HPLS (2024) 46:39
https://doi.org/10.1007/s40656-024-00634-6

ORIGINAL PAPER

The modern synthesis and "Progress" in evolution: a view from the journal literature

Charles H. Pence[1]

Received: 21 January 2023 / Accepted: 18 September 2024

Abstract The concept of "progress" in evolutionary theory and its relationship to a putative notion of "Progress" in a global, normatively loaded sense of "change for the better" have been the subject of debate since Darwin admonished himself in a marginal note to avoid using the terms 'higher' and 'lower.' While an increase in some kind of complexity in the natural world might seem self-evident, efforts to explicate this trend meet notorious philosophical difficulties. Numerous historians pin the Modern Synthesis as a pivotal moment in this history; Michael Ruse even provocatively hypothesizes that Ernst Mayr and other "architects" of the Synthesis worked actively to eliminate Progress from evolutionary biology's scientific purview. I evaluate these claims here with a textual analysis of the journals *Evolution* and *Proceedings of the Royal Society B* (a corpus of 27,762 documents), using a dynamic topic modeling approach to track the fate of the term 'progress' across the Modern Synthesis. The claim that this term declines in importance for evolutionary theorizing over this period can, indeed, be supported; more tentative evidence is also provided that the discussion of 'progress' is largely absent from the British context, emphasizing the role of American paleontology in the rise and fall of 'progress' in 20th-century evolutionary biology.

Keywords Evolution · Progress · Modern Synthesis · Ernst Mayr · Digital humanities · Textual analysis

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40656-024-00634-6.

✉ Charles H. Pence
charles.pence@uclouvain.be

[1] Institut supérieur de philosophie, Université catholique de Louvain, Place du Cardinal Mercier 14, 1348 Louvain-La-Neuve, Belgium

Published online: 06 November 2024

1 Introduction

Does evolutionary theory imply that the history of life on earth is a story of progress? On the one hand, there is a natural tendency to answer in the affirmative: after more than a billion years during which life consisted only of single-celled, prokaryotic organisms, the tree of life now hosts a bewildering array of complexity and diversity. And yet, on the other hand, this natural tendency is resisted by a variety of results from contemporary biological science. Natural selection offers us no guarantee of a march toward ever-increased adaptation – the risk of a population's becoming stranded on a local optimum has been noted from the very first days of evolutionary theory (Wright, 1932) and remains hotly debated today (Coyne et al., 1997, 2000; Gavrilets, 2010; Wade & Goodnight, 1998). Exactly what notion of complexity might be said to progressively increase over evolutionary time (and hence, how to fill in the details of our intuition that this indeed takes place) is also extremely difficult to describe in general (McShea, 1996; McShea & Brandon, 2010). The case becomes even more challenging when we recognize that adaptation is, and must be, relative to an environment, and compared to the kind of macroevolutionary time-scale on which we would like to invoke progress, environments are constantly changing. The vague feeling of a large, global trend toward progressive evolutionary transformation is therefore often cited as a classic misconception or illusion, impeding a clear understanding of evolutionary theory (Hull, 1988; Gould, 1989; but see Desmond, 2021).

1.1 Capital versus lowercase "progress"

Before we can get clear on exactly how evolution and progress are related, though, we have to begin by better understanding what we mean by progress in the first place. It's not at all clear that all the invocations of the term 'progress' in the last paragraph can plausibly be taken to refer to the same concept. Discussions of progress have, especially since the work of J. B. Bury (1920), tended to separate two different notions which might be at work for a particular thinker or discipline. First, we have a kind of local, limited, relative understanding of progress – the sense in which we might talk about progress in computing power or progress in automobile design. This is, of course, an interesting concept to study, and can offer a fruitful analytic lens to investigate directional trends; various instances of progress in this sense might be present in particular evolutionary histories. But this is not the kind of heavyweight, morally loaded concept of progress "toward the better" – again following Bury, often now baptized Progress with a capital-P – that is taken to be at least potentially problematic. This more profound construct of Progress is taken to underlie and legitimate a host of these other, more specific, local-scale instances of small-p progress. In Bury's words,

> the idea of the Progress of humanity, from which all these particular progresses derive their value, raises a definite question of fact, which man's wishes or labours cannot affect any more than his wishes or labours can prolong life

> beyond the grave. This idea means that civilization has moved, is moving, and will move in a desirable direction. (Bury, 1920, p. 2)

Put differently, whatever one particular electrical engineer or automobile designer might do to effect progress in the local sense will be independent from this larger question of Progress. If the world genuinely is driven "in a desirable direction," this will be so regardless of whether or not any particular person is around to witness it. Indeed, Bury claims that "belief in it is an act of faith" (1920, p. 4), though one that can be supported by a particular interpretation of history and a variety of claims about the nature of life and the physical or mental capacities of individuals.

As Michael Ruse tells the tale in his wide-ranging history of the notion of Progress in the sciences, *Monad to Man*, belief in Progress is a product of the 1700s and 1800s:

> following on this belief in the possibility of scientific advance, connected also with the way that the new science was demanding fundamental rethinkings of theology, a belief in the possibility of ongoing moral and social improvement – in short, a belief in Progress – arose in the eighteenth and nineteenth centuries. (Ruse, 1996, p. 23)

As we see in the examples of both Bury and Ruse, then, Progress has always been tied up with our views about the natural sciences, both because they themselves have seemed, at least to many, to be one of the places where human activity shows most clearly a tendency to progress in the local sense, and, more profoundly, because what those sciences tell us about ourselves and our nature might be taken to underwrite Progress in the more general sense.

1.2 Progress in evolution

More particularly, evolutionary theory – with its presentation of a novel understanding of the "deep time" history of life on earth – has often been seen as a testbed for both the question of small-p biological progress and large-P Progress in general. (All the more so, given that the theory arose within the historical period that Ruse identifies as key to the development of the concept as a whole.) Even in contemporary analyses, this deep connection persists. Matthew Nitecki, for example, approaches a discussion of Progress both within and outside the sciences by considering two competing views of evolutionary progress: Bertrand Russell, a staunch opponent, against Henri Bergson, perhaps its most eloquent defender (Nitecki, 1988, p. 4). Whether or not evolution shows us evidence of progress – for Russell, being oriented toward *prior* states of the environment, rather than the future, this was impossible, while for Bergson the *élan vital* drives evolution toward the emergence of mankind – is the fulcrum, for Nitecki, around which discussion of all of Progress turns.

Why does such a link seem so appealing? The ordering of organisms into the "great chain of being," as classically analyzed by Lovejoy, was an idea with us for thousands of years in what we might call our "folk-biological" understanding of the living world (Lovejoy, 1986; Martsa, 2003). Such a ranking from bacteria to man could, assuredly, be interpreted as evidence for Progress – even as a "crucial

experiment" for the theory of Progress. But the evidence is mixed. At the same time, as Ruse notes, our everyday experience with the biological world does not seem to be one that directly supports an idea of Progress, because the characteristics for which we would search are found only with difficulty in non-human organisms – first and foremost, "Progress is a theory or philosophy about human beings," and thus the hunt for support for Progress in the biological world must look like a hunt for "a rise in features on which human achievements depend" (Ruse, 1996, pp. 38–39).

Those features are occasionally found in our closest evolutionary cousins, but many of them tend not to fossilize. When they are discussed by biologists in their scientific works, this will often not be explicitly presented as a defense of Progress, even if the biologists at issue are among its staunch defenders. For those on the other side of the argument, skepticism about progress could arise from a more general suspicion about Progress, rather than the other way around. Take, for instance, David Hull, who wryly writes before launching an attack on the idea of evolutionary progress that "perhaps overpopulation, pollution, the greenhouse effect, the depletion of the ozone layer, AIDS, and the continued persecution by governments and organized religions alike of those who are most vulnerable are all illusions or inconveniences, but I do not think so" (Hull, 1988, p. 28).

