Evidential reasoning in historical sciences: applying Toulmin schemes to the case of Archezoa

Thomas Bonnin

Received: 26 January 2018 / Accepted: 1 March 2019 / Published online: 3 April 2019
© The Author(s) 2019

Abstract
This article is a study of the role and use of evidence in the evaluation of claims in the historical sciences. In order to do this, I develop a “snapshot” approach to Toulmin schemas. This framework is applied to the case of Archezoa, an initially supported then eventually rejected hypothesis in evolutionary biology. From this case study, I criticize Cleland’s “smoking gun” account of the methodology of the historical sciences. I argue that Toulmin schemas are conceptually precise tools that allow for the building of enriched reconstructions of evidential reasoning. From the application of this framework, I discuss three ways in which the construction and use of facts in the historical sciences are theory-laden. Despite its inherent limits, TS are heuristically useful tools to identify epistemic moves that could be further investigated. It also sheds light on the positive roles of speculation in the historical sciences. Finally, I argue that it provides a context-specific and individuated understanding of hypothesis evaluation in the historical sciences. Overall, I think the application of Toulmin schemas to cases of evidential reasoning in the historical sciences is a promising descriptive and heuristic tool for philosophers of science.

Keywords Philosophy of historical sciences · Evidential reasoning · Evolutionary biology · Toulmin schemas · Theory-ladenness of facts

Introduction
The historical sciences, understood as the study of the “events and entities which populate the deep past” (Currie and Turner 2016, p. 43), include a broad range of disciplines from palaeontology to cosmology. All these disciplines are united by a similar epistemic challenge. They all have to draw inferences about past phenomena which, by definition, are physically forever out of reach. Historical scientists,
therefore, rely on the interpretation of the remains of these past phenomena to validate and refute their hypotheses. This opens up (at least) two questions: (a) how are these lines of evidence generated? (b) How are they mobilized in the assessment of a given hypothesis?

This article primarily addresses the second question by applying “Toulmin schemas” (henceforth TS), a framework proposed by Toulmin. This framework, I argue, is a useful tool to drive textual reconstructions and build visual representations of the dynamics of evidential reasoning in the context of the historical sciences. The “Toulmin schemas” section provides a detailed description of the various components of this framework and the slight amendments I make on Toulmin’s original formulation. One of the proposed amendments is terminological. The notion of “data” is replaced by “facts”. The second, more substantial, change concerns how the framework is used. Instead of building single diagrams providing a static summary of the evidential situation around a claim, I propose a dynamical approach. This is done by building a series of diagrams at various key points of the evidential assessment of a claim, providing snapshots of the situation and enabling us to represent the shifts in the evidential picture and in the security of an inference.

The “Case Study: the rise and fall of Archezoa” section applies the “snapshot approach” to Toulmin schemas to the case of Archezoa. The Archezoa hypothesis, formulated by evolutionary biologist Cavalier-Smith in the 1980s, postulates four eukaryotic taxa to be contemporary descendants of some of the earliest eukaryotic lineages. The construction of TS at key evidential junctures provides a historical reconstruction of the initial support and eventual rejection of this hypothesis. It does so by identifying the alterations in the conceptual and factual landscape that led to such changes.

Lastly, the “Toulmin schemas, archezoa and the historical sciences” section provides a philosophical analysis of this reconstruction. In this section, I both expand on the virtues and limits of TS as a tool for the analysis and visualization of evidential reasoning and attempt to draw lessons from this case for our understanding of the historical sciences. While Archezoa presents an apparently paradigmatic case of historical science, I argue that the present account stands in sharp contrast with Cleland’s “smoking gun” account of its methodology. My analysis has five related “take-home messages”:

(a) Toulmin schemas are conceptually precise tools which allow us to enrich our description of evidential reasoning.
(b) This framework is particularly useful to display and suggest three ways in which the construction and use of facts in the historical sciences are theory-laden.
(c) Albeit limited as a representational tool, the snapshot approach to TS is heuristically useful to open investigations into what lies behind each identified epistemic move.
(d) This reconstruction of the case of Archezoa emphasizes the virtues of speculation in the historical sciences.
Toulmin schemas provide context-specific and individuated (non-comparative) understandings of hypothesis evaluation.

Before reaching these conclusions, this article starts off with an abstract description of Toulmin schemas.

**Toulmin schemas**

Toulmin schemas are put to use to reconstruct the dynamics of evidential reasoning via the interaction of a set of six components: facts, warrants, claims, backing, rebuttals and qualifiers. Despite choosing a slightly different terminology, this usage essentially follows Toulmin’s initial formulation and Chapman and Wylie’s recent application to the context of archaeology (2016, pp. 33–40). Toulmin’s intention was to provide an enriched picture of the structure of arguments, with more components than the traditional syllogistic structure (namely “minor” and “major premises”, and “conclusions”). TS, which are inspired by jurisprudential reasoning, are designed to better capture concrete cases of constructions of arguments. The visual layout of the components and their relations is provided in Fig. 1.

The first three components form the tripartite basis of an argument. A claim is made about a target phenomenon from a set of facts. The inferential leap from facts to claims is supported by a set of warrants.

**Facts**

“Facts” are the equivalent of “data” in Toulmin’s original formulation, in which the latter are defined as “the facts we appeal to as foundation for the claim” (Toulmin 2003, pp. 90–91). My proposed terminological shift is motivated by an endorsement of Leonelli’s relational account of data. In this view, data are defined as “any product of research activities […] that is collected, stored, and disseminated in order to be used as evidence for knowledge claims” (Leonelli 2016, p. 77). In other words,
“data” are (1) potential evidence built to (2) travel through several epistemic contexts. Here, “facts” are considered as (1) actual evidence which (2) grounds claims in a specific context. They are the comparatively more robust grounds upon which a more tentative claim can be formulated.¹

**Warrants**

Facts “do not speak on their own”. Warrants are what mediate the interpretation from facts to claims. Starting from a set of facts, warrants are “general, hypothetical statements, which act as bridges and authorize the sort of step to which our particular argument commits us” (Toulmin 2003, p. 91). The differences between facts and warrants are gradual rather than clear-cut. Following Toulmin, the first difference concerns the warrants’ higher degree of generality. Secondly, facts are always explicitly stated in the defence of a claim, whereas most of the warrants are usually kept implicit (Toulmin 2003, pp. 92–93).

