Conceptions of Good Science in Our Data-Rich World

KEVIN C. ELLIOTT, KENDRA S. CHERUVELIL, GEORGINA M. MONTGOMERY, AND PATRICIA A. SORANNO

Scientists have been debating for centuries the nature of proper scientific methods. Currently, criticisms being thrown at data-intensive science are reinvigorating these debates. However, many of these criticisms represent long-standing conflicts over the role of hypothesis testing in science and not just a dispute about the amount of data used. Here, we show that an iterative account of scientific methods developed by historians and philosophers of science can help make sense of data-intensive scientific practices and suggest more effective ways to evaluate this research. We use case studies of Darwin's research on evolution by natural selection and modern-day research on macrosystems ecology to illustrate this account of scientific methods and the innovative approaches to scientific evaluation that it encourages. We point out recent changes in the spheres of science funding, publishing, and education that reflect this richer account of scientific practice, and we propose additional reforms.

Keywords: hypothesis testing, data-intensive science, iteration, science education, science funding

Scientists have been debating for centuries the nature of proper scientific methods, especially the role of hypothesis testing in scientific practice (Laudan 1981). These debates are being reinvigorated as many fields of science, including high-energy physics, astronomy, public health, climate science, environmental science, and genomics, are increasingly using data-intensive approaches (Bell et al. 2009, Baraniuk 2011, Winsberg 2010, King 2011, Porter et al. 2012, Mattman 2013, Khoury and Ioannidis 2014, Katzav and Parker 2015). *Data-intensive science* has been described as research in which the capture, curation, and analysis of (usually) large volumes of data are central to the scientific question; it has also been defined as research that uses data sets so large or complex that they are hard to process and analyze using traditional approaches and methods (Hey et al. 2009, Critchlow and van Dam 2013).

Although the term *data intensive* is relatively new, historians of science point out that scientists have been capturing, curating, and analyzing large volumes of data for centuries in ways that have challenged existing techniques (Muller-Wille and Charmantier 2012). For example, the disciplines of natural history and taxonomy provide important historical examples of data-intensive research; as Strasser (2012) put it, "Renaissance naturalists were no less inundated with new information than our contemporaries" (p. 85). However, contemporary data-intensive science is also characterized by new computational methods and technologies for creating, storing, processing, and analyzing data and also by the use of interdisciplinary teams for designing and implementing research to address complex societal challenges (Strasser 2012, Leonelli 2014). Consequently, in some areas of science (e.g., astronomy), there can be particularly sharp distinctions between historical and current data-intensive approaches, whereas in other areas of science (e.g., natural history), there are fewer differences (Evans and Rzhetsky 2010, Haufe et al. 2010, Pietsch 2016).

Contemporary examples of data-intensive science include collecting evidence for the existence of the Higgs boson, sequencing the human genome, developing computer models of climate change and carbon sequestration, and identifying relationships between social networks and human behaviors. Despite these high-profile examples and the increasing availability of large data sets for many science disciplines, there are concerns that contemporary data-intensive research is bad for science or that it will lead to poor methodology and unsubstantiated inferences. For example, data-intensive research has been criticized for being atheoretical, being nothing more than a "fishing expedition," having a high probability of leading to nonsense results or spurious correlations, being reliant on scientists who do not have adequate expertise in data analysis, and yielding data biased by the mode of collection (Boyd and Crawford 2012, Fan et al. 2014, Lazer et al. 2014).

Such concerns actually reflect deeper and more widespread debates about the centrality of hypothesis-driven research that have challenged the scientific community for centuries. Most contemporary scientific disciplines share a commitment to a hypothesis-driven methodology (see

BioScience XX: 1–10.
doi:10.1093/biosci/biw115 Advance Access publication XX XXXX, XXXX

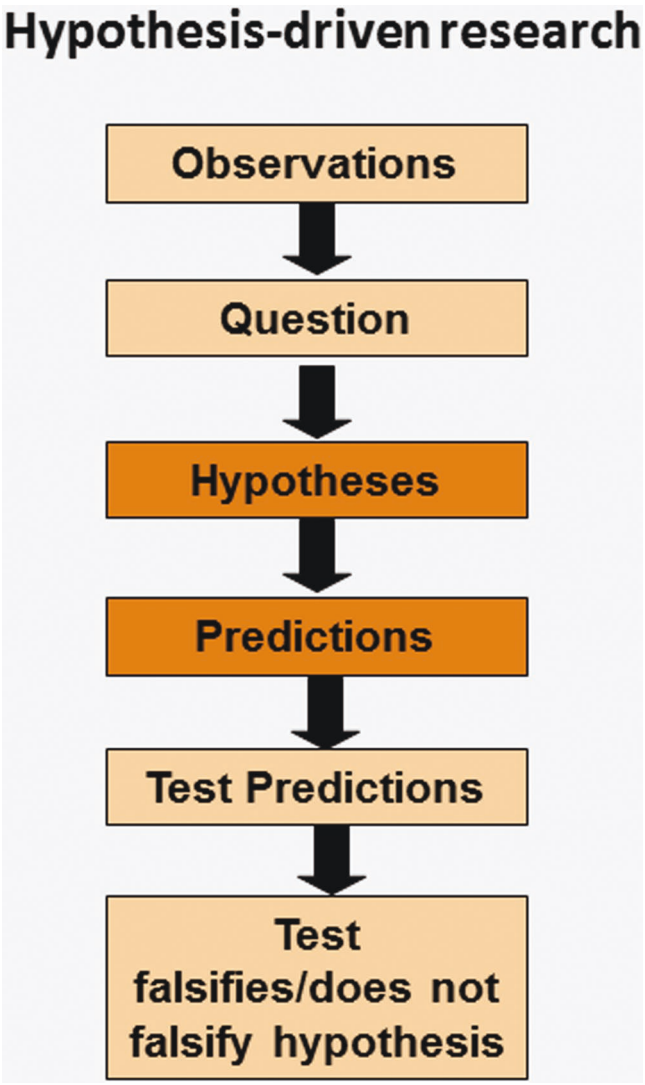

Figure 1. Linear account employed in many descriptions of the scientific method.