Evidence for a belief or lack thereof in Progress will thus be teased out only with difficulty from the historical record of scientific practice. Ruse, for instance, takes great pains to reconstruct the social and cultural environment in which evolution was developed, arguing that it encouraged linkages between (biological) progress and Progress. Evolution, he claims, began its days as an enterprise loaded with concern for Progress:

> Progress was the philosophy of the day; Darwin was submerged in it; and the indications are that it seeped over – more precisely, flooded – into his science. Moreover, at that time people were thinking hard about the causes of Progress, and this too is Darwin's concern, especially as the distinctively Darwinian, relativistic notion of progress. (Ruse, 1996, p. 158)

Darwin himself seems to have been of two minds about the question. His writings contain both admonitions to himself to avoid all use of "higher" and "lower" (scribbled into the margin of his copy of the Progress-laden *Vestiges of the Natural History of Creation* by Robert Chambers; Mayr, 1988, p. 251) as well as a famous claim in the conclusion of the *Origin* that "as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend to progress towards perfection" (Darwin, 1859, p. 489). Robert Richards, contra Ruse's focus on Darwin's "relativistic" progress, has long emphasized this latter, Progress-inspired aspect of Darwin's thought, linking his work on progression in embryology to historical precedents in nineteenth-century morphological and developmental works (e.g., Richards, 1992, pp. 81–90).

But as many historians have noted, evolutionary theory in 1859 was a long way from becoming a free-standing scientific discipline, one which could appeal to a structure of academic societies, conferences, textbooks, and so forth. Instrumental in, or perhaps even synonymous with, this process of discipline formation was the Modern Synthesis, the crucial period over the middle of the twentieth century when

evolutionary biology saw itself become a clearly defined and at least somewhat independent discipline (Smocovitis, 1994a). Numerous historians have noted the Synthesis as being pivotal not only for biology as a whole, but for discussions of evolutionary progress in particular (Bowler, 1992; Provine, 1988; Smocovitis, 1996).

What exactly took place during this period? What did the "disciplining" of evolutionary biology, in Vassiliki Betty Smocovitis's terms, really consist in? Of course, this question is much too complex for a single article or even a single book (see, e.g., Smocovitis, 1994a, 1994b, 1996; Cain, 1994). But one aspect, highlighted especially by William Provine and Ruse, reflects directly on the question of Progress. A "proper" scientific discipline could not carry with it the kind of folk-biological commitment that Progress entailed; it had rather to emphasize "experimentation, or epistemic activities with the virtues of experimentation: hard evidence, the exercise of control, and the use of measurement" (Ruse, 1996, p. 446). Provine calls this the "evolutionary constriction," the process by which biologists agreed that some variables (like population size, structure, and so forth) were important for evolution in nature, and others – *especially* Progressive or purposive forces – "played no role at all" (Provine, 1988, p. 61).

Holding such a line about the content of the nascent discipline of evolutionary biology was made more difficult by the fact that many of the architects of the American synthesis, including Ernst Mayr (Hey, 2011; Mayr, 1990; Sloan, 1985), G. Ledyard Stebbins (Stebbins, 1969), and (perhaps especially) Julian Huxley (Huxley, 1942; Provine, 1988), were in fact believers in Progress. This was, however, a belief to be expressed elsewhere, outside the confines of academic evolutionary work. Further, and more speculatively, Ruse has argued that we should expect a difference between the American context represented by Mayr and the journal *Evolution* founded during the Synthesis, and the context in the United Kingdom: "the question of progress did not arise," he claims, for the English school that grew up around R. A. Fisher and E. B. Ford, while in the United States, "partly for personal reasons, partly because the paleontological record was always a central component of the American synthesis – the question of progress did arise" (Ruse, 1996, p. 448).[1]

A number of historical perspectives concerning this "evolutionary constriction" have been offered. First, we might underline the importance of an "internalist" reading based on the very content of evolutionary theory itself. Smocovitis, in the case of Julian Huxley, notes that "a view of natural selection as a mechanistic principle" or a "nonteleological" theory is difficult to reconcile with the "purposiveness or directionality [in] evolution" that he also had hoped to find (Smocovitis, 1996, p. 144). If scientific legitimacy entails being mechanistic and non-teleological, then any defense of evolutionary progress has to rely on human social or technological progress as a kind of "crowning achievement" of evolutionary theory, rather than natural selection or evolution in a narrow sense. H. J. Muller, for instance, writes that "if, then, we wish evolution to proceed in ways we consider progressive, we ourselves must become the agents that make it do so" (cited in Smocovitis, 1996,

[1] Smocovitis, notably, disagrees, seeing the US and UK responses as largely similar (Smocovitis, 1996, p. 207).

p. 158); Stebbins argues that "human evolution must be regarded in an entirely different context from the evolution of any other species of organism [because] our qualities spring just as much from our cultural heritage as from our biological nature" (Stebbins, 1969, p. 109).

Ruse, by contrast, emphasizes an externalist reading, claiming that the removal of Progress from the disciplinary content of evolution was intentionally pursued to increase evolution's scientific legitimacy, giving significant agency here to Mayr. His argument for this claim draws largely from three sources. First, he explores Mayr's correspondence (including with authors and reviewers) during the founding and his editorship of the journal *Evolution,* noting that it is loaded with a deep mistrust for anything that looks like teleology, directed trends, or orthogenesis – in short, anything that looks like a way to render more scientifically plausible the idea of Progress. Second, he investigates an apparent gap between the popularized writings of the architects of the Synthesis and their academic work, arguing that in the former they made space for the very notions of Progress that they hoped to banish from the latter. And finally, he claims that they carefully laid the groundwork for scientific investigations of at least progress, if not Progress, by constructing bodies of theoretical work on subjects like evolutionary trends. In this sense, while Progress could be avoided in the scientific literature itself, "a foundation for the progressionism [that they would return to in their other, popular writings] was certainly well dug" (Ruse, 1996, p. 449).[2]

Provine, for his part, offers yet a third reading, from a different externalist perspective, arguing that Julian Huxley only supported the concept of progress in evolution because of his preexisting social and ethical value commitments:

> I will argue that Huxley's idea of progress in evolution is merely the imposition of his cultural values upon evolution, that the modern synthesis in evolution is scarcely a synthesis at all and should be renamed, that ethics cannot be founded upon any notion of "progress" in evolution, and that the process of evolution gives no meaning in life. (Provine, 1988, p. 49)

In the case of Huxley, who was one of the most socially and culturally active architects of the Modern Synthesis, we are likely in a position to explore these values-and-science relationships in greater detail than that of other biologists of his day.

Adjudicating between these (and other) views of what happened to progress (and Progress) in the Modern Synthesis is clearly a complex enterprise. An apparently simple claim – that Progress disappeared from the discipline, properly speaking, of evolutionary biology – points to an intricate cluster of historical questions that are amenable to a wide variety of approaches. Most importantly, was this actually the case – was Progress actually eliminated from evolutionary thought, or not? When

[2] Smocovitis and Ruse directly disagree on these claims, Smocovitis rejecting any narratives in terms of "'threat' models or special 'interest' models for the rise of scientific disciplines," particularly as evolutionary biology never rose to the level of an independent academic department in the Anglo-American university context (Smocovitis, 1996, p. 68).

did this happen, and who was responsible? Was this story the same in the various national contexts in which evolutionary theory was discussed and developed?

2 Exploring progress empirically

In this paper, I will pursue another way in which we might try to answer these questions. As we saw the story reconstructed above, the elimination of Progress should be detectable over the middle decades of the twentieth century, the period of the Modern Synthesis. It could also, at least potentially, be different in the American and British contexts, with a claim by Ruse that the signal will be weaker in the UK than in the US. And if that change really shaped the contours of the field that developed, then it should be detectable not only "behind the scenes" – in the correspondence concerning the review or acceptance of manuscripts – or in popular presentations of evolutionary theory, but also in the content of the professional literature itself.

