

Explanatory Coherence? Optimistic Induction at the Cost of Mysticism¹

Frank Zenker

University of Lund

Department of Philosophy & Cognitive Science

Kungshuset, Lundagård

22 222 Lund, Sweden

frank.zenker@fil.lu.se

Abstract: Paul Thagard (2007, “Coherence, Truth and the Development of Scientific Knowledge,” *Philosophy of Science* **74**, 28–47) has presented a truth-conducive account of explanatory coherence. It is claimed to be naturalistic and to be supported by the historical development of modern empirical theories. This account is attested to offer an inadequate treatment of changes by 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 such changes demand a diachronic specification of (partial) theory-identity. Else, they remain phenomena-to-be-explained in a complete account of truth-approximation (*verisimilitude*). Meanwhile, “broadening” and “deepening” of empirical theories fail as

¹ An extended version of this paper is available at www.frankzenker.de.

inductively strong indicators of an epistemologically preferable (because presumably truth-approximating) mode of theory-change.

Key words: coherence, explanation, naturalism, representation, truth-approximation, 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, p. 28f.)

1. Introduction

Addressed to readers acquainted with Thagard's (2007) exposition of explanatory coherence, we forward three lines of criticism:

- (1) Thagard understands diachronic relations between empirical theories analogously to his synchronic view of reduction. This invokes explanation by lower level ("underlying") mechanisms. Thus, the technical term 'deepening' also relates predecessor-successor theory pairs in *revolutionary change*, e.g., Newtonian

Mechanics and General Relativity. However, in this very case, the successor offers neither a mechanism for, nor an explanation of gravitation or of gravitational *forces*.

- (2) In qualifying optimistic induction to a cautious level by weakening ‘truth’ to ‘approximate truth,’ one apparently “leaves room” for – infrequent, but expectable – *radical* theory change (Kuhn’s “revolutions”). This gap is not motivated by explanatory coherence. Rather, the gap motivates induction to remain cautious. As a “gap filling” strategy, a naturalistic account of our ability to “switch” theoretical frameworks is at least equally promising.
- (3) It is unclear what considerations of explanatory coherence contribute above ‘comparative empirical adequacy,’ provided cognitive and sociological models already explain scientific development without invoking truth(-approximation).

2.1 Naturalized Coherence

Following Thagard, a proponent of explanatory coherence may claim truth-conduciveness, yet deny that coherence is truth, that coherence is irrelevant to truth, and that coherence leads to probability which leads to truth (p. 44).² The thrust behind explanatory coherence is a

² *Vis à vis* recent negative results with probabilistic models of coherence (e.g., Bovens & Hartman 2003; Olsson 2002, 2005, 2007;

cautiously optimistic induction “(...) from the history of science, under the assumption that natural science is the major source of human knowledge” (p. 29). The induced thesis is: ‘Scientific theorizing is not merely technologically effective, but ultimately successful.’ Science progresses towards truth in the sense of a comparative increase in explanatory coherence indicating a comparative approximation to truth.

To qualify as naturalistic, the account must square with the historical development of empirical theories as man-made representations *cum* artifacts (measuring instruments). After all, knowledge of comparative truth-approximation may not in principle lie beyond the *cognitive* constraints of theory users. Ensuring the “message remains receivable,” Thagard introduces *broadening* and *deepening* (roughly: subsumption of new facts under a theory, *T*, and explanation of *T* by underlying mechanism(s) postulated in *T*’).

These terms designate less than fully reliable indicators of *epistemologically preferable* theory-development. Their partial

Shogenji 1999), this is not a weak claim. Thagard’s characterization of this research is fit to mislead. It made previously voiced assumptions precise and has not been conducted in the spirit of: “[I]t should be possible to establish a connection between coherence and truth by means of an intermediary connection between coherence and probability” (p. 31). Rather, that was the hypothesis – now weakened. For an early version of explanatory coherence, see Thagard (1989).

unreliability arises from historically infrequent cases of *new* empirical theories eventually reaching wide acceptance within a scientific community. These cases provide the grounds for Thagard's cautious (though equally inductive) optimism *vis à vis* 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 empirically more accurate predictions than *T*. (ii) A superseded theory, *T*, had been partly true (or: had true parts), because *T'* retains some of the vocabulary or mathematical structure of *T*.