**Qualifiers, rebuttals and backings**

Toulmin recognized that warrants only provide equivocal support to given claims on the basis of a set of facts:

[some] warrants entitle us in suitable cases to qualify our conclusion with the adverb ‘necessarily’; others authorize us to make the step from data to conclusion either tentatively, or else subject to conditions, exceptions, or qualifications […]. It may not be sufficient, therefore, simply to specify our data, warrant and claim: we may need to add some explicit reference to the degree of force which our data confer on our claim in virtue of our warrant (Toulmin 2003, p. 93).

This, for Toulmin, justifies the addition of *qualifiers* in his framework. Qualifiers, in this view, are used to specify the security of the inference to the claim. They are visually signified by the dotting of the arrows linking facts to claim (see Fig. 1). The spacing is meant to be an informal representation, not a quantitatively accurate estimation of the security of the link. By contrast, pale arrows designate “broken” links. This is when, for instance, warrants have lost their applicability in this context (see Fig. 2b–e).

In a complementary role, *rebuttals* are components that make explicit the potential fragility identified by qualifiers. They are used to indicate “circumstances in which the general authority of the warrant would have to be set aside” (Toulmin 2003, p. 94). In other words, rebuttals indicate specific circumstances in which the claim made would turn out to be invalid.

¹ This specific usage of “facts” is by no means an attempt to provide a general definition of this notion (see Mulligan and Correia 2017 for a recent review).
Fig. 2  a–e TS representing the epistemic situation of the Archezoa hypothesis at various times. a Represents the initial support to the hypothesis, b the later rejection of Parabasalia, c the rejection of Archamoebae, d the rejection of Microsporidia and e of Metamonada. New facts and warrants are in bold letters.
Backings complete this conceptual arsenal. These are further facts that can be brought to ensure the applicability of the warrants by specifying that the circumstances in which the warrants are applied are the right ones. Backings are secondary facts used in support of warrants. Backings are distinguished from facts functionally. The former is not necessary to the formulation and the sustained existence of a given claim, contrary to the latter. In Toulmin’s view, backings remain implicit in cases the applicability of warrants is “conceded without challenge” (Toulmin 2003, p. 98).

As already noted by Chapman and Wylie, Toulmin’s proposal emphasizes warrants as the potential source of fragility to the whole argumentative construction (Chapman and Wylie 2016, p. 35). The incapacity of warrants to license foolproof inferences justifies the addition of qualifiers, the identification of rebuttals and the need to provide backings. I think, however, that more emphasis should be given to the potentially substantial repercussions of changes in the factual and theoretical grounds for the solidity of claims. My snapshot approach to the construction of TS is an attempt to account for the dynamic nature of the factual and theoretical grounds on which claims are defended. In their
application of Toulmin’s framework, Chapman and Wylie provide *synchronic* reconstructions of evidential reasoning: they summarize evidential arguments in one static diagram (see 2016, p. 35: Fig. 1.2 and p. 70: Fig. 2.3). By contrast, I build TS at different key stages of the evidential assessment of a claim, thereby building a series of snapshots. This way, I aim to extend their approach to capture a more *dynamic* picture that includes shifts in the security of claims, in evidential strength and in the relevance of facts.

In the next section, this snapshot approach to TS is directly applied to provide a historical reconstruction of the case of Archezoa.

**Case study: the rise and fall of Archezoa**

The Archezoa hypothesis, associated with the work of Cavalier-Smith, designates a hypothesis in evolutionary biology with two interrelated components. First, it is a taxonomic hypothesis about the classification of contemporary eukaryotes. Archezoa, in its initial formulation, regroups four phyla: Archamoebae, Metamonada, Microspora and Parabasalia (Cavalier-Smith 1987a, p. 56). All these four phyla are protists (unicellular eukaryotes). The basis for this classification is the shared possession (or rather absence) of several morphological traits, the most important one being “to completely lack any trace of mitochondria” (Cavalier-Smith 1987b, p. 17).

This claim does not simply mean a *current* absence of mitochondria, but an absence of *traces* of mitochondria that could indicate a *past* presence of this organelle.

This taxonomic inference has important evolutionary overtones. Protists, as unicellular eukaryotes, constitute privileged loci of investigations into several evolutionary questions. They are of interest for investigations into the origin of multicellularity—a trait exclusive to eukaryotes—since the first multicellular lineages stemmed from these organisms (for a recent review on the topic, see Sebé-Pedrós et al. 2017). Protists are also studied in relation to the origin of eukaryotic cells. It is assumed from their unicellularity that they contain the descendants of some of the most “primitive” (in the sense of earliest emerging) eukaryotic lineages.

In investigations into the origin of eukaryotes, two traits, the possession of mitochondria and of a nucleus, occupy a substantial part of the scientific community’s attention (for a recent review, see López-García and Moreira 2015). These two cellular structures are indeed considered as defining traits of eukaryotic cells and are observed in nearly all of the representatives of eukaryotes. This is why archezoans, defined as eukaryotes without mitochondria, were considered as “living representatives of the earliest phases of eukaryote evolution” (Cavalier-Smith 1987b, p. 17). In other words, members of Archezoa were hypothesized to be contemporary remains of one of the oldest eukaryotic lineage: one that formed after the emergence of the nucleus, but before the acquisition via endosymbiosis of mitochondria. This way, the Archezoa hypothesis fits with a “mito-late” (Ettema 2016) or “autogenous” (O’Malley 2010) explanation of the origin of eukaryotes, one in which the appearance of several eukaryotic cellular structures precedes the acquisition of mitochondria.
The aim of this section is to provide both a textual and a visual reconstruction (by using TS) of the evidential reasoning behind the initial support and eventual rejection of the Archezoa hypothesis. This reconstruction builds on existing historical accounts, most helpfully on O’Malley’s study (2010, p. 216). My account essentially agrees with her analysis and uses similar resources. Her reconstruction provides an accurate summary of the various lines of evidence that together led to the rejection of the claim. My proposed reconstruction, driven by the snapshot approach to TS, extends hers by providing a more temporally anchored reconstruction of the initial support and taxon-specific rejection of Archezoa. It aims at further substantiating the evidential reasoning behind “the demolition of the Archezoa hypothesis” (O’Malley 2010, p. 216).