In short, this is a question ripe for the tools of the digital humanities, as references to the term 'progress'[3] in published documents should give us another, complementary way to quantify and explore these trends. My goal here, then, is to apply one of the tools of unsupervised text-mining (i.e., analyses that function without any prior training by the investigator) – in particular, dynamic topic modeling – to journal publications that might be able to illuminate the ways in which 'progress' was used over these pivotal years. In particular, if we use methods that can help us disentangle the various senses of 'progress' that appear in the pages of these journals, we should be able to learn more about in which contexts the term was used. Was 'progress' more common earlier rather than later in discussions of evolution, "driven out" by the Modern Synthesis? Is there a visible distinction between Progress and progress at this textual level, or no? Do invocations of 'progress' correlate with other aspects of the field of study or the methodology used? And how has the situation changed in the years since the Synthesis period?

Drawing on resources from the Sciveyor project (Ramsey & Pence, 2016; Pence, 2016), I constructed a corpus of journal publications containing the full text (i.e., not merely the abstracts) of the entire print run of both *Evolution* and *Proceedings of the Royal Society B* (hereafter *PRSB*).[4] The *Evolution* corpus contained 10,961 documents, published between 1947 and 2015; the *PRSB* corpus contained 16,801 documents, published between 1905 and 2014. The rationale for choosing these two journals was two-fold. First, and most obviously, *Evolution* was already the target of Ruse's analyses, and hence evaluating the role of 'progress' in its pages is of intrinsic interest as a test of his claims. Second, *PRSB* can give us a comparative case that differs from *Evolution* in at least two interesting ways: first, it is a UK-based rather

[3] I will refer to the term 'progress' in single quotes, to distinguish it from the local and global concepts of progress and Progress, respectively.

[4] I am unfortunately unable to share the corpus itself (i.e., the document full-text) due to the copyright-restricted licensing agreement negotiated with JSTOR. All other generated source code and data for this paper are available open-access at: https://doi.org/10.6084/m9.figshare.27060640.v1.

than a US-based journal, and second, it is of larger scope, including more contemporary life science than simply evolution-focused documents. In short, the two journals combined offer an excellent inferential base for exploring the kinds of hypotheses that I briefly introduced above.[5]

2.1 Methodology

The full text of both *Evolution* and *PRSB* was extracted from scans of the journal publications by JSTOR, and provided through their Data for Research service (Snyder, 2010)[6]. The text of these documents is of good quality, but by no means without errors. Optical character recognition, the conversion of PDF page images to plain text, was performed by JSTOR, and they no longer know which software, what version, or what settings were utilized to perform OCR for a given document (pers. comm.), hampering efforts to correct for systematic bias or error. Manual correction of mistakes in nearly twenty-eight thousand full-text documents is, of course, prohibitive; I was therefore left with no choice but to run analyses on the text as it was available, aided by some basic, automated cleaning, following standard practice in the field.

In order to prepare for further analysis, the text was first broken into words, and stop words (i.e., common words like "and" or "the") were removed. All words of less than three letters were also eliminated, as were any words containing a non-letter character. (These last two steps radically reduce the amount of noise resulting from incorrect OCR.) The text was also analyzed by the Python spaCy package (Honnibal et al., 2020) in order to tag each word with information about its part of speech; only nouns, verbs, adjectives, adverbs, proper names, and foreign words were preserved. Finally, each word was converted to its lemmatized form, which changes derived forms of words (like "swimming" or "swam") to a single, basic form (like "swim"). After performing this cleaning, the *Evolution* corpus is represented by 36,961,765 words (tokens), of which 226,143 are unique (types); the *PRSB* corpus consists of 43,415,852 tokens and 280,962 types.

If we want to understand the different ways in which 'progress' has been used in the literature, we need to not only look at the term 'progress' itself – we have to include the words that surround it, in an effort to determine how it is used in context. A common method for doing so is to use *topic modeling*, an analysis method which derives a series of probability distributions over the word-types in the corpus (these are "topics") and a probability distribution for each document representing

[5] Note that this corpus includes *all* documents published in these journals, which includes book reviews, obituaries, and prize notices. Because these kinds of texts are apt to show the orientation of their authors toward progress/Progress when they write in a more general or popular vein, they were intentionally left as part of the corpus, rather than being filtered out at this stage. As we will see below, the analysis methods are often able to identify and analyze them separately. (Front and back matter are also, less fortunately, included; they are filtered out by the topic model below and easily discarded.).

[6] Readers uninterested in the methodological details are welcome to skip to the next subsection.

the "mix" of topics used to select its words (Blei, 2012).[7] In an idealized model, we imagine that documents are "written" first by selecting a topic (using one of the the latter set of document distributions over topics), then by selecting a word from the topic thus chosen (using the corresponding topic distribution over word types); we then iterate until we have chosen enough words to make up the document. Notably, this fictional picture of document construction represents documents without any syntactic structure or ordering, as what is sometimes called a simple "bag of words." The number of topics, k, is chosen by the investigator in advance; other free parameters are inferred automatically by the model. As it happens, the "topics" that are derived in this way, when interpreted by investigating the collection of words to which they assign high probability, seem to track genuine topics, that is, subjects that documents are about (Boyd-Graber et al., 2017).

Topic modeling has been used to great effect throughout meta-science and the history and philosophy of science, to chart the changing focus of scientific publications (e.g., Blei & Lafferty, 2007; Griffiths & Steyvers, 2004; Malaterre et al., 2019). But traditional topic modeling comes with one significant disadvantage: there is no way for the topics to change over time. That is, a topic might describe a distribution over words that picks out, say, documents about "molecular biology" or "natural selection" – but the words that make up this topic do not change across the corpus. Such a topic model could let us see that discussions of natural selection were important at one time and became less important later on, for instance, but it would not be able to tell us that 'progress' was an important term for discussion of natural selection in the 1950s but then ceased to be one in the 1980s. In these topic models, the term 'progress' receives one constant probability in each topic, regardless of date of publication.

To perform a more complex analysis, we must turn from classic topic modeling to *dynamic topic modeling*. As developed by David Blei and John Lafferty (2006), dynamic topic modeling allows topics to *change* over time, modeling topic evolution from a sequential series of documents – in this case, the corpora of *Evolution* and *PRSB*, divided up into five-year blocks.[8] Blei and Lafferty's original code is freely

[7] This is not the only method, to be sure, that could be used to look at words in context. Another important such tool is co-occurrence analysis, which investigates which words occur significantly more often at a short distance from a "key word" of interest than they do elsewhere in documents within the corpus (Evert, 2005). An anonymous reviewer suggested a co-occurrence analysis of 'progress' as a way to check the robustness of the topic modeling results that I will present in what follows. The results of that analysis are available in Table S4. When we investigate those results, we see that co-occurrence analysis (in addition to being inherently noisier than topic modeling) offers us some insights that, while interesting, are more challenging to interpret. For instance, in *Evolution*, over the 1950–1954 period, 'progress' is strongly associated with 'ape', 'linnaeus', and 'chromosomes'. I take it that this offers some suggestive indications toward confirming the topic modeling results that I will present below (e.g., the presence of progress in large-scale theoretical discussions of evolutionary theory, as well as, early in the journal's print run, in classical genetics), but without the further contextual information offered by topic models, I think that it is risky to over-interpret such results. In any case, I see no radical divergence between the results of the co-occurrence analyses and the interpretations of the dynamic topic models that I will describe in this and the following section.

[8] A chunk size of five years was selected in order to ensure that there was enough text available at each time-point for the topic model to provide a meaningful analysis.

available, and has been integrated into the widely used Python-language topic modeling package Gensim (Řehůřek & Sojka, 2010).