2.2 The Link between Coherence and Truth

On the account criticized, coherence is relevant to truth, unless this relation is explicated as a coherence theory of representations. Coherence, we learn, is not a relation merely between representations – is “not a purely mental matter” (p. 30). The strongest opponent to Thagard's position will find that truth-bearers are “eternal abstract entities” (p. 30), i.e., Platonic ideas. So, empirical theories – naturalized to man-made representations – would be non-starters. Representations compare to each other and to an external or possible world, not to Platonic ideas.

To this opponent, Thagard argues *ad ignorantiam* when conceding that truth-bearers *may* be Platonic entities (*ibid.*). Further, we learn, when construed naturalistically, but explicated in terms which

only reference representations, truth would find a condition of possibility in the emergence of cognizing beings (p. 29). So, shouldering this claim as a commitment in explicating the relation between coherence and truth would seem to incur the burden of explicating *emergence*. A critic might find, this move presents the naturalist as (illegitimately) seeking to shift the *onus of proof* for emergence, in support of the thesis ‘There is a true world independent of representations of it.’

2.3 Explanatory Coherence and “Not-So-Right” Induction

In induction, one’s mood may be over-pessimistic (*pessimistic meta-induction*). After all, science provides decent empirical theories, although they are *false* (in part or whole), because anomaly-ridden or citing contrary-to-fact conditions.³ Moreover, contemporary theories expectably develop into better ones. In acknowledgement, Thagard maintains “we need a more epistemologically satisfying induction that can tell us when we can take a coherent theory to be true” (p. 34).

He relegates this mood of induction – cited as Newton Smith’s (1981), and based on what novices could consider “freak show-examples” of failed theories of the past (Laudan 1981) – to the over-pessimistic end. As freaks appear the theories of crystalline spheres, phlogiston, and caloric (p. 33f.). The ensuing discussion of Whewell’s

³ For example, General Relativity breaks down in the center of a black hole; Newtonian Mechanics treats all objects as point-masses.

(1968) consilience of inductions – “a special kind of broadening” (p. 36), likewise for Wilson’s (1998) – relegates the latter to the over-optimistic end. As fools appear the wave theory of light and Newton’s Theory of Gravitation (p. 35).

So, both moods of induction are disqualified. Thagard emerges in the middle: The right kind of coherence demands the right kind of induction – a cautiously optimistic one.

3. Enter Lakatos

Lakatos (1978) grants the stronger thesis: Theories are *born* false; they do not eventually *become* false. Moreover, as a normal rather than a pathological condition, they also receive constant improvement (read: change). Anomalies and non-sensical results are tolerated. Lakatos is well-interpreted to adopt a version of Whewell’s consilience-criterion, though not as a criterion of truth or truth-approximation.

When Thagard uses ‘theory,’ Lakatos might have suggested ‘research program,’ i.e., a *series* of theories. His core *vs.* belt-distinction refers to differentially retractable parts of an ultimately dynamic whole. This entity does not add up to one theory over time. Importantly, *degenerative programs* (having ceased to deliver theoretically progressive problems shifts that are empirically confirmed) can “stage a comeback.” In brief, the over-optimistic induction problem does not arise;

the “result” of (over-)pessimistic induction is neutralized from the beginning.