**Initial support**

The evidence in support of the classification came from the observation in Archezoa members of a series of apparently primitive traits that singles them out from the rest of eukaryotes. This included the possession of “prokaryotic” 70S (instead of “eukaryotic” 80S) ribosomes, the absence of well-developed Golgi dictyosomes, and the absence of mitochondria. Moreover, molecular phylogenies, built from the comparison of ribosomal RNA (rRNA) sequences, positioned Archezoa members close to the origin of eukaryotes. This set of facts, warranted by the assumption that these traits are both markers of primitivity in eukaryotes and that they were exclusively present in these eukaryotic taxa, provided the initial support to the Archezoa classification.

Moreover, the use of these warrants was backed by what was, at the time, considered as the mainstream explanation of the origin of eukaryotes (Cavalier-Smith 1987a, b). Since, in this view, the origin of the nucleus precedes the origin of mitochondria, scientists were expecting to find primitive lineages that have never possessed mitochondria. This relation of support is bidirectional: the inferred existence of Archezoa constituted a powerful fact in support of this “mito-late” explanation.

The security of this inference, however, was not foolproof. The claim for the existence of the taxon was threatened by three related rebuttals. Further research could result in the discovery of traces of mitochondrial presence in purported Archezoa members, such as mitochondria-specific genes, proteins and structures. It would also be possible to interpret the facts in the light of alternative warrants which contradict the initial claim. While the traits shared by Archezoa members are considered as marks of primitivity, it was also possible that they came from secondary simplifications. In such cases, these traits would indicate a rather late origin to these organisms, stemming from lineages that have lost some of their “more complex” traits. The exclusivity of the primitive traits shared by Archezoa members could also be undermined by showing that later eukaryotes also possess such traits. Figure 2a provides a TS summarizing the initial evidential support for the Archezoa hypothesis.

As stated in the introduction, the initial uncertainties around the validity of the hypothesis have grown stronger over time and led to the eventual rejection of the existence of Archezoa. The reasoning behind this rejection is specific to each
proposed member. To discuss and represent the “downfall” of this hypothesis, it is thus better to zoom in at the level of each specific taxon.

**Parabasalia and the status of hydrogenosomes**

Parabasalia’s inclusion in Archezoa was the least initially supported one. rRNA-based phylogenies provided only mixed support to their early positioning within eukaryotes: sometimes supporting it (Sogin 1989), sometimes supporting a secondarily simplified status (Qu et al. 1988; Perasso et al. 1989). Further, and contrary to other members of Archezoa, scientists have shown the presence of well-developed dictyosomes and hydrogenosomes in its members (Cavalier-Smith 1987a, p. 23). Well-developed dictyosomes, as described above, were argued to be absent in Archezoa members. Hydrogenosomes are cellular organelles responsible for energy production in anaerobic conditions and excrete hydrogen gas as a metabolic waste product (hence their name). The evolutionary origin of this structure was uncertain and debated at the time. Is it evolutionarily related to mitochondria or has it emerged independently of them? A positive answer to this question implies that possessors of hydrogenosomes (which includes Parabasalia) could not have emerged before the origin of mitochondria.

While initially unresolved, further research resulted in a firmer understanding of the evolutionary status of hydrogenosomes and established their common descent with mitochondria. The history behind the formation of a consensus about this particular claim is something that deserves a separate study. For now, I would like to briefly mention the opinion of one of the chief detractors of a common origin of mitochondria and hydrogenosomes, who recalled that in 1992 “convincing data started appearing that showed my hypothesis to be way off the mark” (Müller 2007, p. 9).

These altered factual and theoretical grounds (represented in Fig. 2b) resulted in the rejection of Parabasalia from Archezoa. This rejection had downstream consequences on the security of the taxonomic inference for the other purported Archezoa members. If Parabasalia is not part of Archezoa but possesses 70S ribosomes and well-developed dictyosomes, then the presence of these two does not constitute facts on which to infer the belonging of their possessors to Archezoa. Such traits were neither considered as marks of early origin nor as exclusive to Archezoa members anymore.

**Archamoebae as advanced eukaryotes**

The credibility of the presence of Archamoebae within the Archezoa taxon initially suffered from having mainly been negatively defined. Archamoebae was devised to accommodate “amitochondrial amoebae that could not be placed” in the three other initial archezoan phyla (Cavalier-Smith 1991, p. 27). The exclusion of Archamoebae from Archezoa came from phylogenies built with rRNA sequences that contradicted their early positioning in the eukaryotic tree (Morin and Mignot 1995) and from the
isolation of genes of mitochondrial origin in *E. histolytica* (Clark and Roger 1995). In light of these new facts (Fig. 2c), Archamoebae were in fact considered to be “relatively advanced eukaryotes that have almost certainly evolved by the secondary loss of mitochondria” (Cavalier-Smith and Chao 1996, p. 557), corresponding to the possibilities evoked in the rebuttals.

**Microsporidia as degenerate fungi**

The inclusion of Microsporidia in Archezoa was initially supported by rRNA phylogenies (Vossbrinck et al. 1987) and their possession of prokaryote-like ribosomes (Vossbrinck and Woese 1986). The risk, as mentioned in the rebuttals, was that this primitivity was not the mark of genuine early evolution, but rather the outcome of a secondary process of simplification that derived from their adaptation to a parasitic lifestyle (Cavalier-Smith 1993, p. 964). This suspicion was confirmed by the finding of spliceosomal RNA, a specific marker of mitochondrial presence, in Microsporidia (DiMaria et al. 1996). In addition to this, phylogenies built with sequences from the families of chaperone proteins (Germot et al. 1997) and of tubulins (Li et al. 1996; Keeling and Doolittle 1996; Roger 1996) confirmed their later origin in the eukaryote tree. These updated factual grounds pointed to a counter-claim entailed in the rebuttals (Fig. 2d). Instead of being primitive eukaryotes, Microsporidia were now seen as “heavily degenerate fungi” (Cavalier-Smith 1998, p. 227).