As with classic topic modeling, the number of topics, *k*, is the most important parameter that must be selected by the investigator. A standard method for doing so is to evaluate what is called the *coherence* of topics – roughly, a measure of the extent to which words that the topic deems highly probable "travel together" within papers (Röder Both & Hinneburg 2015). To determine the value of *k* most suitable for our purposes here, topic models were trained with values of *k* ranging from 2 to 25,[9] the coherence of the model in each five-year block was evaluated, and then the average coherence of the models over all time periods was compared (supplemental Figure S3). In the case of both the *Evolution* and the *PRSB* topic models, the 23-topic model was selected as having the highest average coherence. The 23-topic models were therefore used for both journals.

2.2 Interpreting the models

For classic topic models, one of the primary ways in which we can analyze the "content" of a topic is to look at which words it picks out as highly probable; these are usually taken to be distinctive of what the topic is about. For dynamic topic models, these probabilities can change over time. On the one hand, this is a disadvantage: because the probabilities of words within topics aren't static, instead of reading *k* lists of most-probable words to interpret *k* topics, we now have to perform this interpretive step for *k* topics *at each time-point* in the model. As it turns out in this case, however, this process was no more difficult than the interpretation of static topic models, requiring a bit of finesse but retaining, in general, logical and thematic coherence across the topics over time.[10] And we get a compensatory advantage from models of this kind – namely, we can formulate more fine-grained questions. For our purposes, we can start by asking: Which topics give *at any point in time* a significant probability to 'progress,' and how does the importance of 'progress' *change over time within those topics*? We can reinforce this analysis, then, by looking at which documents, at each time, have a high probability value for those topics – that is, which documents are importantly "made up of" the words selected by a topic, and hence are "about" that topic. Overviews of these topics are presented in a series of figures, which show, in Figs. 1 and 2, the changing values for the probability of 'progress' within each topic, and then, in Figs. 3 and 4, that same value for the probability of 'progress' within each topic, multiplied by the probability that the corresponding topic appears within the corpus as a whole (and, hence, an indicator of the importance of 'progress' to the corpus overall, rather than to the topic taken in isolation). Citation information for top documents in each topic is provided in supplemental Table S2.

[9] Dynamic topic modeling is significantly more memory-intensive than traditional topic modeling, where *k* might be evaluated to as large as 200 or 250 topics. A 25-topic model was the largest that I could train on hardware readily available to me.

[10] I thank an anonymous reviewer for pushing me to consider this peculiarity of dynamic topic models.

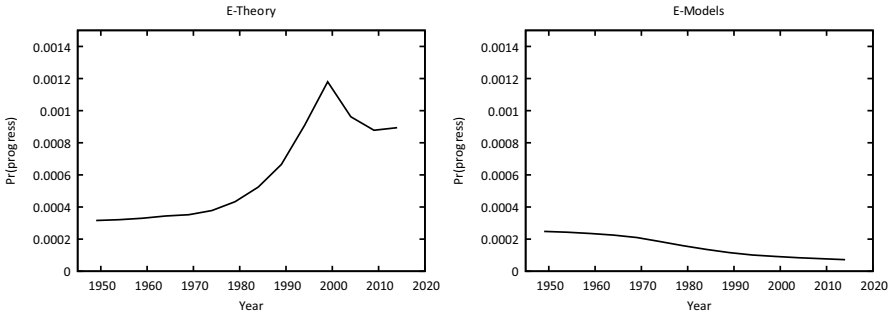

Fig. 1 The probability of 'progress' in the topics from *Evolution*, E-Theory and E-Models

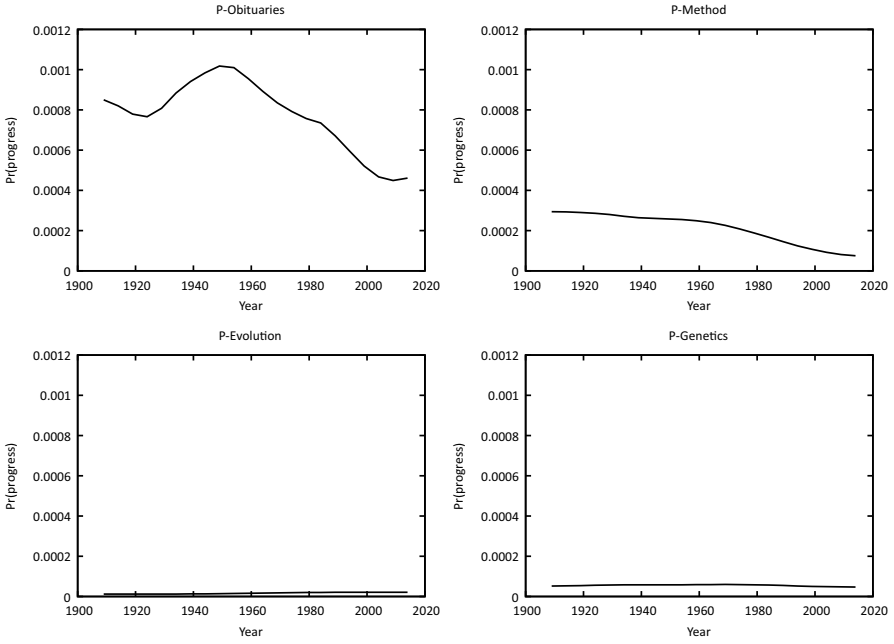

Fig. 2 The probability of 'progress' in the topics from *PRSB*, P-Obituaries, P-Method, P-Evolution, and P-Genetics. The P-Evolution curve is nearly indistinguishable from the x-axis

Let's start by looking at what the model of *Evolution* can tell us. Two topics of interest have significant[11] probability values for 'progress' at some point during

[11] I'm using "significant" in the colloquial sense, not in the sense of statistical significance. Since every topic is a probability distribution over *every* type, "progress" will have a non-zero probability in every topic. By "significant" or "important" topics for "progress," I mean those topics that, for at least some time points, rise above the background level of probability that "progress" receives in the mass of other topics.

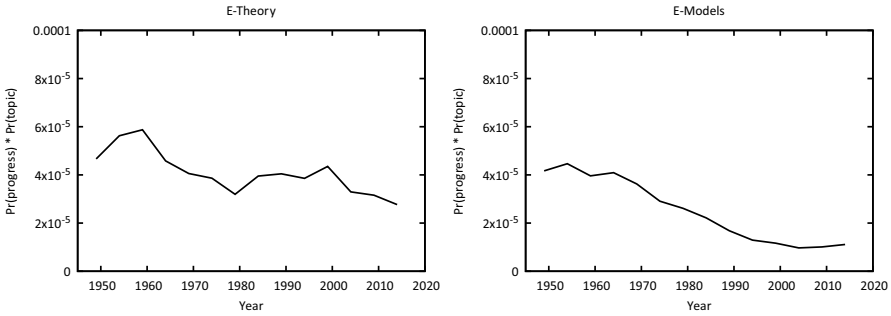

Fig. 3 The probability of 'progress' in the topics from *Evolution*, E-Theory and E-Models, multiplied by the probability that E-Theory and E-Models as a whole appear within the corpus

