherence, even explanatory coherence along the lines I have suggested, is a poor guide to truth. (...) [Yet] the temporal induction that recent theories will turn out to be true seems rather shaky [too], because there just might not have been enough time for superior theories to come along and demonstrate the weaknesses of current theories. (...) [At any rate] we need a more epistemologically satisfying induction that can tell us when we can take a coherent theory to be true. (Thagard 2007: 34)

This relegates the induction – cited to be that of Newton Smith (1981), and based on what the novice might consider “freak show-examples” of failed theories of the past (see Laudan 1980) – to the over-pessimistic end. Amongst those are the theories of crystalline spheres, phlogiston, and caloric (Thagard 2007: 33f.). The ensuing discussion of Whewell’s (1968) consilience of inductions – “a special kind of broadening”, so is Wilson’s (1998) (Thagard 2007: 36) – then relegates the very same to the over-optimistic end of the spectrum. After all, both the wave theory of light and Newtonian gravitation – not only Whewell’s favorite example – have been superseded (Thagard 2007: 35).

Thus, both moods of induction on behalf of the truth of empirical theories (over-optimism and over-pessimism) are argued to have already disqualified themselves. Tha-

---

<sup>7</sup> The operationalization of Thagard’s coherence theory is not our concern here. Briefly, under the E1-E7 constraints and given some algorithm, (hypothesis-/evidence-)statements are separated into an accepted and a non-accepted group, such that coherence increases for one group over the other. For criticism, see Amini (2003), centrally arguing that real life cases turn out to be NP hard.

<sup>8</sup> For example, General Relativity literally breaks down in the center of a black hole. Newtonian Mechanics treats objects as point-masses, rather than respecting their extension in space.

gard will emerge in the middle. The right kind of coherence goes along with the right kind of induction: a cautiously optimistic one.

## 5. Enter Lakatos

Lakatos grants the stronger thesis: All theories are born false; they do not eventually become false. As the pragmatic normal condition, rather than the pathological one, they also receive constant improvement; anomalies and certain non-sensical results are known, but tolerated.

Lakatos is well-interpreted to have adopted his own version of Whewell's consilience-criterion, though not necessarily as a criterion of truth or truth-approximation, which Thagard renders as: "Consilience requires a hypothesis to increase its explanatory coherence, not merely by explaining a new fact, but by explaining a kind of fact different from ones previously explained" (Thagard 2007: 35).

Moreover, where Thagard, with Whewell, uses the term 'theory', Lakatos uses 'research program', i.e., a series of theories. With his core vs. belt-distinction, one can refer to more or less retractable parts of an ultimately dynamic entity. This entity – and that will remain the problem for Thagard's account – does not add up to *one* theory over time.

Finally, Lakatos allows that degenerative programs (those having ceased to deliver theoretically progressive problems shifts which can be empirically confirmed) can "stage a comeback". Hence, the over-optimistic induction problem arises only on a Whewellian view, while the result of pessimistic induction is neutralized from the very beginning.<sup>9</sup>

One can enrich this background with historical information: The widely accepted story about the development of empirical science, specifically mathematical physics in the 20<sup>th</sup> century, is paradigmatically exemplified by the transition to the General Theory of Relativity from the (under limiting-conditions structurally converging, but semantically diverging, technically simpler and scope-wise smaller) Newtonian Theory of Gravitation. Now, the following quote from Thagard should sound much less novel than it otherwise does against his pages 28 to 35:

(...) I think he [Whewell] was on the right track in looking for temporal properties of developing theories that might mark them as good candidates for truth. At least we can say that the theory of universal gravitation and the wave theory of light are *not so totally false as* the theories of crystalline spheres, phlogiston, and caloric turned out to be. Below I will try to identify a sense in which Newtonian gravitation is partly true (...). (Thagard 2007: 35, *italics added*)

To draw a line between a progressive and a degenerative problem shift is to trace the difference between doing something only apparently new by means of something old vs. doing the same old plus something in addition with something apparently new. Lakatos offered a qualification for a progressive problem shift. He meant to rule out that theoretical progress (prediction of new facts, roughly) could be claimed on the basis of a mere empirically corroborated *reapplication* of an old part of a research program, or by

---

<sup>9</sup> Premised on the pessimist-optimist distinction, Lakatos undercuts it as a (dialectical) realist.

extending the old part into phenomena which had been among the intended applications of the theory all along.

On this background, we can now easily grasp the notion of a new theory that deepens an extant theory, without conflating the development of a research program with the dislodgement of one program by another.

