CROSSWORDS AND COHERENCE

DAVID ATKINSON AND JEANNE PEIJNENBURG

A COMMON OBJECTION TO COHERENTISM is that it cannot account for truth: if a true theory and a false theory are equally coherent (and equally simple, comprehensive, and so forth), coherentism gives us no reason to prefer the former over the latter. By stretching Susan Haack's crossword metaphor to its limits, we show that there are circumstances under which this objection is untenable. Although these circumstances may seem remote, they are in accordance with our most fundamental and ambitious physical theories. Coherence might be truth conducive after all.

I

Susan Haack famously launched foundherentism as a fitting portrayal of the way in which we build up knowledge of the world.[1] A foundherentist explicitly has an ecumenical intent, encompassing requirements of both coherentists and foundationalists. Like the coherentist, a foundherentist stresses that our beliefs about the world must hang together in one way or another; and like the foundationalist, she emphasizes that they must somehow be grounded in the world around us. Foundherentism is thus an amalgam which draws on important insights of two factions that are often pictured as being opposed to one another.

In a happy clarification of the nature of foundherentism, Haack introduced her metaphor of the crossword puzzle. The solution of a crossword combines features that are reminiscent both of foundationalism and coherentism. It mirrors the coherentist creed by demanding that words must intersect other words on the same letter, while it reflects the foundationalist stance by the existence of 'clues' that provide grounds, as it were, for each entry in the puzzle.

---

[1] Susan Haack, *Evidence and Inquiry: Towards Reconstruction in Epistemology* (Oxford: Blackwell, 1993).

The Review of Metaphysics 251 (March 2010): XXX–XXX.

While it is true that many scientific theories can be successfully modeled on the basis of this crossword metaphor, our most advanced and ambitious theories of the world resist such modeling. A Theory of Everything, in particular, will typically fail to fit a crossword metaphor of the sort that Haack has in mind. Such a TOE, as it is familiarly dubbed, aims to account for all the elementary particles in nature, including those that mediate nature's basic forces, like the photon in electromagnetism. As such it is called on to explain why there are quarks and leptons and gauge bosons, but not how, for example, a storm in the north Atlantic can be triggered by a butterfly flapping its wing in the Amazon. Given the expectation of many scientists that there can only be one TOE, the best way to apply the crossword metaphor to a TOE is to picture the latter as a crossword that has a unique solution which can be found without using any external clues. Such crosswords transcend the Haackian metaphor, and in this paper we will explore their potentialities. More particularly, we will stretch the crossword metaphor in two directions: first to the point where only one solution is possible, and then to the point where the foundationalist requirement disappears. After we have shown that crosswords exist which are unique and clue-free, we speculate on their appositeness as metaphors for a TOE.

We start, in Section II, by referring to a standard objection to coherentism, namely that it does not enable us to choose between different coherent systems of propositions (or beliefs): since we can always come up with a rivaling system that is equally coherent, the existence of a unique coherent system is excluded. On the basis of the crossword metaphor we show that there are circumstances under which this objection misses its mark: sometimes a particular system of propositions is indeed unique. We argue in Section III that such a system might not only be unique but also clue-free, in the sense that it is not grounded in external clues or data. In Sections IV and V we speculate about the origin of the structure of the world-crossword (where 'world-crossword' refers both to a TOE and to a metaphor for a TOE). In Section VI we tentatively conclude that coherentism might be truth conducive after all.

For the record, our argument should not be read as an unmitigated defense of coherentism, be it a coherentist theory of truth or a coherentist theory of justification (as Haack said, the distinction seldom constitutes a problem, and for the present enterprise it is wholly unimportant[2]). Our aim is more restricted than that. We first query a standard objection to coherentism, showing that circumstances might obtain under which this objection breaks down: there can exist a system of propositions that is coherent, unique, and not grounded, and this system might serve as an empirical theory of the world. We then explore the possibilities under which a coherentist account may be tenable.