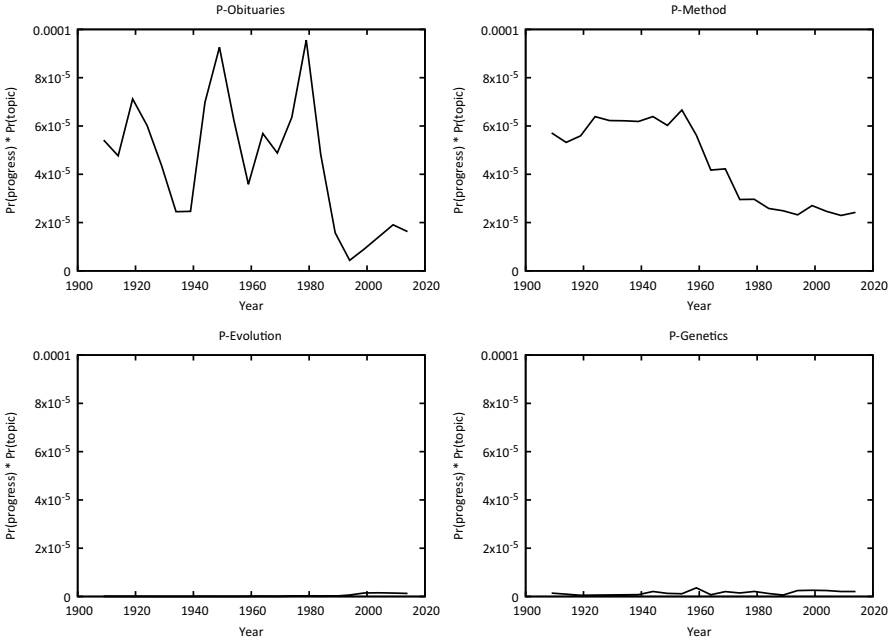

Fig. 4 The probability of 'progress' in the topics from *PRSB*, P-Obituaries, P-Method, P-Evolution, and P-Genetics, multiplied by the probability that each of those topics as a whole appears within the corpus. The P-Evolution curve is, again, nearly indistinguishable from the x-axis

the print run of *Evolution*: topic 13 and topic 17.[12] Topic 13, prior to around 1970, picks out highly theoretical papers in evolutionary biology, with a tendency as well

[12] Topic 14 also has a weak signal for "progress" since 2000, but this picks out only front and back matter, and has thus been ignored. A graph of "progress" for all topics in *Evolution* is found in supplemental Figure S1.

to involve paleontological data (the most probable document for this topic in the journal's first five years is "Towards a Modern Synthesis"; Dobzhansky, 1949). Its most probable words for 1960, for instance, are "evolution," "early," "know," "time," and "primitive," and other terms like "fossil," "process," and "development" appear further down the list (see Table S1 for the most probable words for all relevant topics at each time point). Especially after 1980 (one might hazard a guess that this coincides with a decline in the publication of these highly theoretical papers in the journal overall), the topic tends to pick out book reviews, especially of theoretical, historical, or philosophical works – for example, the most probable document for the topic in the 1990s is a book review of Robert Richards' *The Meaning of Evolution* (Crumly & Richards, 1992). In very recent years, it picks out documents dealing with the public reception, dissemination, or education of evolutionary theory, especially in the context of creationism (e.g., Burton, 2011). For ease of reference later on, I'll refer to this topic as E-Theory (with the "E" standing for *Evolution*).

Topic 17, on the other hand, describes mathematical modeling results for natural selection, fitness, and genetic drift in population genetics. Its most probable words always include "population" and "selection," and other very common terms include "model," "rate," "fitness," and "frequency." Papers that draw heavily on this topic almost all involve evolutionary models, dynamics, fixation probabilities and so forth (e.g., Brown & Vincent, 1987; Li, 1959). While some changes in the topic are of note (such as the fact that "model" only begins to appear as a key term in the mid-1970s), its most probable words remain largely static over time. I'll refer to this topic as E-Models.[13]

For *PRSB*, then, 'progress' is important to two topics: topics 1 and 14.[14] The first may be introduced very quickly: it picks out obituaries (such as the "Obituary Notices of Fellows Deceased" which were published annually), presidential addresses, and commemorative or memorial lectures (e.g., Ziman, 1969). In recent years, as that content leaves the journal, topic 1 becomes a reliable detector for the "Back Matter" in each issue. I'll call this topic P-Obituaries.

Topic 14 is most closely related with the methodology and results sections of papers. In early years, it includes words like "case," "result," "time," and "give," and more recently these are supplemented by "figure," "value," and "number." As one might expect, the most probable papers for this topic are therefore drawn from across the life sciences, exhibiting no strong affinity with a particular field (we see, for instance, botany; Blackman & Smith, 1911; physiology; Greenwood, 1918; allometry; Sholl, 1950; evolution; Smith 1960; and molecular biology; Murray, 1971). In very recent years, the topic narrows in on studies in mathematical ecology of environmental variability and change (e.g., Schoolmaster & Snyder, 2007). The unifying thread seems to be that this topic groups together some of the most

[13] Note that I do not intend the "E-Theory" and "E-Models" topics to distinguish "theory" from "models" according to any kind of philosophical criterion or theory of scientific modeling. These are simply terms that I've chosen to designate the content of the topics as I've interpreted them. The sense of "theory" that I have in mind is thus roughly "more abstract discussion of the nature of evolution," and that of "model" is roughly "mathematical model (largely from population genetics).".

[14] A graph of "progress" for all topics in *PRSB* is found in supplemental Figure S2.

mathematized or formalized topics at each time point in the journal. I'll refer to this topic as P-Method.

Finally, it is topic 3 that is most important for evolution in *PRSB*, and topic 16 that is most closely connected with the study of inheritance, heredity, and genetics. Topic 3 ("P-Evolution") includes probable terms like "specie" (the lemmatized version of "species"), "genus," "character," "population," and "evolution," while topic 16 ("P-Genetics") includes probable terms like "gene," "population," "strain," and "mutation."

These, then, are the topics for which 'progress' is important (E-Theory, E-Models, P-Obituaries, and P-Method), and the topics in which evolutionary theory or genetics seem to be discussed, but without significant reference to 'progress' (P-Evolution and P-Genetics).

Let's explore, then, how 'progress' appears (and fails to appear) across these six topics. Figures 1 and 2 indicate the probability over time that the term 'progress' would be selected by each of the six topics under discussion. While these values should only be compared relatively, they nonetheless let us see the importance of 'progress' within each of these topics and the changes in that importance over time; values very close to zero indicate that the term has nearly no probability to be selected in the context of that topic. Figures 3 and 4, then, are those same probability values, multiplied by the likelihood that documents in the corpus as a whole select each topic. (Colloquially, we might say that Figs. 1 and 2 describe probabilities of 'progress' where the unit of analysis is the topic; Figs. 3 and 4 move the unit of analysis to the corpus.) In *Evolution*, 'progress' seems to be weakly distinctive of E-Theory until about 1980, then becomes strongly distinctive of that topic up until the present (Fig. 1). This observation should be tempered, however, by the fact that the E-Theory topic as a whole becomes less prevalent in the corpus, as can be seen by the shape of the combined probability curve for E-Theory in Fig. 3. (Notably, this is the only topic for which the shape of the curve in Figs. 1 and 2 differs strongly from the shape of the corresponding curve in Figs. 3 and 4.) Notions of progress (or Progress) therefore seem to have been a persistent, if minor, topic of theoretical discussion in *Evolution* in the early years of Modern Synthesis evolutionary theorizing, which experience a resurgence within a much more limited scope in the contemporary period. 'Progress' was also very weakly distinctive of E-Models in the very earliest days of the journal. But this dies out quickly, becoming minimal by 1975 and almost entirely undetectable after 1990.

In *PRSB*, the term 'progress' is always very strongly distinctive of P-Obituaries (Fig. 2). It is more weakly distinctive of P-Method, with probability values continually falling across the journal's run. Things are much clearer in the case of P-Evolution and P-Genetics, where 'progress' has almost no detectable signal.