One may enrich this background with historical information. The widely accepted account of the development of empirical science (specifically mathematical physics in the 20th century) is paradigmatically exemplified by the transition to the General Theory of Relativity from the – under limiting-conditions structurally converging, semantically diverging, technically simpler and scope-wise smaller – Newtonian Theory of Gravitation. Now, the following quote may sound far less novel:

(...) 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 (...). (p. 35, *italics added*)

Separating a progressive from 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 by means of something apparently new. Lakatos offered a qualification for a progressive problem shift. Theoretical progress (prediction of *new* facts) may *not* be claimed on the basis of a mere empirically corroborated *reapplication* of an old part of a research program, or on the basis of extending the old part into *already intended applications*.

This background provides a clearer understanding of a new theory *deepening* an extant theory. We avoid mistaking the development of a program for its dislodgement by another program.

4.1 Truth-Approximation as Coherence *qua* Deepening

To assess explanatory coherence comparatively is to rank historical stages of theory-development according to their comparative degree of explanatory-coherence. Thereby, a predictive candidate might seem at hand that could fill the gap between truth and coherence. The strategy is to assess if theories broaden and, crucially, deepen their explanations (see p. 37). Regrettably, Thagard is silent on the *individuation* of explanations or facts. Content with a cautiously optimistic induction, and refraining from a universal generalization,⁴ he states:

[i]t 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. (p. 37)

The immediate consequence: Newtonian gravitation had not been successfully deepened before 1915.⁵ Alas, deepening is allegedly observ-

⁴ “[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” (p. 37).

⁵ One may note attempts at reducing gravitation to electro-dynamics in the 19th century (Jammer 1961, Roseveare 1982, Zenker 2009).

able in various places (modern chemistry, physics, medical theories, neuroscience, molecular biochemistry). Alas, the objection remains inconclusive. After all, “[t]he deepening maxim obviously does not apply to the most fundamental level in subatomic physics (...)” (p. 39). But neither can it convince to read that General Relativity deepened Newtonian Mechanics (p. 40). In general, the separation between what is basic and what is reduced (or derived) remains unclear.

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: (...) 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. (p. 39)

This characterizes *synchronic* reduction, where only diachronic reduction could address historical development. Moreover, the sense in which “much understanding” is gained should be questioned. General Relativity (GR) renders space-time geometry as a function of gravitational mass distributions.⁶ GR eradicates not gravitation, but gravitational *forces*. In Lakatos’ terms, it is a new research program, featuring core assumptions incompatible with the Newtonian program.⁷

⁶ Slogan: Geometry is gravitized, rather than gravitation geometrized. See Cao (1998, ch. 5).

⁷ In explicating the *core* of GR, $F = ma$ undermines deductive consistency.

4.2. Cautious Optimistic Induction

Lakatos' core vs. belt distinction has been refined in the structuralist view (Balzer et al. 2000, Gähde 2002, Lenk 2002). Here, one may separate a minor from a major revision. The latter changes core or *a priori* assumptions, e.g., $F = ma$.⁸ As Thagard acknowledges, an approximately true theory *may* have parts all of which are false. Therefore, as a synchronic whole, an empirical theory remains fully revisable in all its parts. These parts may be *differentially* revisable. And after differentiating the falsity of parts, one could meaningfully say: Parts of a successor-theory are less partly false (or more nearly true) than parts of a predecessor-theory. Thagard seems to allow as much:

First, it [the induction] allows for the possibility that a major instance of a deepened theory could turn out to be false. (...) Second, it allows for the possibility (...) 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. (p. 41; *italics added*)

Evidenced in the “final version of the deepening maxim” (p. 41), the last consideration motivates 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 (p. 41; *italics added*).

⁸ The “neutral” term is ‘core-replacement,’ an expression adopted from Lakatos (Stegmüller 1969).

But note the open question: How does a broadened and deepened theory (only the singular is used) receive and, over time, maintain identity? It cannot help to offer criteria – “original hypotheses about parts, properties, relations, and changes remain substantially the same” (p. 41) – that are known to fail the (partial) identity required.