## 6. Truth-Approximation as Coherence by Deepening

Thagard holds that, if explanatory coherence were assessed comparatively, such that historical stages of theory-development could be ranked according to their comparative degree of explanatory-coherence, then a predictive candidate is at hand to fill the gap between truth and coherence.

Instead of trying to defeat the pessimistic induction by arguing that the discarded theories are at least partly true, my strategy is to admit their falsehood and look for features that mark current theories as promising candidates to avoid joining phlogiston and caloric in the dustbin of history. (Thagard 2007: 38)

The strategy, then, is to assess if theories broaden and, crucially, deepen their explanations. Here, broadening pertains to subsuming more facts, and deepening pertains to subsuming explanations, such that “[e]xplanatory coherence leads to truth when a theory not only is the best explanation of the evidence, but also broadens its evidence base over time and is deepened by explanations of why the theory works” (Thagard 2007: 37).

However, Thagard is silent on the individuation of explanations and facts, but rests content with a cautiously optimistic induction which refrains from raising the above quote into the status of a universal generalization.<sup>10</sup>

It is remarkable that none of the theories that I discussed in connection with the pessimistic induction that theories turn out to be false were ever deepened. That is, no underlying mechanisms were identified for how entities such as phlogiston and caloric worked. (Thagard 2007: 37)

The immediate consequence is that – of all empirical theories (!) – Newtonian gravitation had not been successfully deepened before 1915.<sup>11</sup> Yet, next to deepening being allegedly observable in various places (modern chemistry, physics, medical theories, neuroscience, molecular biochemistry), our objection remains inconclusive. After all, we read: “The deepening maxim obviously does not apply to the most fundamental level in subatomic physics, but still has ample room for application in other areas of physics as well as chemistry, biology, and the social sciences (Thagard 2007: 39).

Hence, if one accepts Newtonian mechanics as basic physics at some point in time, the objection does not stick. But neither can it convince to read that General Relativity deepened Newtonian mechanics (Thagard 2007: 40). In general, the separation between what is basic and what is derived or reduced receives an inadequate treatment.

---

<sup>10</sup> “[W]e do not need a universal generalization here: it would be enough if we could show from a survey of the history of science that broadened and deepened theories rarely turn out to be false” (Thagard 2007: 37).

<sup>11</sup> However, there were several attempts at reducing (the ultimately unexplained) gravitation – *hypotheses non fingo* – to electro-magnetism throughout the 19<sup>th</sup> century. See Roseveare (1982), Jammer (1961).

I am not assuming the traditional philosophical view of reductionism, according to which the deeper theory serves to deductively predict what goes on at the higher level: the complexity and sensitivity to chaos of higher- and lower-level systems may make such predictions impossible to achieve. But we gain much understanding nevertheless by noting that the mechanism at the upper level works as it does because of the operations of the parts at the lower level. (Thagard 2007: 39)

This only characterizes synchronic reduction, while one needs a diachronic reduction concept to address historical developments. Moreover, it is questionable if any understanding is gained upon learning that general relativity (GR) renders space-time geometry as a function of gravitational mass distributions<sup>12</sup>, thus eradicating (not gravitation, but) gravitational forces? In Lakatos' terms, GR is a new research program featuring core assumptions which are incompatible with the Newtonian program.

## 7. Deliverance: Cautious Optimistic Induction

Having witnessed Thagard express a preference for mechanistic causal explanation, we are prepared for the inductively supported first version of the deepening maxim:

The deepening maxim can then be specified as the induction *that theories can be judged to be true if they have been deepened by having the mechanisms they describe decomposed into more fundamental mechanisms for which there is independent evidence*. As we have seen, inductive support for the deepening maxim includes the germ theory of disease, the neuronal theory of brains, molecular cell biology, molecular genetics, and the atomic theory of matter. Attention to mechanistic explanation serves to spell out and support my cautiously optimistic induction about the connection between coherence and truth. (Thagard 2007: 40; *italics added*)

The above maxim, which offers only a sufficient condition for the truth of a theory,<sup>13</sup> will be problematic, once the implications of Lakatos' approach are spelled out. On the core vs. belt distinction – which is rough, but can be refined, for example on the structuralist view of empirical theories (Balzer et al. 2000, Lenk 2002) –, one can separate minor from major revisions; the latter change core- or *a priori*-assumptions, e.g.,  $F = ma$ .<sup>14</sup>

As Thagard acknowledges, an approximately true theory may have parts all of which are false. Therefore, conceived of as a synchronic whole, an empirical theory remains fully revisable in all its parts, though each part is not necessarily equally revisable. Therefore, one will seek to further differentiate the falsity of parts, in order to speak of the parts of the successor-theory being less partly false (or more nearly true), than the parts of the predecessor-theory. This shows in the two-pronged result of Thagard's cautious optimistic induction:

First, it [the induction] allows for the possibility that a major instance of a deepened theory could turn out to be false. I do not expect fields such as molecular medicine, genetics, and atomic theory to be radically overthrown, but it could happen. Second, it allows for the possibility, which seems to

---

<sup>12</sup> Slogan: Geometry is gravitized, rather than gravitation geometrized. See Cao (1998: ch. 5).