Finally, it is important to check these probabilities as revealed by the topic model against one more important piece of data from the two journals: the overall frequency with which 'progress' is used in journal publications (Fig. 5). While there is a slight peak in the first few years of *PRSB* here, the values for the two journals are reasonably comparable (between 0.1 and 0.4 occurrences of 'progress' per document). This confirms that since we are dealing with similar (and not too small) numbers of term occurrences in the two corpora, the probabilities described by the

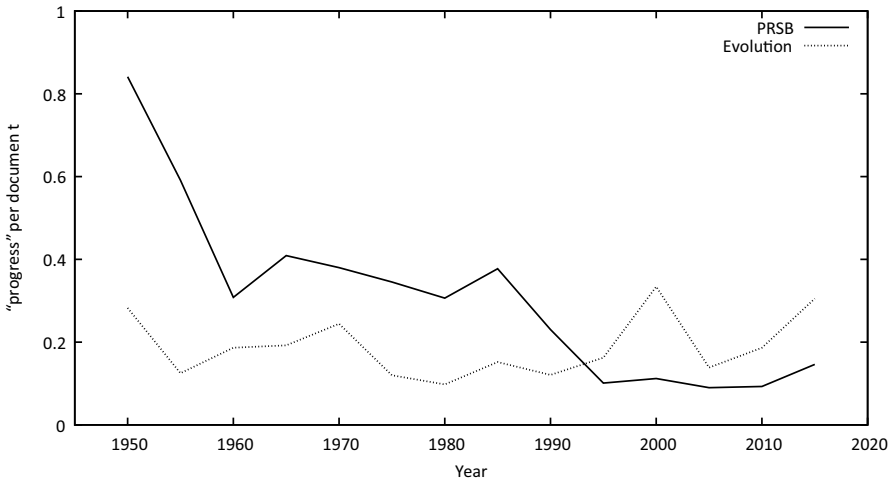

Fig. 5 The number of occurrences of 'progress' per document in the corpus, plotted over time. Values presented here are combined into five-year averages, for the same year ranges as the dynamic topic models (i.e., 1947–1949, 1950–1954, etc.)

topic model can be trusted – we have not asked the model to extrapolate from too small a sample size, the term appearing at the lowest in around one out of every ten documents.

3 Discussion

How can we use these empirical results to evaluate the historical claims that are my focus here? Let's begin with *Evolution* and, in particular, the idea that Progress was eliminated from the nascent academic work of constructing evolutionary models. The data presented above confirm that there was, indeed, a weak but notable tendency to describe evolutionary models in "progressive" terms when *Evolution* was founded. Enough documents were published doing so to make 'progress' weakly distinctive of E-Models until around 1975. But we also see that this influence disappears from the pages of *Evolution* almost entirely by around 1990 – evolutionary modeling indeed lost whatever association that it might have had with 'progress' in the first decades of the twentieth century. Of course, a claim like Ruse's that this was an *intentional* change, with Mayr as its agent, cannot be confirmed by this result, but the data are at least *consistent* with Ruse's assertion, and an empirical understanding of what actually happened in the pages of *Evolution* is necessary to be able to entertain or evaluate a hypothesis about intentions.

The trend as seen in E-Theory is also interesting. The term 'progress' does seem to have appeared in a significant way in the more theoretical parts of evolutionary biology in the journal's first years. We might read this as confirming the observations of Smocovitis (1996, 2016), Ruse (1996), and others that, when many of the architects of the Modern Synthesis expanded their target audience and wrote for

a more general audience (which, in this period, often involved the construction of broader, more theoretical arguments about evolution), they reintroduced the belief in Progress that many of them in fact held, but did not discuss in the kinds of papers that might appear under E-Models. Equally interesting is the topic's recent history. The tenor of the most popular documents in E-Theory (Table S2) shows that evolutionary theorists have dedicated more pages of their journals to the public dissemination and defense of evolutionary theory (with a notable nod to the growing global battle against the teaching of creationism in the science classroom). We see, in parallel, that discussion of progress (here almost certainly with a lowercase p) has taken on an increasingly outsized role as a tool for the historical, philosophical, or popular contextualization of evolutionary thought.

Now, let's turn to *PRSB*. 'Progress' was, when such content made up an important part of the journal, a common laudatory phrase applied to recently deceased scientists or the recipients of scientific prizes. And, although it seems that this kind of talk has fallen out of favor, methodologies (especially formal or mathematical methodologies) were described as constituting 'progress' in various contexts.

When it comes to evolutionary theory, however, 'progress' is almost entirely absent from *PRSB*, as we see in Fig. 2. The analysis here thus offers tentative evidence that could prompt further study of Ruse's other claim: there may indeed be a difference in the way in which 'progress' appeared within the US and the UK contexts. While *Evolution* is marked by shifts in the importance of 'progress' over time in topics like E-Theory and E-Models, the term simply seems to be a non-issue in any topic in *PRSB* that deals substantially with evolutionary theory. Removed, perhaps, from the American interest in paleontology and orthogenesis, British biologists seemed to have much less investment in the debate over what to do with evolutionary progress and its relationship to Progress. Of course, this empirical claim must be tempered by the international nature of journal publication. Assuredly, US authors published in the pages of *PRSB* and UK authors published in the pages of *Evolution*. But the difference between the signal of 'progress' in the two journals remains, nonetheless, suggestive.

3.1 Small-P progress and large-P Progress

Clearly, one important ambiguity remains in these analyses: distinguishing small-p biological progress from large-P Progress in evolutionary theory. It seems straightforward enough that *neither* would have been an important question for *PRSB* in the context of evolutionary theory: if discussions of evolution don't invoke the term at all, then it's immaterial in which sense they're (not) invoking it.

But for *Evolution*, the question is more important. When we look at the shifts in the graphs of E-Theory and E-Models in Fig. 1, how should we understand those curves? Can the topic model begin to give us some information about whether these uses of 'progress' – the uses kept in the early days of E-Theory and slowly discarded from E-Models – are more likely to refer to progress or to Progress?

Extracting such nuance is somewhat difficult. In part, this is a result of the fact, noted by Ruse and discussed above, that commitments to Progress will be difficult

to discern in any practicing biological work. A journal publication is simply not the space in which a scientist would normally explore their underlying philosophical commitments, no matter how sincerely and deeply held. There are also technical constraints arising from the kind of analysis undertaken. Because topic models, as noted above, treat a document as a simple "bag of words," disregarding their order and internal relationships, information about which words are closer to or farther from one another is discarded, some of which might be crucially important in this case.

One thing that we can do, however, is inspect the changes in words over time other than 'progress' in the two topics E-Theory and E-Models. The changes in the probabilities of every word between 1947 and 1979 were computed within these two topics – these two dates corresponding to the founding of the journal and the point at which, roughly, E-Models had ceased to discuss 'progress' entirely and E-Theory was becoming a topic about philosophy, history, and science in culture. The top five most extreme increases and decreases in probability within those topics over this period, along with a few further interesting words drawn from the top twenty, are presented in Table 1. In E-Theory, we see fairly clearly the topic's shift to review content: the largest-increase word between 1947 and 1979 is "book," and it is accompanied by "chapter," "review," and "press." The largest-decreased words, on the other hand, tell an interesting story. We see terms like "primitive," "modern," and "form" which, especially when taken with "time" and "genera," seem to underline the fact that the kind of theoretical papers that were important for this topic (and hence were invoking 'progress') were paleontological. (While it might be tempting to view "primitive" and "modern" here as indications of Progress, it is far more likely that they are being used as paleontological terms of art.)

The disappearance of 'progress' in E-Theory, then, looks largely to have followed the decline of a kind of progressive (here with a small-p) perspective toward

Table 1 The top-five most extreme increased and decreased probabilities, and (after the "*and*") selected further changes from the top twenty, for words in the topics E-Theory and E-Models, between the 1947–1949 time-block and the 1975–1979 time-block

Topic	Increasing Words	Decreasing words
E-Theory	book: + 0.003514	time: − 0.001870
	theory: + 0.002712	primitive: − 0.001693
	chapter: + 0.002214	know: − 0.001582
	evolutionary: + 0.001942	genera: − 0.001557
	biology: + 0.001718	rodent: − 0.001523
	and	*and*
	review: + 0.001681	form: − 0.001298
	press: + 0.001376	man: − 0.001212
	major: + 0.001098	modern: − 0.000710
E-Models	model: + 0.005866	gene: − 0.005723
	selection: + 0.005175	type: − 0.003758
	fitness: + 0.003236	series: − 0.002843
	rate: + 0.002562	adaptive: − 0.002205
	evolution: + 0.002529	fact: − 0.002152
	and	*and*
	equation: + 0.001189	natural: − 0.001736
	probability: + 0.001131	normal: − 0.001348
	random: + 0.001120	great: − 0.001335

the theoretical study of paleontology. At the very least, there are no terms here that seem to strongly indicate any change in a commitment to Progress.