II

A classic objection to any coherentist theory of truth is that it gives no good reason for preferring one coherent system of propositions over another. The roots of this objection are long, stretching back at least to Russell's critique of monistic truth theories.[3] Imagine that we have two systems of propositions, S and S*, each equally coherent (and equally comprehensive, and simple, and so on). Let S contain the (complex) proposition that Bishop Stubbs wore episcopal gaiters and was peacefully lying in bed when he gave up the ghost (*p*), while S* contains the proposition that said bishop wore a pistol belt before he was hanged for murder (*q*). Which of the two systems is to be preferred? Defenders of a correspondence theory of truth have of course no problem in answering this question. They will simply argue that S is to be preferred over S* because *p* is true and *q* is false. Coherentists, however, cannot take this route, since for them there is nothing to truth over and above coherence.

The above objection to coherentism is based on the assumption that there can be more than one coherent system. No doubt this assumption is very often fulfilled: in the majority of cases it will be possible to concoct a competing, but equally coherent system of propositions. This alternative system need not have been formulated yet—the point is whether it could, in principle, be formulated. The assumption that it can always be formulated is what may be called the

---

[2] Haack, *Evidence*, 13.

[3] Bertrand Russell, "The Monistic Theory of Truth," in his *Philosophical Essays* (London: Longmans, Green and Company, 1910; reprinted London-New York: Routledge,1996), 131–46.

principle of plurality. It states that a rivaling, but equally coherent system is always possible, and it underlies the classic objection that coherentism fails to give us the means to choose between the two alternatives.

Clearly, if the plurality assumption were false, the classic objection would lose its bite. For then only one coherent system of propositions would be possible and the need for choice would not arise. In the nineteenth century, Hegelian defenders of a monistic theory of truth might have had such a situation in mind when they claimed that a proposition cannot be called ‘true’ unless it applies to the one and only existing system as a whole, namely the Absolute itself. Russell made short shrift of these ideas and no doubt he was right. Nevertheless, the claim that there can be a unique system of propositions is not prima facie absurd. It may be given a sensible interpretation, as can be demonstrated by extending Haack’s crossword metaphor in the following way.

In an ordinary crossword puzzle each clue is typically ambiguous. For example, if the clue is ‘a flower (eight letters)’, then the solution might be ‘bluebell’ or ‘hyacinth’, ‘daffodil’, ‘geranium’, ‘aubretia’, and so on. The ambiguity is then diminished by the requirement that the eight-letter word in question has to cohere with other words. Still, if a crossword puzzle is very simple, it might well have several, mutually incompatible solutions. Consider for example the crossword in Figure 1, and let the clue for the horizontal row be ‘a bird’, and for the vertical column ‘a boy’s name’.

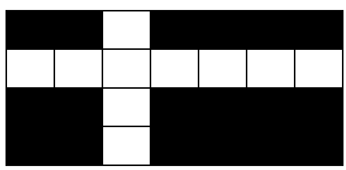

Figure 1

Then we may fill in ‘swallow’, ‘sparrow’ or ‘seagull’ for the first clue, and ‘Carl’, ‘Dave’ or ‘Jack’ for the second one; this already yields nine different solutions for the entire puzzle— and there are clearly many more.

Obviously, crossword puzzles are usually more interesting than the all too simple one of Fig. 1, which hardly deserves the name. In general, two parameters determine how complex a crossword is. The first is the *extension* or *size*, by which we mean a pair of integers, say (c, r), indicating that the crossword has c columns and r rows. The second is the *structure* or *pattern*, by which we mean the number and placing of the blocked-out squares. So the crossword of Fig. 1 has the same size as that of Fig. 2a but a different structure, while the crossword of Fig. 2b has a different size than either and therefore also a different structure.