Peters R 1991, Weinberg 2010, Keating and Cambrosio 2012, Fudge 2014). Definitions for hypotheses vary across disciplines (ranging from specific to general and quantitative to qualitative; Donovan et al. 2015), but we define hypothesis-driven methodology in terms of the linear process canonized in many textbooks and represented in figure 1.

Although this linear scientific process continues to be held up as an exemplar in many textbooks and grant proposal guidelines (Harwood 2004, O'Malley et al. 2009, Haufe 2013), recent commentaries from scientists and historians and philosophers of science have argued that historical and contemporary scientific practices incorporate a much more complex, iterative mixture of different methods (e.g., Kell and Oliver 2004, Glass and Hall 2008, Gannon 2009, O'Malley et al. 2010, Forber 2011, Elliott 2012, Glass 2014, Peters DPC et al. 2014, Pietsch 2016). These scholars argue that focusing primarily on a linear, hypothesis-driven account of science impoverishes the scientific enterprise by encouraging scientists to focus on narrowly defined questions that can be posed as testable hypotheses. For example, hypothesis-driven approaches are particularly helpful for choosing between alternative mechanisms that could explain an observed phenomenon (e.g., through a controlled experiment), but they are much less helpful for mapping out new areas of inquiry (e.g., the sequence of the human genome), identifying important relationships among many different variables, or studying complex systems. According to those who accept an iterative account of scientific methods, attempting to draw a sharp distinction between hypothesis-driven and data-intensive science is misleading; these modes of research are not in fact orthogonal and often intertwine in actual scientific practice (e.g., O'Malley et al. 2009, Elliott 2012, Peters DPC et al. 2014).

Unfortunately, the historical and philosophical literature on iterative scientific methods has not been well integrated into recent accounts of data-intensive research, nor have the implications for evaluating research quality been fully explored. We address both of these gaps by showing how data-intensive research can be conceptualized more effectively using iterative accounts of scientific methods and by showing how these accounts encourage innovative approaches for evaluation. We argue that the key to assessing the appropriateness of data-intensive research—and, indeed, any scientific practice—is to evaluate how it is situated within broader research practices. Scientific practices should be evaluated on the basis of the significance of the knowledge gap that they address and the alignment between the nature of the gap and the approach or combination of approaches used to address it. In order to better reflect scientific practices and to accommodate all scientific approaches, including data-intensive ones, we point out recent changes and propose additional reforms in the spheres of funding, publishing, and education.

Debates over scientific methods

Contemporary debates over data-intensive methods are merely the latest episode in a long-standing conflict over the proper roles of hypotheses in scientific research. In the seventeenth century, figures such as Robert Boyle and Robert Hooke espoused the use of hypotheses, whereas Francis Bacon and Isaac Newton argued that investigators could easily be led astray if they proposed bold conjectures rather than working inductively from the available evidence (Laudan 1981, Glass 2014). These examples illustrate the long history during which hypothesis-driven science has waxed and waned in popularity (figure 2; Laudan 1981). Most scientists did not favor the use of hypotheses during the eighteenth century, but this perspective changed dramatically over the next 100 years (Laudan 1981). By the late nineteenth century, largely descriptive disciplines such as natural history were beginning to be dismissed as a form of "stamp collecting" (Johnson 2007). Popper's (1963) emphasis on the hypothetico-deductive (H-D) method proved hugely influential during the twentieth century, and

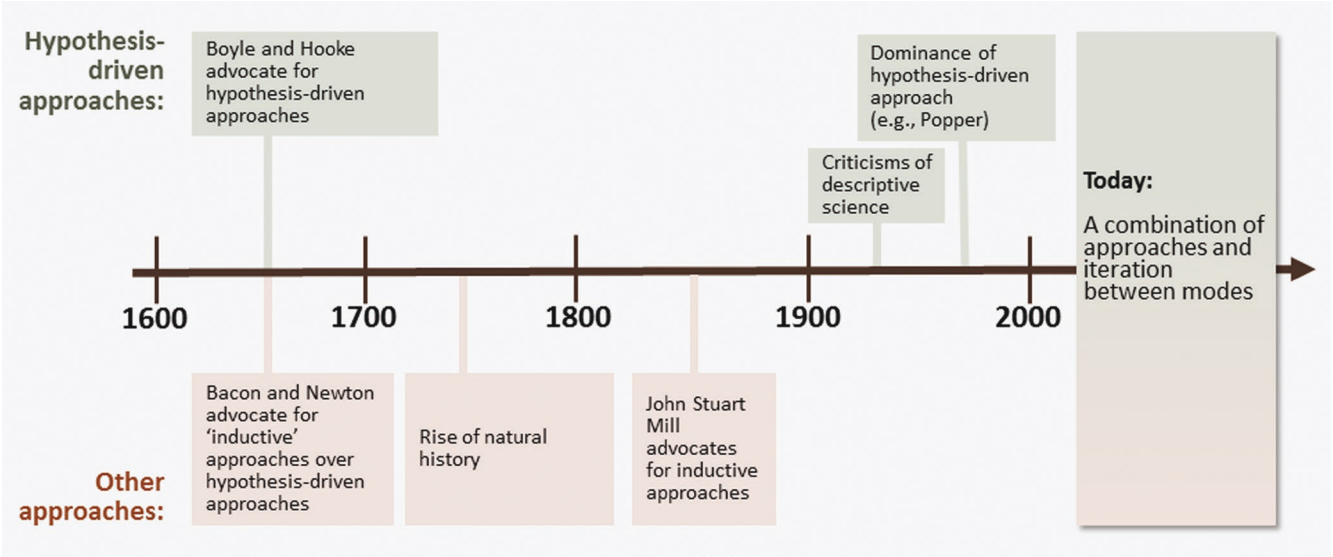

Figure 2. A depiction of the waxing and waning of hypothesis-driven approaches.

most textbooks continue to focus on hypothesis testing as the core of the scientific method (see figure 1; Harwood 2004). Although some scientists, publishers, and funders have remained loyal to a Popper-informed account of the scientific method that privileges hypothesis-driven research, many today are questioning this focus and mirroring the methodological debates embodied in previous time periods (Hilborn and Mangel 1997, Kell and Oliver 2004, Glass and Hall 2008, Peters DPC et al. 2014).