**What about metamonada?**

Metamonada’s inclusion in Archezoa was initially well-supported. It was grounded on rRNA phylogenies built with sequences from *G. lamblia* (Sogin et al. 1989). As other members were progressively removed from Archezoa, the support of the group as a whole generally decreased. The claim was now only weakly supported by molecular phylogenies and the absence of traces of mitochondrial presence in Metamonada. Rejection of Metamonada from Archezoa came from the rejection of this latter fact. Phylogenies built with sequences of the cpn60 gene “grouped solidly [members of Metamonada] with the mitochondrial and α-proteobacterial sequences” (Roger 1999, p. 152), pointing to the ancient presence of mitochondria in Metamonada. In parallel to this, there was the discovery of a gene involved in protein synthesis “specifically related to the homologue from *Trichomonas vaginalis* (Hashimoto et al. 1998)” (Roger 1999, p. 152). Since *T. vaginalis* is a member of Parabasalia, this gene was taken as a marker of close evolutionary proximity between Metamonada and Parabasalia, therefore arguing against the primitiveness of the former. The addition of these new facts led to the rejection of the claim that Metamonada are primitively amitochondrial eukaryotes (Fig. 2e).
Epilogue—What remains of archezoa?

Once all of the purported Archezoa members were removed from this taxon, what remained of the once popular hypothesis and of its evolutionary underpinnings? The name “Archezoa” remained as a phylum within protists, grouping Metamonada with Parabasalia. They were still, for a while, claimed to be the earliest (albeit post-mitochondrial) eukaryotes (Cavalier-Smith 1998, p. 206). This claim was also eventually rejected and led Cavalier-Smith to drop the name in 2003 (p. 1745).

Cavalier-Smith defended, at the dawn of Archezoa, the logical independence of the classificatory and the evolutionary sides of this hypothesis (Cavalier-Smith 2002, p. 318). The claims that were rejected, according to him, are the grouping of four distinct taxa together and the idea that these organisms are contemporary remnants of early eukaryotic evolution. It did not, accordingly, fully eliminate his mitochondria-late explanation of the origin of eukaryotes, which he still defends (see Cavalier-Smith 2014 for a recent formulation). With the growth of alternative hypotheses (for instance, the “hydrogen hypothesis” from Martin and Müller 1998), the debates about the origin of eukaryotes and about the place of mitochondria in the picture are still open and active.

In the next section, I use this application of TS to the case of Archezoa to draw insights about the applicability and usefulness of this framework and to critically assess existing positions in the literature about the historical sciences.

Toulmin schemas, Archezoa and the historical sciences

Reconstructing cases of evidential reasoning requires selecting, in the scientific literature, the relevant elements to include. In the previous section, the case of Archezoa was reconstructed by building a series of Toulmin schemas. Cavalier-Smith’s publications on this issue were scanned to find the facts, warrants, backings, rebuttals and the degree of qualification associated with each claim, as well as how these components have changed over time. This section has two interrelated aims: (1) to show why I think TS provide a good conceptual framework to provide historical reconstructions of cases in the historical sciences and (2) to make explicit the sort of image of evidential reasoning in the historical sciences contained in TS as shown by its application to the case of Archezoa.

Conceptual clarity

The first apparent appeal of TS as a framework is, I think, its ability to be terminologically precise without being cumbersome. Facts and backings are, in principle, of a similar nature. They could probably be regrouped under the more general category “lines of evidence”. However, they play distinct roles in TS. Similarly, both warrants and rebuttals have a capacity to bring interpretative light on facts. A purported archezoan trait, for instance, could be interpreted as a mark of early origin or as a sign
of secondary simplification. These two types of principles, however, play a separate role in this context since one underpins the validity of the claim (warrant) while the other points to how it might be undermined (rebuttal).

This conceptual precision also extends to the evaluation of the security of an inference from facts to claims. Qualifiers and rebuttals have the virtue of substantiating the fallibility of claims. The historical sciences are marked by a pervasive underdetermination (Turner 2007; Currie and Turner 2016) that goes as deep as to the level of potential data (Wylie this issue). Rebuttals are therefore a particularly important component in evidential reasoning. They help us point out the limit of the security of a given interpretation and the alternative warrants with which a set of facts could be interpreted. Not making this component explicit runs the risk of having an exaggerated optimism on the validity of a claim. In the same vein, qualifiers are important to visually substantiate the uncertainty surrounding a claim. The use of qualifiers, I think, can also be helpful when it comes to understanding disagreement about claims. In this usage, they could display how scientists place different weights on some components. For instance, it can help to visualize disagreements about the relevance of some facts and backings as well as the different weighing of warrants.

I argue that this increased conceptual precision improves on how evidential reasoning is currently presented in the philosophical and scientific literature, which usually rests on a rather undifferentiated dichotomy between “theory/hypothesis” and “evidence”. Between the usually recognized role of facts and claims, Chapman and Wylie argued that “Toulmin’s central point is that the inferential work of warrants should be recognized as critical to the appraisal of substantial arguments” (Chapman and Wylie 2016, p. 35). This centrality of warrants is clear in the initial support to the Archezoa hypothesis, which relied upon the solidity of the assumptions that the traits identified were indeed marks of primitivity and not possessed by other eukaryotes. TS thus explicitly display a first type of theory-ladenness:

\[ TL_1 \]: Warrants mediate the support to claims brought by facts.

However, the above reconstruction makes it look as if (except for Parabasalia) the sheer discovery of new facts triggered the downfall of each Archezoa lineage. This impression is an artefact resulting from a lack of space. It would have been possible to construct more snapshots, notably after each rejection of a lineage from Archezoa, to display the weakened authority of warrants, which progressively undermined the support for each of the abovementioned claim, which was then decisively rejected when new relevant facts were brought in the picture.