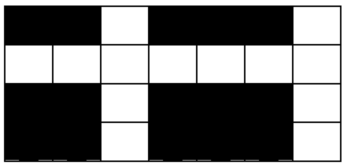

Figure 2a

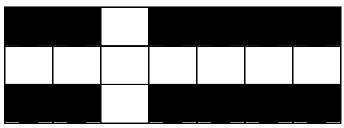

Figure 2b

As the complexity of the crossword increases, the ambiguity in general decreases: it becomes more and more difficult to come up with different solutions. If the size of the crossword is finite, then the number of words that can be filled in on the basis of the given clues will also be finite. Jointly these words form a proper subset of the *vocabulary* or *dictionary*, namely the words that one is allowed to use. This vocabulary is in turn a proper subset of all the *possible strings* that can be made with the letters of the alphabet that one is using, for example the twenty-six letters of the English alphabet. Naturally, the subset will have finitely many elements. After all, even the Oxford English Dictionary, with all its historical citations of archaic word usage thrown in for good measure, contains only a finite number of words. And although some of these words are inordinately sesquipedalian, like 'floccinaucinihilipilification' for example, they are all of a finite length. The number of coherent ways of filling in a finite crossword, with a finite alphabet, irrespective of lexical constraint, is finite. In the end, if the crossword puzzle is sufficiently complicated, there might be only one solution. In this sense the existence of a unique system of propositions or words is possible; and stretching Haack's crossword analogy helps us to see that this possibility is not prima facie excluded.

III

The fact that a finite crossword puzzle of sufficient complexity might have a unique solution calls the classic objection to coherentism into question. It nullifies the principle of plurality, and thereby one of the pillars on which the objection rests. If the plurality principle is false, so that a unique solution is possible in a particular case, then the coherentist no longer needs to choose a solution, since only one is available.

However, the mere fact that a unique solution is possible does not help the coherentist enough. For the difficulty remains that a unique solution is normally achieved by using clues, and clues are so to speak a foundational affair. There is only one way in which the existence of a unique solution could really help the coherentist, and that would be if a unique solution could exist without any clues. Only then could it be claimed that a system of propositions need only be coherent, and that foundationalist grounding is unnecessary.

It goes without saying that clues are usually very important. But it is also clear that there are cases in which not every clue in a puzzle is necessary. It can well happen that, when most of the words have been inserted, a few of the remaining words, which are by that time partially filled in, admit only one completion according to the most comprehensive dictionary in existence. In such a case one could drop the clue in question, for it is redundant. The crucial question for us is: could it happen that *all* the clues are redundant? Is it conceivable that, for a crossword with a particular structure, and for a given dictionary or vocabulary list, *no clues at all* would be needed for the successful, unique completion of the puzzle?

To illustrate what can happen, consider a crossword of size (3, 2) and suppose that the entire vocabulary consists in only the words:

me, he, it, she, him, her, his, its.

Given this vocabulary, it depends on the structure whether there is just one way to fill in the clueless crossword (Fig. 3):

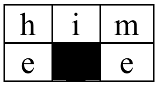

Figure 3

or whether there are more ways than one (Figs. 4a, 4b, 4c), or none at all (Fig. 4d).

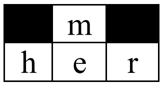

Figure 4a

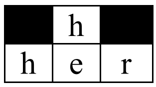

Figure 4b

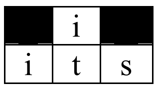

Figure 4c

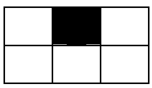

Figure 4d

Of course, if the vocabulary were increased, uniqueness might no longer hold for the structure of Fig. 3; and the structure of Fig. 4d might perhaps be filled in, with use of some of the new words. However, it could happen that, even with the whole contents of the Oxford English Dictionary at our disposal, we might be able to fill in a large crossword of a sufficiently complex structure in one, and only one way. And if there is only one way to complete the crossword with a given vocabulary, then the clues have lost their function—they are completely superfluous.

The above example of a crossword puzzle with a unique, clue-free solution involves a structure and a vocabulary that are both finite. This is no coincidence: it seems clear enough that examples are easier to find when one disregards crosswords with infinite vocabularies and infinite structures. Infinite crosswords are rather exotic, and are likely to trigger intuitions that might well lead us astray. Moreover, there is no need to consider them here. For, and this is an essential point, the restriction to finitude is eminently reasonable if we focus on the system for which a crossword is supposed to function as a metaphor, namely a representation of the universe itself.