In particular, despite the huge potential for new data-intensive methodologies to generate knowledge (King 2011), the advent of these techniques has raised questions about the appropriate relationships between hypothesis-driven and observationally driven modes of investigation (Kell and Oliver 2004, Beard and Kushmerick 2009). Again, historians of science have shown that this debate is not a new one and that scientists have struggled for centuries with storing, analyzing, and standardizing large quantities of data (Muller-Wille and Charmantier 2012). Nevertheless, contemporary data-intensive science raises additional issues because of its extensive use of statistical and computer science methodologies and interdisciplinary teams (Strasser 2012), thereby adding further dimensions to debates about appropriate scientific methods.

A richer account of scientific practice

Many concerns about data-intensive research can be addressed by defining scientific practice more broadly (figure 3), as has been argued in recent historical and philosophical studies of scientific methods. Taking this view, the fundamental goal of science is to address gaps or challenges facing our current state of knowledge. Hypothesis testing is one approach for filling these knowledge gaps, but science proceeds in other ways as well (Chang 2004, Franklin 2005, O'Malley et al. 2009, Elliott 2012, O'Malley and Soyer 2012). Scientists attempt to answer research questions with observations, field studies, or integrated databases (Leonelli 2014); they engage in exploratory inquiry or modeling exercises to detect patterns in available data (Steinle 1997, Burian 2007, Elliott 2007, Winsberg 2010, Katzav and Parker 2015); or they create new tools, techniques, and methods (Baird 2004, O'Malley et al. 2010)—all of which in turn enable them to test hypotheses, answer questions, or gather additional data more effectively.

This multiplicity of different research approaches is not new, but it has become even more prominent in contemporary data-intensive research. Historically, it was often most efficient for scientists to work from hypotheses that guided their inquiry in the most promising directions. But with the advent of high-throughput technologies and data-mining techniques that make data less expensive to generate and analyze, other approaches that are more inductive also play a fruitful role in scientific research (Franklin 2005, Servick 2015). Broad hypotheses or background assumptions may still provide guidance about what sorts of questions or exploratory inquiries are likely to be most fruitful, but these are not the sorts of specific hypotheses envisioned by most hypothesis-driven accounts of scientific method (Franklin 2005, Leonelli 2012, Ratti 2015). Because it is difficult (often impossible) for an individual scientist to become an expert in all of these contemporary approaches and methods, good science also incorporates the most appropriate disciplines and collaborators, thus making the development of effective—and often interdisciplinary—scientific teams more essential than in the past, and the resulting research reflects a combination of methods originating from multiple disciplines (Cheruvelil et al. 2014, NRC 2015).

An important feature of the scientific methods illustrated in figure 3 is that they are often employed in an iterative fashion in order to address complex research challenges