**Two types of implicit theory-ladenness**

In addition to \( TL_1 \), I argue that TS allows the identification of two other, implicit, forms of theory-ladenness of the facts mobilized in the defence of claims. The first concerns the claim-driven evidential relevance of facts. The second concerns the methodological and conceptual resources employed in the construction of such facts.
They are both used to argue against Cleland’s lack of emphasis on the various types of theory-ladenness of the generation and use of facts in the historical sciences.

Across this case study, all of the facts appealed to are directly derived from contemporary remains. They were built from the discovery of structures (such as certain genes and cellular structures) and the comparison of homologous contemporary DNA and protein sequences that enabled scientists to uncover relevant characteristics in each Archezoa member. Evidential relevance for each of these facts was grounded on the Archezoa hypothesis, which underpins the expectation of finding shared characteristics in the downstream descendants of these closely related lineages. The application of TS to the case of Archezoa thus implicitly sheds light on another form of theory-ladenness of facts:

**TL₂**: Claims and warrants establish the evidential relevance of facts.

An illustration of **TL₂** concerns the role of the possession of hydrogenosome as a fact refuting the belonging of Parabasalia in Archezoa. Here, it is in establishing an evolutionary link between the latter and mitochondria that evidential relevance was generated for this fact.

The nature of the facts appealed to in the case of Archezoa seem to bear affinities with Cleland’s account of evidential reasoning in the historical sciences. In her view, these sciences proceed mainly by providing common cause explanations, which explain “observable phenomena in terms of unobservable causes […]” (Cleland 2001, p. 987). In both the support and rejection of Archezoa, evolutionary biologists are indeed inferring a common cause (the ancestors of these organisms originated before/after the origin of mitochondria in eukaryotes) to the body of contemporary traces they observe. Closer looks at the case and at Cleland’s account, however, reveal limits to this affinity.

The facts mobilized in the assessment of the Archezoa claim are not only theory-laden in the **TL₁** and **TL₂** senses, but also because none of these facts constitutes “raw discoveries”. Instead, as hinted above in “Introduction” section, facts can be seen as sedimented claims that were themselves defended in other contexts. They are often the product of complex (and sometimes contested) methodologies. Phylogenies built on the comparison of gene sequences are one illustration of this claim. The production of such facts follows a series of steps of data selection and comparison that are all mediated by increasingly sophisticated computational and theoretical background knowledge (Suárez-Díaz and Anaya-Muñoz 2008; O’Malley 2013, 2016; Bonnin and Lombard forthcoming). Improvements in various aspects of this method explain the volatile character of the phylogeny-derived facts in the case of Archezoa. Therefore, facts, in the case of Archezoa, are theory-laden in a third sense:

**TL₃**: The constitution of facts is unavoidably methodologically mediated.

By this, I mean to bring emphasis on the primary importance of the network of models, instruments and conceptual commitments which are used to constitute and give credibility to facts. This complex epistemic network is often implicit (hence trusted) when facts are used in defense of claims. In the case of Archezoa, shifts in
methodological commitments ($TL_3$) and the crystallization of new claims ($TL_2$) are not explicitly visible in the various snapshots of the evidential situation. Nevertheless, these implicit processes have visible consequences on the relevance and interpretation of facts.

None of these three forms of theory-ladenness is, I think, foregrounded in Cleland’s analysis. Her use of the *smoking gun* metaphor to describe the typical line of evidence in the historical sciences is symptomatic of this relative disregard. A smoking gun is “a trace(s) that unambiguously discriminates one hypothesis from among a set of currently available hypotheses as providing ‘the best explanation’ of the traces thus far observed (Cleland 2002, p. 481)”2. The emphasis on the lack of ambiguity has already been criticized for providing inflated expectations about the effects of lines of evidence (see O’Malley 2016, in the context of molecular phylogenetics). Further to this, I think that this metaphor does not emphasize any of the theory-ladenness of facts discussed above, $TL_3$ in particular. In the literary passage that has given the metaphor its meaning, it does not require any methodological mediation to infer from the smoke stemming from a gun that a shot has been fired from this gun in the near past and to identify the murderer as the person still holding the smoking gun.2. The criticism here is that the trick for historical scientists is not to “find” smoking guns dispersed in nature (contra Cleland 2002, p. 490) but to construct facts that are theory-laden in three ways: by the often sophisticated methodologies behind their constitution and interpretation ($TL_3$), by the claims underpinning their evidential relevance ($TL_2$) and by the inferential work of warrants in supporting or undermining a given claim ($TL_1$).

This triple theory-ladenness points to another issue with Cleland’s account. In her view, the finding of new traces (smoking guns) is seen as the main source of change in the assessment of claims. However, if the continuous discovery of new facts is indeed a key aspect of this process, one should not disregard the parallel primary importance of methodological and theoretical changes. The rejection of Parabasalia from Archezoa, for instance, is a theoretical change that removed the evidential relevance of a set of facts (the possession of 70S ribosomes and well-developed dictyosomes) and at the same time decreased the inferential security of the Archezoa membership of the other taxa as a whole. The same facts were thus given different roles (or no roles) with this change in accepted claims. This crucial importance of theoretical and methodological changes on the assessment of the “same” traces was also beautifully illustrated in cases of secondary retrieval of “legacy data” in the context of archaeology (Wylie 2017).

To summarize, while it is true that the case of Archezoa presents an instance of common cause reasoning based on the interpretation of relevant contemporary traces, the following case differs from Cleland’s methodological picture by placing much more emphasis on the various forms of theory-ladenness behind the use of these traces. In this respect, I argue that this case study fits better with views

---

2 “We rushed into the captain’s cabin […] there he lay with his brains smeared over the chart of the Atlantic […] while the chaplain stood with a smoking pistol in his hand at his elbow” (Conan-Doyle 1893, cited in Safire 2003).
defending the centrality of middle-range theories and investigative scaffolds in the historical sciences.

A middle-range theory (MRT), initially conceptualized by Binford (1977) in the context of archaeology, is the theoretical package that “tells us how an event’s footprint at a time is made and then transformed” (Currie and Sterelny 2017, p. 19; see also Kosso 2001; Jeffares 2008). In other words, MRTs create inferential links between contemporary remains and their common cause in the past. A set of MRTs underpins the possibility, for instance, to build molecular phylogenies and to evaluate the security of the outcomes of such analyses. It corresponds to the TL$_3$ type of theory-ladenness.