Recent findings in astronomy strongly support the idea that our universe is finite, both in past time and current spatial extent. Even more important for our present enquiry is the discovery that this finite universe can only contain a finite amount of information, where information can be seen as a string of 'bits', that is, units that can take on one of two values, traditionally written 0 and 1. In 1972 Jacob Bekenstein showed that black holes have a well-defined entropy (or informational content). This idea formed the basis for his demonstration, nine years later, that there is a maximum to the informational content of any finite volume of space that contains a finite amount of total energy (including rest-energy or mass). This maximum is known as the Bekenstein bound, being the upper bound on the entropy, and thus on the number of bits of information that a given volume with finite total energy can contain.[4] According to the Bekenstein bound, the world can be described by a finite string of zeros and ones. The smallest string from which this finite string can be deduced is the world's informational content; and of course this content will be finite too.

The Bekenstein bound is undoubtedly a force to be reckoned with. It is deeply rooted in modern physics, involving not only the Planck constant but also the speed of light, and it calls upon both quantum theory and the theory of relativity. In fact, as things stand at present, no Theory of Everything would be viable without taking into account the Bekenstein limitation of informational content. If we follow modern physicists and take the Bekenstein bound seriously, then we

[4] Jacob Bekenstein, "A Universal Upper Bound on the Entropy to Energy Ratio for Bounded Systems," *Physical Review* D 23, (1981): 287–98.

must accept that any future TOE will have to envisage our world as an incredibly complex, but informationally finite, many-dimensional, coherent system of bits. And since it needs only to explain a finite number of fundamental facts, a TOE itself need contain only a finite amount of information; in other words, it will also be encodable as a finite string of zeros and ones.

We can picture the TOE as a world-crossword, that is, a sort of many-dimensional crossword wrapped on a hypersphere, in which the letters are bits, conveniently combined into larger 'bytes'.

Table 1: Crossword metaphor applied to TOE

	Crossword Puzzle	**Theory of Everything**
1	twenty-six letters	two Bekenstein bits
2	stopping block	stop byte
3	strings of letters	strings of bytes
4	words	elementary particles

Table 1 compares the elements of Susan Haack's crossword metaphor with the components of the TOE involving the Bekenstein bound. The basic elements in the TOE are bits, here pictured in the first row. The analogous elements in the crossword are the letters of the English alphabet. Each of these twenty-six letters could of course be reduced to a byte of five bits: there are 32 distinct bytes of this kind, enough to distinguish all the twenty-six letters, with a few bytes left over. One of these residual bytes could stand for what we call a stopping block (see the second row). In the context of a crossword, a stopping block stands for a square that is blocked out. In the TOE column the stop byte is likewise some particular combination of bits that signals the end of the encoding of one item and the beginning of the next. In row 3 the 'strings of letters' could also be written as 'strings of bytes' in the more fundamental language in which letters are themselves represented as bytes. The fourth row lists 'words' for the clue-free crossword, and these constitute a proper (and indeed relatively small) subset of all the possible strings of letters, namely those strings that are to be found in the vocabulary or the dictionary appropriate to the language. The analog of this in the TOE column consists in the subsets of those strings of Bekenstein bytes that constitute the encoding of the basic entities in the world, namely the elementary particles. At present we think that these elementary particles are quarks, leptons, and gauge bosons, but a future theory may perhaps analyze these into yet more basic entities.

Whatever the bytes or 'words' may be that make up the fabric of the universe, they are to be thought of as members of a relatively small subset of the set of all conceivable strings of bytes. Together these 'words' form the vocabulary of the world.

IV

We have argued that there are circumstances under which it is possible to have a system of propositions, or a solution to a crossword, which is unique and clue-free. The example we gave involved a crossword that was finite, but we have shown that this assumption of finitude is not at all unreasonable.