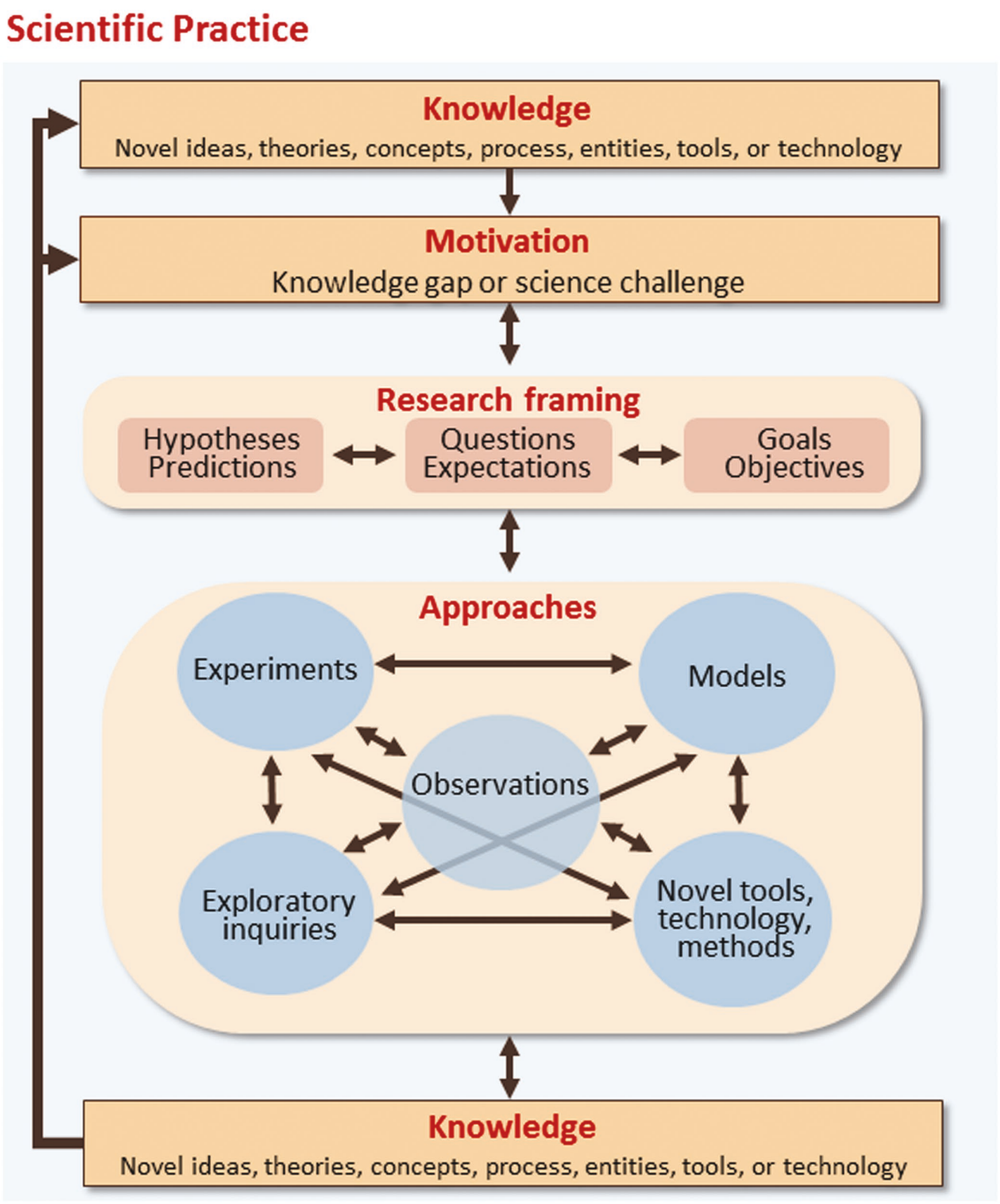

Figure 3. A representation of scientific practice as an iterative process, with many approaches and links (as depicted by two-way arrows). The evaluation or assessment of scientific practices is based on the importance of the knowledge generated, the importance of the gap or challenge addressed, and the alignment of the approaches and methods used to conduct the science.

(Chang 2004, O'Malley et al. 2010, Elliott 2012, Leonelli 2012). Although some contemporary data-intensive research focuses primarily on the repeated use of inductive methods and machine-learning algorithms (Evans and Rzhetsky 2010, Lazer et al. 2014, Pietsch 2016), much of it involves a combination of different approaches. O'Malley and colleagues (2010) argued that not only data-intensive research but also scientific practice as a whole should be characterized as an iterative interplay between at least four different modes of research: hypothesis-driven, question-driven, exploratory, and tool- and method-oriented. As inquiry proceeds, initial questions are specified, whereas others are revised or give rise to new lines of research. In an effort to address these questions, new equipment and techniques are often developed and tested, frequently generating new questions and altering old ones. In the course of investigating questions and developing new techniques, exploratory approaches are often central (O'Malley et al. 2010). These exploratory efforts, which can include experimentation, data mining, and simulation modeling, often involve the systematic variation of experimental parameters or analysis of datasets in search of important regularities and patterns (Elliott 2007, Winsberg 2010). In many cases, this web of activities generates the sorts of tightly constrained contexts in which specific hypotheses can be fruitfully tested, but this may be just one component of a much broader scientific context. In fact, the *methodological* iteration between different approaches results in a process of *epistemic* iteration by which our knowledge is gradually altered and improved (Elliott 2012), as is depicted by the two-way arrows in figure 3 that highlight the links among knowledge, motivation, and the multiple approaches employed by scientists.

One of the primary lessons to be learned from the iterative model of scientific methods is that contemporary research, and especially data-intensive research, incorporates a wide variety of different approaches, which gain their significance primarily from their roles in broader research programs and lines of inquiry. Therefore, evaluating the quality of this work requires much more than looking to confirm that it incorporates a well-formulated hypothesis (Kell and Oliver 2004, Beard and Kushmerick 2009). Instead, it should be evaluated on the basis of the alignment between the nature of the knowledge gap or challenge addressed and the combination of approaches or methods used to address the gap. Research should be evaluated favorably if it incorporates approaches and methods that are well-suited for addressing an important gap in current knowledge, even if they do not focus solely or primarily on hypothesis testing (figure 3).

An iterative model of scientific practice alleviates many common concerns about data-intensive research. The potential for generating spurious correlations becomes less serious when data-generated patterns are identified and evaluated as part of larger research projects that incorporate broader research questions, hypotheses, or objectives and when appropriate techniques and inferences are used to deal with spurious correlations (Hand 1998). These projects are also frequently embedded within conceptual frameworks or theories that facilitate the investigation of underlying causal mechanisms. Some proponents of data-intensive science argue that it can largely replace hypothesis testing, focusing on generating correlations rather than seeking causal understanding (Prensky 2009, Steadman 2013). In contrast, we contend that data-intensive science will typically be most fruitful when it is part of broader inquiries that guide the collection and interpretation of data and that provide additional investigations of the correlations that are generated (Leonelli 2012, Kitchin 2014). Finally, the worry that individual researchers do not have the skill sets to perform data-intensive work can be alleviated by the development of interdisciplinary research teams that can accomplish the iterative tasks required for many contemporary scientific research projects. Admittedly, data-intensive methods can still be used inappropriately, such as when data are collected without standard approaches or quality metadata or when data are simply mined for correlative relationships without attention to spurious correlations (Hand 1998). However, we argue that this is a matter of improper technique or a poorly designed research program, which can occur in any form of scientific practice; it is not a problem inherent in data-intensive methods themselves.

Examples of iterative data-intensive research practices

The interplay between multiple research approaches can be observed across many scientific subdisciplines and time periods. To illustrate, we present two examples drawn from the natural sciences. The first example highlights the historical nature of these debates concerning scientific methods (the study of evolution by natural selection; figure 4a). It shows that even though contemporary data-intensive approaches have unique characteristics, historical research also incorporated iterative and data-intensive components. The second example highlights how methods from contemporary data-intensive ecology are being used to better understand broad-scale ecological research questions and environmental problems (the study of macrosystems ecology; figure 4b). It also illustrates how contemporary data-intensive research incorporates greater use of computational approaches and interdisciplinary teams than did historical data-intensive research.