Investigative scaffolds, in Currie’s terminology, are “a set of claims [that] must already be on the table for new evidence to be relevant” (Currie 2015, p. 188). Currie mainly focuses on the positive effects of such scaffolding claims. In the case of Archezoa, such positive effects are seen in the case of hydrogenosomes: establishing their evolutionary link with mitochondria allowed the presence of hydrogenosomes in Parabasalia to be evidentially relevant. However, this case study enlarges the range of effects of such scaffolds by highlighting the sometimes negative effects they can have regarding a claim’s epistemic security. Here, the acceptance of a given claim suppressed the evidential relevance of some facts. With the rejection of Parabasalia from Archezoa, 70S ribosomes and well-developed dictyosomes ceased to be evidentially relevant traits possessed by contemporary organisms. Both types of investigative scaffolding map the TL$_1$ (when the scaffolds are warrants) and TL$_2$ (when the scaffolds are claims) understanding of the theory-ladenness of facts in the historical sciences.

**Toulmin schemas as heuristic tools for historians and philosophers**

Visual reconstructions of evidential reasoning with TS only explicitly display TL$_1$ but leave TL$_2$ and TL$_3$ implicit. On a similar note, it also only makes it possible to visualize propositional knowledge, relegating material, tacit and embodied forms of knowledge to the background. I think, however, that uprooting the implicit TL$_3$ that underpins the constitution of evidence necessitates an investigation into the material counterparts of these theoretical and methodological choices. As a general point, if TS as diagrams are limited representational tools, I think that they are heuristically powerful as they can help to identify a variety of investigations related to the generation and use of evidence in a given situation. To give a few examples: how are cellular structures (such as dictyosomes) reliably identified in microorganisms? How to distinguish between primitively and secondarily simplified traits? Between ambiguous and unambiguous phylogenetic trees? None of these questions present things that TS, as a descriptive tool for context-specific evidential reasoning, are capable of addressing. However, these questions, I think, can emerge from looking at the series of TS reconstructing the case of Archezoa.

The “snapshot” approach proposed in this article is particularly interesting in this respect. It helps to identify shifts in the interpretation of given facts or changes in the warrants used at a given time. In the case of Archezoa, it allows identifying the
shifting role played by the possession of hydrogenosomes in Parabasalia, in particular how it became evidentially relevant once its common descent with mitochondria has been established. In addition to the descriptive questions it can open (about how the consensus emerged on this question), I argue that clearly identifying such shifts can be helpful for normative purposes, helping to formulate precise questions about the legitimacy of given epistemic moves (similar to what Chang 2012, pp. 1–92, does with the phlogiston/oxygen controversy). Obviously, the construction of TS is not the only way to generate such investigations, but I think that building a series of relevant snapshots is capable of tracking most of the explicit epistemic moves that were made in a given epistemic context, thereby generating a rather extensive set of questions. This is further facilitated by its diagrammatic nature which, I think, makes it a more efficient communication tool that can substitute for less visually tractable and intuitively generative textual descriptions of similar episodes.

On the whole, I think that TS are therefore interesting descriptive and heuristic tools for investigators, such as philosophers and historians, both interested in reconstructing the use of evidence in a given case and in building a heuristically useful starting point into further investigations. Finally, I believe the use of TS has value for scientists, since it accurately deconstructs their evidential strategies, the shifts in factual and conceptual grounds, and helps to find loci of disagreements.

Speculation and historical progress

Emphasizing various types of theory-ladenness and opening up further investigation are not the only virtues of TS as a conceptual and visual tool. I argue that it can also be used to illustrate how the formulation of a “false” hypothesis helps generate a variety of epistemic goods.

This case study, firstly, provides an illustration of the positive uses of speculation in the historical sciences. This runs against a view that stresses the arbitrariness associated with the invention of events that outrun the available evidence in historical explanations (see Cleland 2009, p. 66). In the formulation of Archezoa, Cavalier-Smith was explicitly aware of making “bold claims”. This is visible in the formulation of Archamoebae as a taxonomic gap-filler, but also in his initial recognition of the precarious status of Parabasalia as members of Archezoa.

The Archezoa hypothesis, I think, provides a nice instance of what Currie and Sterelny describe as productive speculation (Currie and Sterelny 2017; Currie 2018). In this view, “going beyond simple description into the bold and the speculative is more fruitful, generates more opportunities for testing, unlocks new avenues of investigation, links our little pockets of knowledge to each other […]” (Currie 2018, p. 287). This is in line with the defence of investigative scaffolding outlined above, as it emphasizes the necessity to put forward claims (this time, speculative ones) to generate evidential relevance. Would there have been so much attention
given to the evolutionary status of hydrogenosomes if the Archezoa hypothesis was not formulated? Would have we attempted to find reliable markers of mitochondrial presence in members of Archezoa? Could have we understood the secondarily simplified nature of Microsporidia if there weren’t boldly postulated to be primitive in the first place?

While I cannot answer “no” with certitude to all such counterfactual questions, my intuition is that it is the case. Without Archezoa, it probably would have taken another speculative hypothesis to bring the attention of the scientific community to such questions. Therefore, while Archezoa has been shown wrong, I think that it undoubtedly opened up lines of investigations that led to the production of various epistemic goods within but also outside evolutionary biology. Surely, knowledge about the functioning of contemporary organisms feeds off an increased understanding of their origins.

Context-specificity and minimal formality

Lastly, I would like to expand on the advantages of conceiving and representing evidential reasoning in a context-specific and minimally formal fashion.³

By applying to a case from the historical sciences a conceptual framework initially devised for the uses of arguments in general, it might be objected to the use of TS that it subsumes disciplinary specificities under a dangerously unified pattern of evidential reasoning. I think, however, that the use of TS is compatible with endorsing methodological forms of scientific pluralism and retains the distinctness of the case under investigation. Firstly, while TS have, I think, fruitfully been applied to the case of Archezoa and already been used in other contexts, I do not mean to argue for the universal applicability of this framework across scientific practices. The scope of applicability of TS is an open, empirical question. I think it would be particularly interesting to assess the applicability of this framework to cases where (1) evidential reasoning is done in a more quantitative fashion and (2) in more “complex” cases, with nested claims and an increased diversity of facts and warrants mobilized. The latter cases would provide good means of testing the limits of the visual tractability of the diagrams constructed.