The immediate upshot is: General Relativity does in no good sense *underlie* Newtonian Mechanics, but is an alternative (if related) framework. It is far from clear that the explanatory relation (if any) to GR’s predecessor is *being more fundamental* in the sense of an epistemic or a natural order (p. 42). In characterizing the relation between Newton’s and Einstein’s theories, no non-metaphorical deepening mechanism has been offered.⁹ Rather, having obtained this relation is regularly treated as an act of “genius” (even by Kuhn¹⁰) and, perhaps, constitutes an anomaly for Nagel-style reduction, which Thagard rejects.

Failing to rationally reconstruct the relation by means of which one recovers the old plus some more from something apparently new

⁹ Approximate reduction, as discussed in Balzer et al. (1984), Balzer (1985), Kuipers (2000), and Batterman (2003), is instructive as to the epistemological cost of comparing at the level of *singular* partial models (*sans* theoretical “super-structure”), in a structural explication of limiting case reduction. None cite mechanisms.

¹⁰ See Kuhn (1962, p. 119); the source is Nersessian (1992, p. 11).

fuels the incommensurability debate.¹¹ No deductive reduction-concept has captured such transitions well. In contrast, non-deductive (limiting case) reduction rather confirms meaning shift, e.g., in the case of *mass*. *Nolens volens*, endorsing explanatory coherence in its alleged relation to truth incurs accounting for the relation which holds when a successor theory is the “more coherent” version of its predecessor, provided original hypotheses about parts, properties, relations (perhaps also change) *are substantially different*.

A cognitive-historical line of research can be traced in the history and philosophy of science.¹² It suggests that the *cognitive* schemas necessary to “bridge” frameworks (paradigms) do not coincide with classical conceptual analysis – nor are schemas arbitrary. Applications of frame-theory (e.g., Chen 2003a, b), adopted from psychology (Barsalou 1992), suggest that systematicity in the replacement of theory-parts can be recovered – without citing gestalt-shift (Zenker 2010).

¹¹ See Friedman (2002, 2003), cf. Thagard (1992). Arguably, the “incommensurability thesis” pivots on ontology, not structure.

¹² At least since Nersessian (1984), including, e.g., Barker (2001), Shapere (1989), Andersen & Nersessian (2000), perhaps Bird (2008).

Both analysts and “producers” of scientific history already acknowledge the *integration* of old parts into new wholes.¹³ For such processes, the term ‘dialectic’ would be fitting – at least for Lakatos. Rather than *explain* or *understand*, endorsing explanatory coherence glosses over them. In doing so, we are prone to cite ‘genius,’ a term which should be “taboo” for the naturalist (see below).

5. Beyond Deep

Forseeably, empirical theories will continue to save corroborated phenomena (within acceptable error-bounds), thus “retain” some of their intended applications, as long as such theorizing is relevant for societies relying on technology.¹⁴ Via engineering, empirical theories also give rise to profitable use; politically, science is expected to be useful in the long run. Why, then, believe one has gotten closer to truth, rather than having obtained empirically more adequate theories?

Presently, physicists endorse theories of quantum-gravity. After all, for the four major “force-field” frameworks (gravitation, electro-

¹³ For an approach to theory development that models integration within a topological account of concepts, see Gärdenfors and Zenker (2010) and Zenker (2010).

¹⁴ Short-term funding under a political impetus for interdisciplinary research sustains and, to this extent, explains deepened and broadened theories. It does not explain their genesis.

magnetism, strong and weak nuclear interaction), it is true: Upon superposition, the frames do *not* cohere.¹⁵ Should these frames be brought to cohere, 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 – which was Thagard’s primary concern –, a new theory *could* still “have” truth or be comparatively closer to it. So, we better stay open to non-deepening truth candidates. Now, would we recognize that quality in a candidate? If we did, then certainly not *because of* explanatory coherence. Would we call the candidate’s genesis creative or mystic? Once deepened, probably ‘creative.’