Our argument might come as good news for a coherentist, who can now defend herself against the classic objection to her position, explained in Section II. But what would the foundationalist say? He might be prepared to accept the hypothesis of finitude, granting that the world-crossword is finite in size, and taking the lesson from science to heart that our universe is an informationally finite coherent system of bits. In other words, the foundationalist may well accept that, in principle, a structure might be possible that allows for a unique completion, so that the 'clues' have indeed become redundant. However, he need not see this acceptance as a reason for relinquishing his foundationalistic convictions. For he can still maintain that the entire system must be connected to something outside that system. After all, there are still three things that the coherentist must apparently rely on without being able to account for them, namely the *size* or *extension* of the system (or crossword), its *structure* or *pattern*, and last but not least *the vocabulary* or *dictionary*. All of these seem to be 'given' by some agency outside the crossword itself. How can a coherentist come to terms with these objections? We shall deal with this question now, discussing first the provenance of the pattern, as well as that of the size. In the next section we deal with the origin of the vocabulary.

Consider again Fig. 3. Like any other crossword, its size and structure seem to function as a straitjacket, as it were, that has been constructed before one comes to consider the words that fit into it, namely ‘him’, ‘he’, ‘me’. The fact is, however, that neither the size nor the pattern need be regarded as constraints that have been imposed ‘from the outside’. For they can be ‘internalized’, in the sense that a uniform string of symbols can be defined that contains them. The size, the structure, as well as the three words ‘him’, ‘he’, ‘me’ can all be placed at the same level, since all can be encoded in a one-dimensional string of bytes (ultimately bits) of information. The first step in this encoding is to augment the alphabet of twenty-six letters by two extra symbols: Ξ, which stands for a blocked-out square, and Λ, that is an end-of-row marker.

All twenty-eight symbols can be accommodated within the byte containing five bits to which we alluded above. With this extended alphabet the second step can be made. For now we are able fully to specify the crossword of Fig. 3 by the one-dimensional code:

Λ him Λ e Ξ e Λ

This is to be read: in the first row (between the first pair of Λ signs), the letters ‘h’ ‘i’ ‘m’, and in the second row (between the last pair of Λ signs), the letter ‘e’, followed by a filled-in block, followed by another letter ‘e’. In this manner we can read off the number of columns (namely 3) and the number of rows (2), as well as the content of every space in the pattern.

In much the same way that we can encode the size and structure of a crossword as a one-dimensional string of bits, so we can encode the size and structure of a TOE as a finite string of Bekenstein bits of information. Both are subject to requirements of coherence, since in neither case are we dealing with arbitrary strings. But there is an important difference between the two. In the case of the crossword, the string cannot be explained by the crossword itself. Of course, we have the rules of the game, as well as the dictionary, and with the help of them we can explain why the string in question is as it is. But both the rules and the dictionary come from outside; they are not part of the crossword puzzle itself. All this is different in the case of a TOE. For a TOE is self-contained: it must prescribe not only its own size and structure, but also the 'words' in which it is written, namely the elementary particles—be they quarks and leptons or some as yet unsuspected fine-grained level of reality.

V

The vocabulary of 'words', in this case the allowed strings in the Bekenstein encoding of the universe, must itself be part and parcel of the theory, in the present case the TOE. Admittedly, we do not yet have such a TOE, but we know some of the things that it has to do. For example, it must surely combine quantum field theory and general relativity, that is, Einstein's theory of gravity. At a more general level, we require the TOE to satisfy two desiderata: it must be self-contained, and it must be unique. String theory, in its halcyon days, claimed to be such a TOE. Indeed it does combine general relativity and quantum field theory into one system. Moreover, it arguably satisfies the first desideratum: except perhaps for the so-called string tension, nothing is put in by hand. Lately string theory has fallen on hard times, however: it altogether fails on the second count, for it no longer arrogates to itself the possession of a unique solution. And justly so, for the lack of uniqueness of modern string theory is in fact unprecedented in the history of science. There are so many ways of 'rolling up' the ten dimensions of string theory into the four dimensions of space-time in which theoreticians have lived since the Einsteinian revolution, that the number of effective theories encapsulated in string theory is more than astronomical. The number has been estimated, not entirely jocularly, at 10^{108}! What is more, none of these string theories are empirically testable. For the new effects that they predict are at distances comparable to the Planck length ($\sim 10^{-35}$ m), and thus very likely forever beyond the reach of experimenters.