The historical study of evolution by natural selection. Darwin's development of the theory of natural selection provides a classic example of research that incorporates multiple approaches. Despite the efforts of some commentators to reconstruct Darwin's research as primarily hypothesis-driven (Ayala 2009), he spent more than two decades performing exploratory work in an effort to identify the patterns that he later explained in *The Origin of Species.* Driven by curiosity and a naturalist's love for nature, as well as a structured observational agenda that he learned from scholars like Humboldt, Cuvier, and Lyell, Darwin's observations during

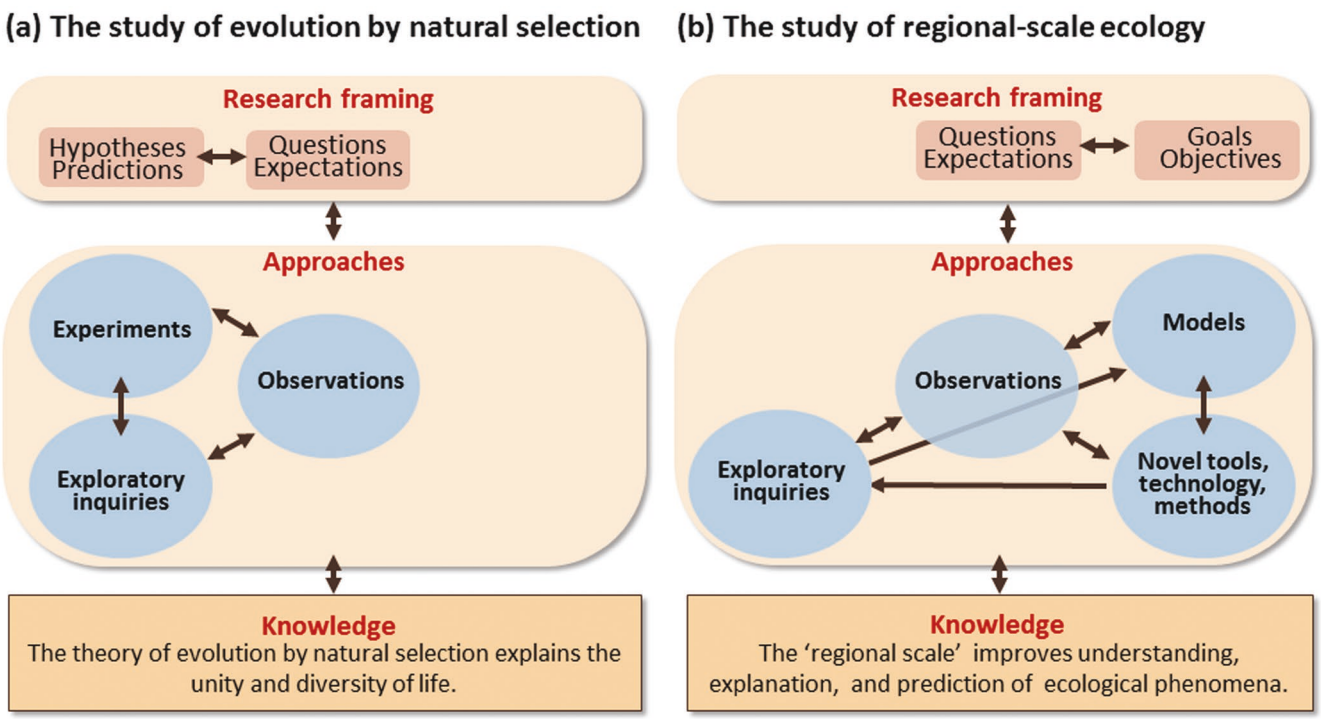

Figure 4. Two examples of iterative scientific efforts using multiple approaches.

his famous voyage aboard the *Beagle* generated questions that guided his inductive data collection over subsequent decades. During that time, he drew upon a wide range of methods and sources (Hodge 1983), including data produced by fellow members of the traditional scientific elite and countless women and other so-called amateurs practicing science outside of the scientific societies and journals of the nineteenth century. In the *Origin*, for instance, Darwin cites animal breeders as an important source of data, and in *Expression of Emotions*, mothers provided observations of their own children to supplement those made by Darwin of his own family (Harvey 2009, Montgomery 2012).

Darwin's use of natural history methods led Frank Gannon to write a tongue-in-cheek editorial pointing out that in today's funding structure Darwin's work would be dismissed as "an open-ended 'fishing expedition'" (Gannon 2009). However, Darwin also engaged in experiments that showed how his theory of evolution could explain the details of sexual form in plant species (Bellon 2013). His combination of methods and compilation of data from a variety of sources proved to be extremely fruitful, and works such as *Origin* (1859), *The Variation of Animals and Plants under Domestication* (1868), *The Descent of Man* (1871), and *Expression of Emotions* (1872) all embody a blend of what are now often held up as distinct approaches: inductive and deductive methods, observation and experiment.

Even in Darwin's own time, he was forced to consciously navigate scientific norms when considering how to present his multi-modal research. For example, following nineteenth-century philosophers of science such as William Whewell and John F. W. Herschel, Darwin organized the *Origin* to conform to the scientific values of the day—namely, demonstrating the strength of a theory by the breadth of facts it explained (Ruse 1975). Arguing from analogy, as Whewell recommended, Darwin began by recognizing an uncontested phenomenon—that artificial selection quickly resulted in drastic structural changes in domestic breeding of animals such as pigeons—and used this accepted truth to compel the reader to accept his inference that natural selection accounted for species changes.

Darwin's use of both inductive and deductive methods also followed Whewell's methodological recommendations. In contrast with more recent accounts of hypothesis-driven science, Whewell insisted that scientists should move through a very gradual inductive process to arrive at successively more general causal laws (Snyder 1999). Only after performing this inductive process did he think that scientists could legitimately move on to test these hypotheses. Thus, Whewell himself encouraged the use of a combination of research modes, and this is reflected in Darwin's works. Philosophers of science have since debated the extent to which Darwin was influenced by different methodologists (including Francis Bacon and John Stuart Mill, as well as Whewell and Herschel) and precisely when Darwin switched from an inductive to a deductive approach during the 20-plus years of gestation of the *Origin* (Ruse 1975, Hodge 1991). Regardless of the exact year when this switch occurred, it is clear that scientists today—like Darwin—often move back and forth between the best aspects of both inductive and deductive logic when formulating and testing a theory. Similarly—and again like

Darwin—scientists also often blend laboratory- and fieldwork, observation and experiment, and data from multiple sources rather than conforming to artificially distinct modes of scientific practice that are sometimes held up as "traditional" to a particular field of science, despite the long history of a multimodal reality.