<sup>13</sup> Stating a sufficient condition for the truth of a theory can be read to indicate the weakness of his position. I interpret the text stronger than the orthodox interpretation of 'if-then-sentences' licenses.

<sup>14</sup> On the *methodological a priori* status of  $F = ma$ , see Friedman (2000). Such can be considered semantic principles (meaning postulates) or, more psychologically, frame-conditions.

have happened in both the atomic theory of matter and Newtonian mechanics, that deepening by virtue of a *more fundamental mechanism* may lead to some revisions in the original theory, with recognition that it is only approximately (partly and nearly) true. (Thagard 2007: 41; *italics added*)

This consideration is included in the “final version of the deepening maxim” (Thagard 2007: 41) by weakening ‘truth’ to ‘approximate truth’:

If a theory not only maximizes explanatory coherence, but also broadens its evidence base over time and is deepened by explanations of why the theory’s proposed mechanism works, then we can reasonably conclude *that the theory is at least approximately true*. This induction is the strongest relation available between coherence and truth (Thagard 2007: 41; *italics added*).

Here, Thagard does not address the question, how a so-broadened and deepened theory – note that only the singular form is used – receives and, over time, maintains identity. True, “(...) it is useful to think of a theory not just as a set of sentences but rather as a representation of parts, properties, relations, and changes” (Thagard 2007: 41), but only offering criteria (“original hypotheses about parts, properties, relations, and changes remain substantially the same” (Thagard 2007: 41)) which are known to fail the (partial) identity required to speak of an underlying mechanism does not yield insight.

For example, in the case mentioned above, the critical claim is that GR does in no good sense underlie Newtonian mechanics. It is an alternative framework. Moreover, and independently of the foregoing, it is far from clear that the explanatory relation (if any) to GR’s historical predecessor is that of being more fundamental in the sense of an epistemic or a natural order (see Thagard 2007: 42). Worse yet, for the relation between the theory of Newton and that of Einstein, no non-metaphorical deepening mechanism has been offered.<sup>15</sup> Rather, having obtained this relation is treated as an act of genius (even by Kuhn<sup>16</sup>) and constitutes somewhat of an anomaly for the Nagel-style approach to reduction, which Thagard rejects.<sup>17</sup>

Failing to reconstruct the relation by means of which one recovers the old plus some more from something apparently new fuels the incommensurability debate. Generally, no deductive reduction-concept has captured such transitions well, while non-deductive (limiting case) reduction struggles with meaning shift, e.g., in the case of *mass*.<sup>18</sup> Thus, whoever wants to argue for (explanatory) coherence in its relation to truth has to provide an account of the relation which lets the successor theory appear as a more coherent version of its predecessor, while original hypotheses about parts, properties, relations, (perhaps also change) *are substantially different*.

A cognitive-historical line of research can be traced in the philosophy of science at least since Nersessian (1984), and including, e.g., Barker (2001), Shapere (1989), Andersen & Nersessian (2000), perhaps even Bird (2008). Put briefly, this line strongly suggests that the cognitive schemas necessary to “bridge” paradigms do not coincide with

---

<sup>15</sup> See the discussion on approximate reduction in Balzer et al. (1984), Balzer (1985) and Kuipers (2000). One can achieve a structural comparison of singular partial models and, thus, give an structural explication of limiting case reductions, but none of these cite (causal or explanatory) mechanisms.

<sup>16</sup> See Kuhn (1962: 119). The observation is in Nersessian (1992: 11).

<sup>17</sup> ‘Core-replacement’, an expression structuralists adopted from Lakatos (Stegmüller 1969), is only a seemingly neutral term which rather hides the triviality of one’s “reconstruction” of this relation.

<sup>18</sup> While it often goes unnoticed, Kuhn’s incommensurability thesis hinges entirely on ontological considerations, rather than those pertaining to mathematical structure. See Friedman (2002, 2003).

those of philosophical conceptual analysis. Nor are these schemas arbitrary. For example, applications of frame-theory, e.g., Chen (2003a, b) which is adopted from psychology (Barsalou 1992) suggest that systematicity can be recovered when studying the replacement of theory-parts without resorting to gestalt-shift. As Kuhn himself had to admit later in life, it fails to enlighten the issue.