Concerning the former assertion, the application of TS to the case of Archezoa presents hypothesis evaluation in the historical sciences as informal (non-quantitative). This runs against existing accounts that emphasize probabilistic evaluations of historical hypotheses (Sober 1988; Cleland 2002). This is, however, compatible with viewing historical scientists as methodological omnivores (Currie 2015, 2018): the variety of their “evidential diet” prevents too strict a formalization of their evidential grounds. It might in principle be possible to use a probabilistic framework to reconstruct the case of Archezoa and track the series of shifts in evidential relevance and inferential security. This is something I haven’t proposed since (a) the concerned

³ The defense of the necessity to attend to the specificities of a given epistemic context to understand evidential reasoning in the historical sciences also pervades Chapman and Wylie’s (2016) own usage of TS and Currie’s recent book (2018).
scientists evaluate these hypotheses in a similarly informal fashion and (b) my limited knowledge prevents me to derive such precise quantitative values.

Framing hypothesis evaluation in probabilistic terms is often associated with a view of hypothesis evaluation as a comparative matter. The present case, however, presents the individual worth of the Archezoa hypothesis, cut off from possible competing hypotheses on the topic. This by no means denies the importance of competing hypotheses in the generation of lines of evidence. Alternative accounts may well have been at the origin of the generation of some of the lines of evidence against the Archezoa hypothesis. The case I present here displays the possibility of understanding hypothesis evaluation individually. By this, I do not mean that such evaluation is never comparative, but rather that, contrary to Cleland and Sober’s views, it doesn’t necessarily have to be the case. The case of Archezoa shows that historical scientists are capable of doing more than merely assessing comparative worth.

In his initial formulation, Toulmin presented his general framework as capable of accounting for disciplinary specificities. The locus of such specificities is found, according to him, in the backings employed in each field, as “the kind of backing we must point to if we are to establish [a warrant’s] authority will change greatly as we move from one field of argument to another” (Toulmin 2003, p. 96). The present usage of TS pushes this specificity at the level of individual claims.

While the case of Archezoa relies on methods and concepts generally employed in evolutionary biology, it is in the assessment of this particular claim that evidential relevance is generated for the facts, and that warrants and backings are mobilized. The roles played by these various bits of knowledge are heavily context-dependent. As already pointed out by Chapman and Wylie, “[w]arrants are themselves claims that depend on further substantive arguments; they are not purely formal inference rules, nor are they ‘self-authenticating’ as Toulmin puts it [Toulmin 2003, p. 91]” (Chapman and Wylie 2016, pp. 22–23). This also echoes Leonelli’s relational account of “data”, in which “the same objects may or may not be functioning as data, depending on which role they are made to play in scientific inquiry” (Leonelli 2015, p. 817; see also Chapman and Wylie 2016, p. 93; Canali under review). This context-specificity is, in my view, the right level of grain to capture the temporally unfolding opportunistic blending of evidential resources that occurs in the appraisal of a given claim.

Conclusions

This article aims to provide, after Chapman and Wylie in the context of archaeology, another application of Toulmin schemas to understand how evidential reasoning works in the historical sciences. This time, this framework was applied to reconstruct the case of Archezoa in evolutionary biology and successfully tracked the shifts from initial support to the eventual rejection of this claim. This was done by providing snapshots of the evidential situation at key points of the assessment of this hypothesis. I argued that TS are promising tools for historical and philosophical investigations. It is a precise descriptive tool that drives the identification of relevant
Evidential reasoning in historical sciences: applying Toulmin…

facts for historical reconstructions. The diagrams are visually tractable representations of situations that can nicely substitute for cumbersome textual descriptions. Finally, while TS are not explanatory tools, the snapshot approach to their construction is heuristically useful to uproot several investigations about the various epistemic moves it identifies. I thus argue that this minimal formalization of the “context of justification” provides an excellent starting point for various complementary forays into the “context of discovery”.

Acknowledgements I thank Adrian Currie for his invitation to contribute to this special issue and his overarching support. Thank you to the participants of the “Philosophy meets Paleobiology in the Badlands” workshop for their constructive feedback on an early version of this paper. I also thank the audiences of the SPSP, ICOP, Shesvie and of the UK iHPS network for having commented on presentations of this paper. I also thank John Dupré, Dan Nicholson, Sabina Leonelli and Alison Wylie for their feedback on the thesis chapter upon which this article is drawn. I am grateful to two anonymous reviewers for their sharp and constructive feedback. Research for this paper was funded by the University of Exeter and the private material support from Justine Frétigné.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Binford LR (1977) For theory building in archaeology. Academic, New York
Bonnin T, Lombard J (forthcoming) Situer l’analyse phylogénétique entre les sciences historiques et expérimentales. Philosophia Scientiae
Canali S (under review) Making evidential claims in epidemiology: three strategies for the study of the exposome
Cavalier-Smith T (1987a) The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. Ann NY Acad Sci 503:55–71
Cavalier-Smith T (1987b) The origin of eukaryote and archaeabacterial cells. Ann NY Acad Sci 503:17–54
Cavalier-Smith T (1991) Archamoebae: the ancestral eukaryotes? Biosystems 25:25–38
Cavalier-Smith T (1993) Kingdom protozoa and its 18 phyla. Microbiol Rev 57:953–994
Cavalier-Smith T (1998) A revised six-kingdom system of life. Biol Rev 73:203–266
Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. Int J Syst Evol Microbiol 52:297–354
Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (Meta-Algae). Philos Trans R Soc B 358:109–134
Cavalier-Smith T (2014) The neomuran revolution and phagotrophic origin of eukaryotes and cilia in the light of intracellular coevolution and a revised tree of life. Cold Spring Harb Perspect Biol. https://doi.org/10.1101/cshperspect.a016006
Cavalier-Smith T, Chao EE (1996) Molecular phylogeny of the free-living archezoan trepomonas agilis and the nature of the first eukaryote. J Mol Evol 43:551–562
Chang H (2012) Is water H2O? Evidence, realism and pluralism. Springer, Dordrecht
Chapman R, Wylie A (2016) Evidential reasoning in archaeology. Bloomsbury, London
Clark CG, Roger AJ (1995) Direct evidence for secondary loss of mitochondria in entamoeba histolytica. Proc Natl Acad Sci USA 92:6518–6521
Cleland CE (2001) Historical science, experimental science, and the scientific method. Geology 29:987–990
Cleland CE (2002) Methodological and epistemic differences between historical science and experimental science. Philos Sci 69:447–451