The contemporary study of macrosystems ecology. A contemporary example of data-intensive research that involves multiple and iterative approaches comes from the emerging subdiscipline of macrosystems ecology (Heffernan et al. 2014). Most traditional ecological research is conducted by studying organisms and their environments at relatively small scales—such as individual species, communities, or ecosystems—using methods such as lab or field experiments, modeling, field surveys, or long-term studies (Carmel et al. 2013). However, environmental changes such as the spread of invasive species, climate change, and land-use intensification are occurring globally, are the result of relationships and interactions between human and natural systems, and may result in widespread but complicated effects. For example, across regions and continents (at the scales of hundreds of kilometers), there are differences in the direction and magnitude of environmental changes, the underlying geophysical and ecological contexts, and social structures. These differences mean that results from fine-scaled studies in some regions are not likely to apply to other regions and that the study of ecological systems at larger scales—such as regions to continents—is required. Macrosystems ecology fills this gap by explicitly studying fine-scaled ecological patterns and processes nested within regions and continents and employing a variety of methods to do so.

Such multiscaled understanding of ecological systems cannot be achieved through an individual hypothesis test or a field experiment, nor can it be achieved by using only one approach (Heffernan et al. 2014, Levy et al. 2014). For example, to understand the complex relationships among tree growth, human disturbance, and regional and global climate, scientists need to study forests as a whole using multiple methods within a region rather than at the scale of individual trees or stands (Chapin et al. 2008). One approach that ecologists have used to study ecological systems at regional scales is by quantitatively delineating ecological regions that represent a measured combination of geophysical features thought to influence fine-scaled ecological processes (Cheruvelil et al. 2013). However, existing ecological regions have limitations in that they were created for a variety of purposes, using different underlying geophysical and human data and using a variety of methods.

For example, lake water quality is related to both climate and land use. Therefore, scientists have speculated that lake water quality is likely to strongly respond to changes in both climate and land uses. However, the response of lake water quality to such environmental changes is likely to vary among regions and continents. In fact, Cheruvelil and colleagues (2008, 2013) had observed that lake water chemistry varied regionally but that the variation depended on how the boundaries of "regions" were defined. Therefore, they had the overarching goals of developing new ways to define regional boundaries that were based on the geophysical features that are likely important for predicting regional water quality and its response to climate and land-use change (figure 4b). Meeting these goals required the iterative use of multiple research methodologies, data collected by various individuals and groups, and contributions from multiple disciplines.

An interdisciplinary team was created (*sensu* Cheruvelil et al. 2014) that included ecologists, computer scientists, and experts in geospatial analysis and ecoinformatics to create a large, multiscaled database by integrating multiple lake data sources (including field surveys of water quality conducted by state agency scientists, citizen scientists, and university researchers) with geospatial data quantified at the national scale (Soranno et al. 2015). The team used three data-intensive approaches to meet their goal of developing new ecological regions for water quality (figure 4b): First, they developed and tested a clustering algorithm to define regional boundaries (Yuan et al. 2015); second, they used an exploratory data-mining analysis to determine which geophysical features were correlated with the regional boundaries and might lend insight into the underlying mechanisms driving regional variation in lake water quality (Cheruvelil, Lyman Briggs College and Department of Fisheries and Wildlife, Michigan State University, East Lansing, personal communication, 9 November 2015); and third, they used statistical models to quantify how well the regional boundaries captured variation in lake water quality for thousands of lakes in approximately 100 regions (Cheruvelil, Lyman Briggs College and Department of Fisheries and Wildlife, Michigan State University, East Lansing, personal communication, 9 November 2015). Ecological regions were created with a variety of geophysical features that are related to lake water quality, many of which are expected to be strongly affected by changes in climate and land use. Employing multiple scientific practices, rather than solely a hypothesis-driven approach, improved their ability to use the regional scale for understanding, explaining, and predicting ecological phenomena across spatial scales.

Lessons learned from examples of iterative data-intensive research. Together, these two examples illustrate the major points that we have made in this article. First, they show that although scientists have been working with challenging quantities of data for centuries, contemporary data-intensive science incorporates additional features. For example, whereas Darwin received data from numerous sources, he worked primarily on his own (with input from colleagues) to analyze the data. In contrast, the environmental scientists in the second example worked with computer scientists and experts in ecoinformatics in order to make optimal use of contemporary computational tools for integrating, creating, and analyzing data.

Second, these examples illustrate the power of moving iteratively among multiple research methods. What made both

Table 1. Recommendations for promoting iterative data-intensive science.

Components of science	Current norms	Proposed reforms	Recent exemplar of reform
Funding	Proposals are expected to have an organizing hypothesis.	Proposals should be expected to have alignment between knowledge gaps and approaches.	Several institutes of the NIH have introduced long-term funding opportunities that allow investigators to pursue more creative, innovative research projects (e.g., *http://grants.nih.gov/grants/guide/rfa-files/RFA-DE-17-002.html* and *http://grants.nih.gov/grants/guide/rfa-files/RFA-NS-16-001.html*)
	Proposals are expected to describe a linear, non-iterative approach.	Proposals should be expected to describe appropriate iterative use of multiple approaches.	The Biotechnology and Biological Sciences Research Council of the UK describes multiple methods that are integrated into the systems-biology research it funds (*http://bbsrc.ac.uk/research/systems-approach*).
Publishing	Articles are expected to be structured to embody a hypothesis-testing approach.	Articles should be structured to convey the alignment between the identified knowledge gaps and the approaches used.	A new journal, *Limnology and Oceanography Letters*, requires an explicit statement by the authors of the knowledge gaps filled by the study (*www.LOLetters.com*).
	The components of iterative research are difficult to publish on their own (e.g., exploratory analysis, data, methods, code).	Articles focused on any aspect of iterative research should be publishable based on contribution to knowledge, data, or methods development	Recent advent of outlets for a broad range of research products, such as data journals (e.g., *Earth System Science Data, Scientific Data, GigaScience, Biodiversity Data Journal*), online code repositories (e.g., *GitHub, BitBucket*), and online data repositories (e.g., *FigShare, Dryad, TreeBASE*)
Education (K–12, undergraduate, and graduate)	Students are taught mainly about hypothesis testing.	Students should be taught multiple scientific methods and to choose approaches that best align with knowledge gaps.	Reformed teaching approaches, such as authentic science labs (e.g., Luckie et al. 2004, Harwood 2004) and teaching with case studies (e.g., *http://sciencecases.lib.buffalo.edu/cs/collection, http://www.evo-ed.org,* White et al. 2013).
	Students are taught linear, non-iterative scientific methods.	Students should be taught an iterative account of scientific methods.	Dissemination of nonlinear accounts of scientific methods (e.g., *http://undsci.berkeley.edu/article/howscienceworks_02*)