6. Conclusion

When weakening the induction to a degree of reasonable caution, we leave the *genesis* (discovery) of a new and – on some measure – radically different empirical theory, T^* , unaccounted for. So, explanatory coherence cannot serve in guiding one’s appreciation of T^* *vis à vis* then-current theories. After all, appreciating T^* as a comparatively more truth-approximating candidate is tied to T^* having been broadened and deepened. This *post hoc* criterion privileges the *status quo*. Unless one can already locate the *status quo* on a truth-

¹⁵ In other words, superpositioning these fields entails massive coherence loss (see Thagard 2000, ch.2).

distance measure, “improvements” will likewise lack such information. Therefore, the mood of induction should at most remain *neutral* (rather than optimistic).

Whether the integrating state of an empirical theory is “more true” than the integrated state, cannot be independent of what it means for a developmental process to not be false. This requires a naturalized sense of historical progress and regress, as in Lakatos’ degenerating research programs. That sense is not readily available. In its absence, we face an absurdity: As optimistic inducers who exercise caution, we contribute to mystifying an infrequent, but expectable process which is *native* to the development of empirical theories. Of all epistemological positions, this process should be *fully reconstructable* for a naturalist.

Acknowledgments

For helpful comments and criticism, I thank Bengt Hansson, Johannes Persson, Erik J. Olsson, the audiences of the Higher Seminar in Theoretical Philosophy and the 2009 Swedish Philosophy Days at the University of Lund, as well as two anonymous reviewers.

References

- 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.
- Batterman, R.W. (2003), *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. New York: Oxford University Press.
- Barker, P. (2001), “Incommensurability and Conceptual Change During the Copernican Revolution,” in: Sankey, H. & P. Hoyningen-Huene (eds.). *Incommensurability and Related Matters*. Boston: Kluwer, 241–271.
- Barsalou, L. (1992), “Frames, Concepts, and Conceptual Fields,” in: Lehrer, A. & 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 20th Century Field Theories*. Cambridge: Cambridge University Press.
- 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. (2002), "Kant, Kuhn and the Rationality of Science," *Philosophy of Science* **69**, 171–190.
- Friedman, M. (2003), "Kuhn and Logical Empiricism," in: Nickles, T. (ed.), *Thomas Kuhn*. Cambridge: Cambridge University Press, 19–44.
- Gähde, U. (2002) "Holism, Underdetermination, and the Dynamics of Empirical Theories," *Synthese* **130**, 69–90.
- Gärdenfors, P. (2000), *Conceptual Spaces. The Geometry of Thought*. Cambridge, MA: MIT Press.

- Gärdenfors, P. & F. Zenker (2010), "Using conceptual spaces to model the dynamics of empirical theories," in: Olsson, E. & S. Enqvist (eds.), *Science in Flux: Philosophy of Science Meets Belief Revision Theory*. Berlin: Springer (forthcoming).
- Jammer, M. (1961), *Concepts of Mass in Classical and Modern Physics*. Cambridge, MA: Harvard University Press.
- Kuhn, T. (1962, ²1970), *The Structure of Scientific Revolution*. Chicago: Chicago University Press.
- Kuipers, T. (2000), *From Instrumentalism to Constructive Realism*. Dordrecht: Kluwer.
- Lakatos, I. (1978), *The Methodology of Scientific Research Programs*. Cambridge: Cambridge University Press.
- 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*. 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: Oxford University Press.
- 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.

Zenker, F. (2009), *Ceteris Paribus in Conservative Belief Revision*. Frankfurt: Peter Lang.

Zenker, F. (2010), "From Features via Frames to Spaces. Modeling Scientific Conceptual Change without Incommensurability or Apriority," in: Gamerschlag, T., D. Gerland, R. Osswald & W. Petersen (eds), *Proceedings of the 2nd International Conference on Concept Types and Frames in Language Cognition and Science* (CTF09). Duesseldorf (forthcoming).

Word count: 3928 (including footnotes and references), 3211 (without references)

Originally submitted version: 6300 (incl. footnotes and references)