When we turn to E-Models, we see, in the increasing column, the steady mathematization of theorizing in evolutionary biology, with "rate," "equation," "probability," and "random" all rising in probability in this topic over this period. Decreasing words include "gene" and "factor" – indicating a decline in evolutionary models phrased directly in classical-genetic terms (though "allele" is found among the increasing words, replacing the terminology of classical genetics with that of contemporary genetics; see Table S3). The decline in the probability of "natural" is probably best explained as the replacement of the term "natural selection" by other, more specialized kinds of selection (directional, stabilizing, etc.) in evolutionary modeling. But the appearance of "normal" and "great" in this list is, at the very least, a tantalizing signal of a potential removal of a normative dimension from evolutionary modeling during this time. These two words were not necessarily signals of Progress, to be sure. But, I claim, they can incline us to formulate the hypothesis that a removal of normative or evaluative dimensions from biological theorizing might be an element of a broader trend against folk-biological notions in "legitimate" scientific work (see discussion of folk-biology in Machery, 2008).

While this is assuredly not a detection of the rejection of Progress in the literature – something that we could not have really expected in any event – it nonetheless serves as support for the idea that the issue of "non-scientific" language was a live one for evolutionary theory during this period.

4 Conclusion

Let's step back and take stock. We saw at least three different ways in which we might interpret the shift in belief about evolutionary progress that nearly all historians agree took place over the course of the Modern Synthesis. They are, however, radically different – ranging from the content of evolutionary theory and the interpretation of natural selection itself to the prior value commitments of an author like Julian Huxley.

In this paper, I have attempted to evaluate a new set of evidence for these claims, drawn directly from some nearly twenty-eight thousand journal publications in the journals *Evolution* and *PRSB*. Regardless of which of these interpretations of the trend in 'progress' is correct, we should be able to see the signal of this shift at a large scale across these important publication outlets. The method used here – dynamic topic modeling – is particularly suitable for assessing the ways that 'progress' is used in the journal literature. Progress in any sense, we see, was a minor theme in the early years of *Evolution* in two contexts: paleontology and evolutionary modeling. Especially in the latter, talk of 'progress' entirely vanished from the journal by around 1975. In the former, we see that while talk of 'progress' continued and even increased, this was not because *paleontological* papers talked about progress (or Progress) – rather, evolutionary biologists began to use the term to discuss the historical, philosophical, and educational contextualization of their scientific work, and papers with this "progressive" approach to paleontology seem to disappear.

PRSB, on the other hand, serves as a tentative and limited, but nonetheless important test-case for a comparative analysis of the United States and United Kingdom, and we find that progress in any guise never seemed to be a genuine issue in evolutionary theory in *PRSB*. While it would regularly be referred to in the laudatory context of scientific prizes and obituaries, topics concerning evolution and genetics see essentially no mention of 'progress.'

Finally, for a variety of reasons both methodological (the bag-of-words nature of topic models) and theoretical (the low likelihood that biologists would speak of Progress in their journals), we saw little way to test broader assertions about Progress here. What we did see was a bit of potential evidence that the preoccupation with normative or moralizing language extends beyond worries about Progress; this claim would require the support of close reading to be developed further.

To conclude, then, I want to first discuss the relationship between these digital methods and traditional close readings, and finally explore avenues for future work. It is important to underline that the analysis that I offer here is limited in important ways. It does not, and could not, settle the question of how we should interpret the conflicting and complex heritage of Progress in Darwin's own writings (Richards & Ruse 2016; though this could be a promising place for further study, see, e.g., Jiménez-Pazos, 2021). More broadly, it is by no means intended to replace the kind of painstaking close reading and archival study undertaken in the history of biology. Those close readings would be required to defend, for instance, Ruse's contentious claim that eliminating Progress is something that Mayr *intended* to effect in the pages of *Evolution*. What digital methods add, then, is twofold. First, they can confirm that, in fact, such a change in the journal did come to pass.

Second, and turning to future work, we've seen in several cases that these analyses can help us answer what J. T. Burman (2018, 300) has identified as one of the key questions for digital scholarship in the history of science: "How can these tools help you to see what you are interested in such that you can then make better judgments about what to select for further research?" While we were interested here in the shift in 'progress,' changes in E-Models seem to highlight both the process of formalization of evolutionary theory as well as the shift to thinking explicitly in terms of "models" in scientific practice, both of which deserve further study. The decline in invocations of apparently morally loaded terminology in E-Models also would merit further exploration, supplemented by detailed textual study. Only six topics from these two topic models were discussed here, and of those only four in real depth – there are forty others that one might use to shed light on a host of other questions. We must, to be sure, be careful not to be philosophically or historiographically naive in our use of digital humanities tools (Pence, 2022). Nevertheless, the kind of large-scale exploration and confirmation offered here of the disappearance of 'progress' from evolutionary theory over the course of the Modern Synthesis is, I hope, indicative of the power and possibility that these approaches open for interested philosophers and historians of science.

Acknowledgements Thanks to JSTOR for access to the corpora used in this project, by agreement with Sciveyor. Thanks to Jesse Powell for extensive discussion of the ideas early on in the project, and to two anonymous reviewers, Betty Smocovitis, Hugh Desmond, and Michael Ruse for comments.

Funding C.H.P. was supported in this project by the Fonds de la Recherche Scientifique—FNRS under grant no. F.4526.19, and by the Commission mixte permanente Wallonie-Bruxelles Québec under grant no. RECH-INNO-05.

Declarations

Conflict of interest The author has no conflict of interest to declare.

References

Blackman, Frederick Frost, & Smith, A. M. (1911). Experimental researches on vegetable assimilation and respiration IX—On assimilation in submerged water-plants, and its relation to the concentration of carbon dioxide and other factors. *Proceedings of the Royal Society of London Series B, Containing Papers of a Biological Character, 83*(565), 389–412. https://doi.org/10.1098/rspb.1911.0022

Blei, David M., & Lafferty, John D. (2006). Dynamic Topic Models. In *ICML*' 06, (pp. 113–20). ACM Press. https://doi.org/10.1145/1143844.1143859.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77. https://doi.org/10.1145/2133806.2133826

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of Science. *The Annals of Applied Statistics, 1*(1), 17–35. https://doi.org/10.1214/07-AOAS114

Bowler, P. J. (1992). *The eclipse of Darwinism: Anti-Darwinian evolution theories in the decades around 1900*. Johns Hopkins University Press.

Boyd-Graber, J., Yuening, Hu., & Mimno, D. (2017). Applications of topic models. *Foundations and Trends in Information Retrieval, 11*(2–3), 143–296. https://doi.org/10.1561/1500000030

Brown, J. S., & Vincent, T. L. (1987). Coevolution as an evolutionary game. *Evolution, 41*(1), 66. https://doi.org/10.2307/2408973

Burman, J. T. (2018). Digital methods can help you if you're careful, critical, and not historiographically naïve. *History of Psychology, 21*(4), 297–301. https://doi.org/10.1037/hop0000112

Burton, E. K. (2011). Evolution and creationism in Middle Eastern education: A new perspective. *Evolution, 65*(1), 301–304. https://doi.org/10.1111/j.1558-5646.2010.01113.x

Bury, J. B. (1920). *The idea of progress: An inquiry into its origin and growth*. Macmillan.

Cain, J. (1994). Ernst Mayr as community architect: Launching the society for the study of evolution and the journal evolution. *Biology & Philosophy, 9*(3), 387–427. https://doi.org/10.1007/BF00857945

Coyne, J. A., Barton, N. H., & Turelli, M. (1997). Perspective: A critique of sewall wright's shifting balance theory of evolution. *Evolution, 51*(3), 643–671.

Coyne, J. A., Barton, N. H., & Turelli, M. (2000). Is Wright's shifting balance process important in evolution? *Evolution, 54*(1), 306–317.