of these research efforts successful is not the fact that they used a particular approach but rather that the approaches they chose were well designed for addressing important knowledge gaps. In Darwin's case, his research was important because he was addressing one of the most fundamental issues in biology—namely, the processes by which species have changed over time. Similarly, the scientists in our second example were addressing the important societal issue of the response of water quality to environmental changes at macroscales. Encouraging scientists to emulate the iterative approaches embodied in these two examples requires the development of richer conceptions of scientific practice.

Recommendations for promoting good science in our data-rich world

A number of reforms should be made to promote not only iterative data-intensive science but also the scientific enterprise more broadly (table 1). First, funding agencies (and reviewers) should evaluate the quality of proposed research not based on a uniform requirement that it states a specific hypothesis but based on the importance of the knowledge gaps that it identifies and the appropriateness of the methods proposed for addressing those gaps (O'Malley et al. 2009). For example, some recent funding initiatives are placing emphasis on grand challenges (e.g., the human genome project, brain research, personalized medicine, smart cities) that do not lend themselves to solely hypothesis-based approaches. Therefore, rather than expecting researchers to shoehorn proposals into a misleading, linear research format, reviewers should be open to proposals that describe a more realistic, iterative research trajectory. This reform will require developing appropriate grant guidelines and review mechanisms that encourage mixed modes of scientific practice, such as those recently being used by the US National Institutes of Health to fund investigators rather than individual projects (table 1).

Second, rather than expecting articles to be structured to embody a linear hypothesis-testing approach, journal editors and reviewers should be open to publications that are organized around the full range of methods used to address knowledge gaps. Allowing journal articles and other research products to take a greater variety of forms will help alleviate the discrepancies that a number of authors have identified between the structure of scientific articles and the actual practice of research (e.g., Medawar 1996, Schickore 2008). Some journals and online repositories are providing guidelines and mechanisms for scientists to disseminate data and computer code, and the science community as a whole is discussing ways to give scientists credit for a variety of research products that will help advance a broader view of scientific practices (e.g., Goring et al. 2014; see also table 1).

Third, whereas K–12 through graduate science education currently emphasizes a linear, hypothesis-driven approach to science, it should be reformed to incorporate more complex models of the scientific method. For example, students should be taught that hypothesis testing is just one important component of a much broader landscape of scientific activities that need to be combined in creative and interdisciplinary ways to move science forward (Harwood 2004). Including the history, philosophy, and sociology of science in science curricula; teaching science in interdisciplinary ways; and using reformed teaching methods in science courses (e.g., inquiry-based labs, case studies) can introduce students to the multiple methods scientists have historically used—and continue to use—to address significant knowledge gaps (table 1).

Conclusions

The recognition that data-intensive research methods—and indeed, research practices in all areas of science—need to be evaluated as part of broader research programs does much to alleviate common concerns about these and other non-hypothesis-driven methods. Although data-intensive and exploratory efforts to identify patterns in large datasets have the potential to generate spurious results, all methods have their potential problems when used poorly; when used properly, such data-intensive approaches can play a very fruitful role in broader research programs that also test hypothesized processes and mechanisms. The iterative research methods that we have described in this article allow researchers to address more complex questions than they could with hypothesis testing alone. To make these efforts successful, changes are needed in the norms for research funding, publication, and education. In all these areas, more emphasis should be placed on aligning research methods with the knowledge gaps that need to be addressed rather than focusing primarily on hypothesis testing. In addition, scientific practice should be more explicitly recognized as an iterative path through multiple approaches rather than as a linear process of moving through pre-defined steps. Of course, this does not mean that "anything goes"; rather, it facilitates more careful thought about how to fund, publish, and teach the right combinations of methods that will enable the scientific community to tackle the big issues confronting society today.

Acknowledgments

Funding for this work was provided by the Science + Society @ State program at Michigan State University to all authors; the US National Science Foundation's Macrosystems Biology Program (no. EF-1065786) to PAS and KSC; and the USDA National Institute of Food and Agriculture, Hatch Project no. 176820 to PAS.

References cited

Baird D. 2004. Thing Knowledge: A Philosophy of Scientific Instruments. University of California Press.

Baraniuk RG. 2011. More is less: Signal processing and the data deluge. Science 331: 717–719.

Beard D, Kushmerick M. 2009. Strong inference for systems biology. PLOS Computational Biology 5 (art. e1000459).

Bell G, Hey T, Szalay A. 2009. Beyond the data deluge. Science 323: 1297–1298.

Bellon R. 2013. Darwin's evolutionary botany. Pages 131–138 in Ruse M, ed. The Cambridge Encyclopedia of Darwin and Evolutionary Thought. Cambridge University Press.

Boyd D, Crawford K. 2012. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication, and Society 15: 662–679.

Burian R. 2007. On microRNA and need for exploratory experimentation in post-genomic molecular biology. History and Philosophy of the Life Sciences 29: 283–310.

Chang H. 2004. Inventing Temperature: Measurement and Scientific Progress. Oxford University Press.

Donovan S, O'Rourke M, Looney C. 2015. Your hypothesis or mine? Terminological and conceptual variation across disciplines. Sage Open 5: 1–13. doi: 10.1177/2158244015586237

Elliott K. 2007. Varieties of exploratory experimentation in nanotoxicology. History and Philosophy of the Life Sciences 29: 311–334.

———. 2012. Epistemic and methodological iteration in scientific research. Studies in History and Philosophy of Science 43: 376–382.