All the same, the analyst is forced to acknowledge the *integration* of old parts into new wholes.<sup>19</sup> This is a process of change for which the term ‘dialectic’, at least for Lakatos, would seem fitting. Rather than explain this process, by endorsing explanatory coherence, we gloss over it and, as the final section brings out, are prone to do so by citing ‘genius’, a term which should be *taboo* for the naturalist.

## 8. Beyond deep

Thagard’s final discussion on “deepening the deepening maxim” appears altogether rather odd. In defense of Peter Railton, who is quoted to have suggested that “the most plausible explanation of why deepened theories survive and thrive is that they are at least approximately true” (Thagard 2007: 42), the following may be entertained: Any empirical theory will continue to save some corroborated phenomena (within an acceptable error bound) and thus retain some of its intended applications, as long as theorizing these phenomena is relevant for a society that declares to rely on technology. By and large, *via* engineering, empirical theories give rise to profitable use. Politically, present science is expected to do so in the long run. Why then should one believe that – rather than having more accurate, encompassing, fruitful and empirically adequate theories – one has gotten closer to truth?<sup>20</sup>

Consider moreover that contemporary physics seeks theories of quantum-gravity, because for that part of physical reality one has managed to get reasonable response from within the frameworks of the fields of gravitation, electro-magnetism, strong and weak nuclear interaction, it is true: Upon superposition, these fields no longer cohere.<sup>21</sup>

On Thagard’s view, should those fields which are presently postulated as predecessor theories be brought to cohere at some point in the future, then we may cautiously optimistically induce that we will have seen this “deepening on us” – unless, of course, we won’t have seen any deepening. In the latter case, however, according to Thagard, a new theory could still “have truth” or be comparatively closer to it. In brief, we better stay open to non-deepening candidates for truth. Alas, would we recognize them for their quality? Likely not. Would we call the genesis creative or mystic? Once we manage to deepen a theory: likely creative.

If that which integrates a preceding state of theorizing is approximately more true than the integrated state of theorizing depends on what it shall mean for this developmental process to not be false. This requires a naturalized sense of historical progress and

---

<sup>19</sup> For an approach to theory development which models such integration as a form of change within a topological account of concepts (conceptual spaces), see Gärdenfors and Zenker (2009).

<sup>20</sup> As Bengt Hanson points out (personal communication), the current organizational structure of science (short-term third party funding administered under a science-political impetus for interdisciplinarity) sustains and therefore explains the presence of deepened and broadened theories. The genesis of deepening and broadening, however, cannot be explained science-sociologically.

<sup>21</sup> In other diction: Superpositioning these fields entails massive coherence loss (see Thagard 2000: ch.2).

regress – compare Lakatos’ degenerating research programs, i.e., those which can no longer empirically confirm a theoretical problem shift. This sense is not readily available.

Without this sense, however, the absurdity is the following: As optimistic inducers who exercise caution, we ultimately keep mystified an infrequent, but expectably regular process which belongs to the development of empirical theories, a process which – of all epistemological positions – should be fully reconstructable for a naturalist.

## 9. References

- Amini, M. (2003), “Has Foundationalism Failed? A critical review of Coherence in Thought and Action by Paul Thagard,” *Human Nature Review* **3**, 119-123.
- Andersen, H. & N.J. Nersessian (2000). “Nomic Concepts, Frames, and Conceptual Change,” *Philosophy of Science* **67**, 224–241.
- Balzer, W., D. A. Pearce & H.-J. Schmidt (eds) (1984), *Reduction in Science. Structure, Examples, Philosophical Problems*. Dordrecht: Reidel.
- Balzer, W. (1985), “Incommensurability, Reduction, and Translation”. *Erkenntnis* **23**: 255–267.
- Balzer, W., C. U. Moulines & J. D. Sneed (eds) (2000), *Structuralist Knowledge Representation. Paradigmatic Examples*. Amsterdam: Rodopi.
- Barker, P. (2001), “Incommensurability and Conceptual Change During the Copernican Revolution,” in: H. Sankey and P. Hoyningen-Huene (eds). *Incommensurability and Related Matters*. Boston: Kluwer, 241–271.
- Barsalou, L. (1992), “Frames, Concepts, and Conceptual Fields,” in: A. Lehrer and E. Kittay (eds). *Frames, Fields and Contrasts: New Essays in Semantical and Lexical Organization*. Hillsdale, NJ: Erlbaum, 21–74.
- Bird, A. (2008), “Incommensurability naturalized,” in: Soler, L, H. Sankey & P. Hoyningen-Huene (eds.), *Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities*. Dordrecht: Springer, 21–39.
- Bovens, L. & S. Hartmann (2003), *Bayesian Epistemology*. Oxford: Clarendon Press.
- Cao, T. Y. (1998), *Conceptual Developments of 20<sup>th</sup> Century Field Theories*. Cambridge: CUP.
- Chen, X. (2003a), “Why Did Herschel Fail to Understand Polarization? The Differences Between Object and Event Concepts,” *Studies in History and Philosophy of Science* **34**: 491–513.
- Chen, X. (2003b), “Object and Event Concepts: A Cognitive Mechanism of Incommensurability,” *Philosophy of Science* **70**: 962–974.
- Friedman, M. (2000), “Transcendental Philosophy and A Priori Knowledge: A Neo-Kantian Perspective,” in: Boghossian, P. & C. Peacocke (eds), *New Essays on the A Priori*. Oxford: Clarendon Press, 367–384.
- Friedman, M (2002), “Kant, Kuhn and the Rationality of Science,” *Philosophy of Science* **69**, 171–190.