Springer
Cleland CE (2009) Philosophical issues in natural history and its historiography. In: Tucker A (ed) A companion to the philosophy of history and historiography. Wiley, Chichester, pp 44–62
Conan-Doyle A (1893) The adventure of the “Gloria Scott”. In: Conan-Doyle A (ed) The memoirs of Sherlock Holmes. George Newnes, London
Currie A (2015) Marsupial lions and methodological omnivory: function, success and reconstruction in paleobiology. Biol Philos 30:187–209
Currie A (2018) Rock, bone, and ruin: an optimist’s guide to the historical sciences. MIT Press, Cambridge
Currie A, Sterelny K (2017) In defence of story-telling. Stud His Philos Sci A 62:14–21
Currie A, Turner D (2016) Introduction: scientific knowledge of the deep past. Stud Hist Philos Sci Part A 55:43–46
DiMaria P, Palic B, Debrunner-Vossbrinck BA, Lapp J, Vossbrinck CR (1996) Characterization of the highly divergent U2 RNA homolog in the Microsporidian Vairimorpha Necatrix. Nucleic Acids Res 24:515–522
Ettema TJ (2016) Evolution: mitochondria in the second act. Nature 531:39–40
Germot A, Philippe H, Le Guyader H (1997) Evidence for loss of mitochondria in microsporidia from a mitochondrion-type HSP70 in Nosema Locustae. Mol Biochem Parasit 87:159–168
Hashimotó T, Sánchez LB, Shirakura T, Müller M, Hasegawa M (1998) Secondary absence of mitochondria in Giardia Lamblia and Trichomonas Vaginalis revealed by Valyl-TRNA synthetase phylogeny. Proc Natl Acad Sci USA 95:6860–6865
Jeffares B (2008) Testing times: regularities in the historical sciences. Stud His Philos Biol Biomed Sci 39:469–475
Keeling PJ, Doolittle WF (1996) Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Mol Biol Evol 13(10):1297–1305
Kosso P (2001) Knowing the past: philosophical issues of history and archaeology. Amherst, New York
Leonelli S (2015) Data-centric biology: a philosophical study. University of Chicago Press, Chicago
Li J, Katiyar SJ, Hamelin A, Visvesvara GS, Edlind TD (1996) Tubulin genes from AIDS-associated microsporidia and implications for phylogeny and benzimidazole sensitivity. Mol Biochem Parasitol 78:289–295
López-García P, Moreira D (2015) Open questions on the origin of eukaryotes. Trends Ecol Evol 30:697–708
Martin W, Müller M (1998) The hydrogen hypothesis for the first eukaryote. Nature 392:37–41
Morin L, Mignot JP (1995) Are archamoebae true archezoa? The phylogenetic position of Pelomyxa sp. as inferred from large subunit ribosomal RNA sequencing. Eur J Protistol 31:402
Müller M (2007) The road to hydrogenosomes. In: Martin WF, Muller M (eds) Origin of mitochondria and hydrogenosomes. Springer, Berlin, pp 1–12
Mulligan K, Correia F (2017) Facts. In: Zalta EN (ed) The Stanford encyclopedia of philosophy (Winter 2017 Edition). https://plato.stanford.edu/entries/facts/. Accessed on 8 Jan 2019
O’Malley MA (2010) The first eukaryote cell: an unfinished history of contestation. Stud Hist Philos Biol Biomed Sci 41:212–224
O’Malley MA (2013) When integration fails: prokaryote phylogeny and the tree of life. Stud His Philos Biol Biomed Sci 44:551–562
O’Malley MA (2016) Histories of molecules: reconciling the past. Stud Hist Philos Sci Part A 55:69–83
Roger AJ (1996) Studies on the phylogeny and gene structure of early-branching eukaryotes. Dissertation, Dalhousie University
Roger AJ (1999) Reconstructing early events in eukaryotic evolution. Am Nat 154:S146–S163
Safire W (2003) The way we live now: 1-26-03: on language; smoking gun. The New York Times Magazine. https://www.nytimes.com/2003/01/26/magazine/the-way-we-live-now-1-26-03-on-language-smoking-gun.html. Accessed on 31 July 2018
Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I (2017) The origin of metazoa: a unicellular perspective. Nat Rev Genet 18:498–512
Sober E (1988) Reconstructing the past: parsimony, evolution, and inference. MIT Press, Cambridge
Sogin ML (1989) Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. Integr Comp Biol 29(2):487–499
Sogin ML, Gunderson JH, Elwood HJ, Alonso RA, Peattie DA (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from Giardia Lamblia. Science 243:75
Suárez-Díaz E, Anaya-Muñoz VH (2008) History, objectivity, and the construction of molecular phylogenies. Stud His Philos Biol Biomed Sci 39:451–468
Toulmin SE (2003) The uses of argument, updt edn. Cambridge University Press, Cambridge
Turner D (2007) Making prehistory: historical science and the scientific realism debate. Cambridge University Press, Cambridge
Vossbrinck CR, Woese CR (1986) Eukaryotic ribosomes that lack a 5.8S RNA. Nature 320:287–288
Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR (1987) Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326:411–414
Wylie A (2017) How archaeological evidence bites back: strategies for putting old data to work in new ways. Sci Technol Hum Val 42:203–225
Wylie CD (this issue) Overcoming the underdetermination of specimens. Biol Philos

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.