Despite the heroic failure of string theory, the possibility for a future successful TOE remains open; there are tantalizing glimpses of such a possibility in Gerard 't Hooft's holographic principle[5], and

[5] Gerard 't Hooft, "Dimensional Reduction in Quantum Gravity," in *Salamfestschrift: A Collection of Talks from the Conference on Highlights of*

perhaps in speculations concerning cellular automata. But no matter how a future TOE will look, if it is going to be successful it will have to be unique. Moreover, as we said, a TOE must itself prescribe not only size and structure, but also the vocabulary. For example, a crucial difference between the Standard Theory of elementary particles and the TOE is that the Standard Theory still contains unexplained quantities, such as the strengths of the interactions, the number and symmetry of the families of the particles, and so forth. A TOE, on the other hand, would have to account for itself. By definition it may not contain any *ad hoc* elements or constants that are either underivable or derivable from another theory. In fact, it may rely on nothing but its own coherence, and it is by this bootstrap of coherence that a TOE supports itself.

At this juncture, one might ask whether such a TOE can still be called an empirical theory. If it is only coherence that matters, and if the TOE is completely self-contained, how can this theory ever be said to describe the furniture of the world, let alone describe it at its most fundamental level? The answer to this question lies in a switch of direction. In contrast to other empirical theories, our self-contained TOE has not been built up from evidence. Unlike for example the Standard Theory, no data or 'clues' enter the TOE from outside. However, this does not mean that a TOE, once it has been formulated, cannot be found to fit the facts. It certainly can, in the sense that its theoretical constructs and parameters may describe the world as it actually is. Instead of listing the known leptons and quarks and putting them into arbitrarily chosen symmetrical families, as the Standard Theory does, a TOE would contain theoretical constructs corresponding to these particles, probably in terms of more fundamental entities that uniquely cohere with one another. These entities, and the less fundamental particles and forces that derive from them, would form part of the TOE, and they would agree with our observations and experience. The elementary particles are not, as it were, the input, but the output of the theory.

Particle and Condensed Physics (8–12 March 1993 in Trieste), ed. A. Ali, J. Ellis, and S. Randjbar-Daemi (Singapore: World Scientific, 1994), 1–13. A revised version, dating from March 20, 2009, can be found at http://arXiv:gr-qc/9310026v2 .

VI

Received opinion has it that coherence, by itself, is not truth conducive. Coherence might be necessary, but not sufficient for a system of propositions to be called ‘true’—essentially the Bishop Stubbs objection. Against this received opinion, Tomoji Shogenji argued ten years ago that coherence can in fact be truth conducive if it is combined with at least one proposition that is known to be true.[6] Our investigations have suggested that perhaps we might not even need so much: circumstances can exist under which coherence alone would be sufficient for truth. It could be that there is but one way to fill in the world-crossword with ‘words’ that the ultimate theory of the world-crossword itself prescribes. We do not know whether this actually is the case, but the mere possibility is enough to invalidate any premature dismissal of coherentism.

The prospect of the world-crossword being filled in by some future TOE, uniquely determined by the criterion of coherence, may well seem remote. Indeed it is, but it is evinced by our most ambitious scientific theories; it is verily the ultimate goal of unification in physics.[7]

University of Groningen, The Netherlands

---

[6] Tomoji Shogenji, “Is Coherence Truth Conducive?,” *Analysis* 59, no. 4 (1999): 338–45.

[7] We thank Igor Douven for a discussion in March 2008 in Leuven that triggered this paper. We further acknowledge the hospitality of the Department of Philosophy at the National University of Singapore, where in March 2009 we did the writing up. Searching criticism by Tomoji Shogenji and the members of the Promotion Club Cognitive Patterns (PCCP) in Groningen led to a considerable improvement.