- Freidman, M. (2003), "Kuhn and Logical Empiricism," in: Nickles, T. (ed.), *Thomas Kuhn*. Cambridge: Cambridge University Press, 19–44.
- Gärdenfors, P. & F. Zenker (2009), "Using conceptual spaces to model the dynamics of empirical theories," in: Olsson, E.; S. Rahman, T. & Tullenheimo, T. (eds), *Science in Flux: Philosophy of Science Meets Belief Revision Theory* (Logic, Epistemology, and the Unity of Science-Series). Berlin: Springer (to appear).
- Jammer, M. (1961), *Concepts of Mass in Classical and Modern Physics*. Cambridge, MA: Harvard University Press.
- Kuhn, T. (1962), *The Structure of Scientific Revolution*. Chicago: CUP.
- Kuipers, T. (2000), *From Instrumentalism to Constructive Realism*. Dordrecht: Kluwer.
- Lakatos, I. (1978), *The Methodology of Scientific Research Programs*. Cambridge: CUP.
- Laudan, L. (1981), "A Confutation of Convergent Realism," *Philosophy of Science* **48**: 19–49.
- Lenk, H. (2002), Epistemological Remarks concerning the Concepts 'Theory' and 'Theoretical Concepts'. *Protosociology* **17**: 171–187.
- Newton-Smith, W. H. (1981), *The Rationality of Science*. London: Routledge and Kegan Paul.
- Nersessian, N. (1984), *Faraday to Einstein: Constructing Meaning in Scientific Theories*. Dordrecht: Nijhoff.
- Nersessian, N. (1992), "How do scientists think?" In: Giere, R. (ed.), *Cognitive Models of Science* (Minnesota Studies in the Philosophy of Science Vol. 15). Minneapolis: University of Minnesota Press, 3–41.
- Olsson, E. J. (2002), "What Is the Problem of Coherence and Truth?" *Journal of Philosophy* **99**: 246–272.
- Olsson, E. J. (2005), *Against Coherence: Truth, Probability, and Justification*. Oxford: OUP.
- Olsson, E. J. (ed.) (2007), "Editor's Introduction, Coherence and Truth: Recovering from the Impossibility Results". *Synthese* **158**, 267–274.
- Roseveare, N.T. (1982), *Mercury's perihelion, from Le Verrier to Einstein*. Oxford : Clarendon Press.
- Shapere, D. (1989), "Evolution and Continuity in Scientific Change," *Philosophy of Science* **56**: 419–37.
- Shogenji, T. (1999), "Is Coherence Truth-Conducive?" *Analysis* **61**: 338–345.
- Stegmüller, W. (1969). *The Structuralist View of Theories*. Berlin: Springer.
- Thagard, P. (1989), "Explanatory Coherence," *Behavioral and Brain Sciences* **12**: 435–467.
- Thagard, P. (1992), *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.
- Thagard, P. (2000), *Coherence in Thought and Action*. Cambridge, MA: MIT Press.
- Thagard, P. (2007), "Coherence, Truth and the Development of Scientific Knowledge", *Philosophy of Science* **74**: 28–47.
- Whewell, W. (1968), *William Whewell's Theory of Scientific Method*. Pittsburgh: University of Pittsburgh Press.

Wilson, E. O. (1998), *Consilience: The Unity of Knowledge*. New York: Vantage.