Crumly, C. R., & Richards, R. J. (1992). Review of "The meaning of evolution: The morphological construction and ideological reconstruction of Darwin's theory". *Evolution, 46*(5), 1592. https://doi.org/10.2307/2409968

Darwin, C. (1859). *On the Origin of Species* (1st ed.). John Murray.

Desmond, H. (2021). The selectionist rationale for evolutionary progress. *Biology & Philosophy, 36*(3), 32. https://doi.org/10.1007/s10539-021-09806-1

Dobzhansky, T. (1949). Towards a modern synthesis. *Evolution, 3*(4), 376–377. https://doi.org/10.2307/2405724

Evert, Stefan. (2005). *The statistics of word cooccurrences: Word pairs and collocations*. Ph.D Dissertation, University of Stuttgart.

Gavrilets, S. (2010). High-dimensional fitness landscapes and speciation. In M. Pigliucci & G. B. Müller (Eds.), *Toward an extended evolutionary synthesis* (pp. 45–79). The MIT Press.

Gould, S. J. (1989). *Wonderful life*. W. W. Norton.

Greenwood, Major. (1918). On the efficiency of muscular work. *Proceedings of the Royal Society of London Series B, Containing Papers of a Biological Character. 90*(627), 199–214. https://doi.org/10.1098/rspb.1918.0010

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*(suppl 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

Hey, J. (2011). Regarding the confusion between the population concept and Mayr's 'population thinking.' *Quarterly Review of Biology, 86*(4), 253–264.
Honnibal, Matthew, Montani, I., Van Landeghem, S., & Boyd, A. (2020). SpaCy: Industrial-strength natural language processing in Python. https://doi.org/10.5281/zenodo.1212303.
Hull, D. L. (1988). Progress in ideas of progress. In M. H. Nitecki (Ed.), *Evolutionary progress* (pp. 27–48). University of Chicago Press.
Huxley, J. S. (1942). *Evolution: The modern synthesis*. Allen and Unwin.
Jiménez-Pazos, B. (2021). Darwin's perception of nature and the question of disenchantment: A semantic analysis across the six editions of *On the origin of species*. *History and Philosophy of the Life Sciences, 43*(2), 57. https://doi.org/10.1007/s40656-021-00373-y
Li, C. C. (1959). Notes on relative fitness of genotypes that forms a geometric progression. *Evolution, 13*(4), 564. https://doi.org/10.2307/2406137
Lovejoy, A. O. (1986). *The great chain of being*. Harvard University Press.
Machery, E. (2008). A plea for human nature. *Philosophical Psychology, 21*(3), 321–329. https://doi.org/10.1080/09515080802170119
Malaterre, C., Chartier, J.-F., & Pulizzotto, D. (2019). What is this thing called philosophy of science? A computational topic-modeling perspective, 1934–2015. *HOPOS, 9*(2), 215–249. https://doi.org/10.1086/704372
Martsa, S. (2003). On the conceptualization of animal taxonomies across cultures. In L. Komlósi, P. Houtlosser, & M. Leezenberg (Eds.), *Communication and culture: Argumentative, cognitive and linguistic perspectives* (pp. 131–141). Sic Sat.
Mayr, E. (1988). *Toward a new philosophy of biology: Observations of an evolutionist*. Harvard University Press.
Mayr, E. (1990). When is historiography whiggish? *Journal of the History of Ideas, 51*(2), 301–309.
McShea, D. W. (1996). Metazoan complexity and evolution: Is there a trend? *Evolution, 50*(2), 477–492.
McShea, D. W., & Brandon, R. N. (2010). *Biology's first law: The tendency for diversity and complexity to increase in evolutionary systems*. University of Chicago Press.
Murray, James Dickson. (1971). On the molecular mechanism of facilitated oxygen diffusion by haemoglobin and myoglobin. *Proceedings of the Royal Society of London Series B Biological Sciences 178*(1050), 95–110. https://doi.org/10.1098/rspb.1971.0054
Nitecki, M. H. (Ed.). (1988). *Evolutionary progress*. University of Chicago Press.
Pence, C. H. (2016). RLetters: A web-based application for text analysis of journal articles. *PLoS ONE, 11*(1), e0146004. https://doi.org/10.1371/journal.pone.0146004.
Pence, C. H. (2022). Testing and discovery: Responding to challenges to digital philosophy of science. *Metaphilosophy, 53*(2–3), 238–53. https://doi.org/10.1111/meta.12549.
Provine, W. B. (1988). Progress in evolution and meaning in life. In M. H. Nitecki (Ed.), *Evolutionary progress* (pp. 49–74). University of Chicago Press.
Ramsey, G., & Pence, C. H. (2016). EvoText: A new tool for analyzing the biological sciences. *Studies in History and Philosophy of Biological and Biomedical Sciences, 57*, 83–87. https://doi.org/10.1016/j.shpsc.2016.04.003.
Řehůřek, Radim & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of LREC 2010 workshop new challenges for NLP frameworks*. (pp. 45–50). University of Malta. http://is.muni.cz/publication/884893/en.
Richards, R. J. (1992). *The meaning of evolution: The morphological construction and ideological reconstruction of Darwin's theory*. University of Chicago Press.
Röder, Michael, Andreas Both, A. & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on web search and data mining*, (pp. 399–408). ACM. https://doi.org/10.1145/2684822.2685324.
Ruse, M. (1996). *Monad to man: The concept of progress in evolutionary biology*. Harvard University Press.
Ruse, M., & Richards, R. J. (2016). *Debating Darwin*. University of Chicago Press.
Schoolmaster, D. R., & Snyder, R. E. (2007). Invasibility in a spatiotemporally fluctuating environment is determined by the periodicity of fluctuations and resident turnover rates. *Proceedings of the Royal Society B: Biological Sciences, 274*(1616), 1429–1435. https://doi.org/10.1098/rspb.2007.0118
Sholl, D. A. (1950). The theory of differential growth analysis. *Proceedings of the Royal Society of London Series B - Biological Sciences, 137*(889), 470–474. https://doi.org/10.1098/rspb.1950.0056
Sloan, P. R. (1985). Essay review: Ernst Mayr on the history of biology. *Journal of the History of Biology, 18*(1), 145–153.

Smith, J. M. (1960). Continuous, quantized and modal variation. *Proceedings of the Royal Society of London Series B Biological Sciences, 152*(948), 397–409. https://doi.org/10.1098/rspb.1960.0047

Smocovitis, V. B. (1994a). Disciplining evolutionary biology: Ernst Mayr and the founding of the society for the study of evolution and *evolution* (1939–1950). *Evolution, 48*(1), 1–8. https://doi.org/10.2307/2409996

Smocovitis, V. B. (1994b). Organizing evolution: Founding the society for the study of evolution (1939–1950). *Journal of the History of Biology, 27*(2), 241–309. https://doi.org/10.1007/BF01062564

Smocovitis, V. B. (1996). *Unifying biology: The evolutionary synthesis and evolutionary biology*. Princeton University Press.

Smocovitis, V. B. (2016). The unifying vision: Julian Huxley, evolutionary humanism, and the evolutionary synthesis. In G. Somsen & H. Kamminga (Eds.), *Pursuing the unity of science: Ideology and scientific practice between the Great War and the Cold War* (pp. 30–49). Routledge.

Snyder, Ron. (2010). JSTOR data for research API. Notre Dame, IN, February.

Stebbins, G. L. (1969). *The basis of progressive evolution*. University of North Carolina Press.

Wade, M. J., & Goodnight, C. J. (1998). Perspective: The theories of Fisher and Wright in the context of metapopulations: When nature does many small experiments. *Evolution, 52*(6), 1537–1553.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics, 1*, 356–366.

Ziman, J. M. (1969). The Rutherford memorial lecture, 1968. Some problems of the growth and spread of science into developing countries. *Proceedings of the Royal Society of London Series A, Mathematical and Physical Sciences, 311*(1506), 349–69.