Evans J, Rzhetsky A. 2010. Machine science. Science 329: 399–400.

Fan J, Han F, Liu H. Challenges of big data analysis. Natural Science Review 1: 293–314.

Forber P. 2011. Reconceiving eliminative inference. Philosophy of Science 78: 185–208.

Franklin L. Exploratory experiments. Philosophy of Science 72: 888–899.

Fudge D. 2014. Fifty years of J. R. Platt's strong inference. Journal of Experimental Science 217: 1202–1204.

Gannon F. 2009. A letter to Darwin. EMBO Reports 10: 1.

Glass DJ. 2014. NIH grants: Focus on questions, not hypotheses. Nature 507: 306.

Glass DJ, Hall N. 2008. A brief history of the hypothesis. Cell 134: 378–381.

Goring S, et al. 2014. Improving the culture of interdisciplinary collaboration in ecology by expanding the measures of success. Frontiers in Ecology and the Environment 12: 39–47.

Hand DJ. 1998. Data mining: Statistics and more? American Statistician 52: 112–118.

Harwood W. 2004. A new model for inquiry: Is the scientific method dead? Journal of College Science Teaching 33: 29–33.

Haufe C. 2013. Why do funding agencies favor hypothesis testing? Studies in History and Philosophy of Science 44: 363–374.

Haufe C, Burian R, Elliott K, O'Malley M. 2010. Machine science: What's missing. Science 330: 317–318.

Hey T, Tansley S, Tolle K. 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research.

Hilborn R, Mangell M. 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press.

Hodge M. 1991. Darwin, Whewell, and natural selection. Biology and Philosophy 6: 457–460.

Johnson K. 2007. Natural history as stamp collecting: A brief history. Archives of Natural History 34: 244–258.

Katzav J, Parker S. 2015. The future of climate modeling. Climate Change 132: 475–487.

Kitchin R. 2014. The Data Revolution: Big Data, Open Data, Data Infrastructures, and Their Consequences. Sage.

Keating P, Cambrosio K. 2012. Too many numbers: Microarrays in clinical cancer research. Studies in History and Philosophy of Biology and Biomedical Sciences 43: 37–51.

Kell D, Oliver S. 2004. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays 26: 99–105.

Khoury M, Ioannidis J. 2014. Big data meets public health. Science 346: 1054–1055.

King G. 2011. Ensuring the data-rich future of the social sciences. Science 331: 719–721.

Laudan L. 1981. Science and Hypothesis: Historical Essays on Scientific Methodology. Reidel.

Lazer D, Kennedy R, King G, Vespignani A. 2014. The parable of Google flu: Traps in big data analysis. Science 343: 1203–1205.

Leonelli S. 2012. Introduction: Making sense of data-driven research in the biological and biomedical sciences. Studies in History and Philosophy of Biology and Biomedical Sciences 43: 1–3.

———. 2014. What difference does quantity make? On the epistemology of Big Data in biology. Big Data and Society 1: 1–11. doi:10.1177/2053951714534395.

Leonelli S, Ankeny R. 2012. Re-thinking organisms: The impact of databases on model organism biology. Studies in History and Philosophy of Biology and Biomedical Sciences 43: 29–36.

Levy O, et al. 2014. Approaches to advance scientific understanding of macrosystems ecology. Frontiers in Ecology and the Environment 12: 15–23.

Luckie DB, Maleszewski JJ, Loznak SD, Krha M. 2004. Infusion of collaborative inquiry throughout a biology curriculum increases student learning: A four-year study of "Teams and Streams." Advances in Physiology Education 28: 199–209.

Mattman CA. 2013. A vision for data science. Nature 493: 473–475.

Medawar P. 1996. Is the scientific paper a fraud? Pages 33–39 in Medawar P, ed. The Strange Case of the Spotted Mice and Other Classic Essays on Science. Oxford: Oxford University Press.

Montgomery GM. 2012. Gender and evolution. Pages 443–450 in Ruse M, ed. Cambridge Encyclopedia of Darwin and Evolutionary Thought. Cambridge University Press.

Muller-Wille S, Charmantier I. 2012. Natural history and information overload: The case of Linnaeus. Studies in History and Philosophy of Biology and Biomedical Sciences 43: 4–15.

[NRC] National Research Council. 2015. Enhancing the Effectiveness of Team Science. National Academies Press.

O'Malley M, Soyer O. 2012. The roles of integration in molecular systems biology. Studies in History and Philosophy of Biology and Biomedical Sciences 43: 58–68.

O'Malley M, Elliott KC, Haufe, C, Burian R. 2009. Philosophies of funding. Cell 138: 611–615.

O'Malley M, Elliott KC, Burian R. 2010. From genetic to genomic regulation: Iterative methods in miRNA research. Studies in History and Philosophy of Biology and Biomedical Sciences 41: 407–417.

Peters R. 1991. A Critique for Ecology. Cambridge University Press.

Peters DPC, Havstad KM, Cushing J, Tweedie C, Fuentes O, Villanueva-Rosales N. 2014. Harnessing the power of big data: Infusing the scientific method with machine learning to transform ecology. Ecosphere 5 (art. 67).

Pietsch W. 2016. The causal nature of modeling with big data. Philosophy and Technology 29: 137–171.

Popper K. 1963. Conjectures and Refutations. Routledge and Kegan Paul.

Porter JH, Hanson PC, Lin C-C. 2009. Staying afloat in the sensor data deluge. Trends in Ecology and Evolution 27: 121–129.

Prensky MH. 2009. Sapiens digital: From digital immigrants and digital natives to digital wisdom. Innovate 5 (art. 1).

Ratti E. 2015. Big data biology: Between eliminative inferences and exploratory experiments. Philosophy of Science 82: 198–218.

Ruse M. 1975. Examination of the influence of the philosophical ideas of John F. W. Herschel and William Whewell in the development of Charles Darwin's theory of evolution. Studies in History and Philosophy of Science 6: 159–181.

Schickore J. 2008. Doing science, writing science. Philosophy of Science 75: 323–343.

Servick K. 2015. Proposed study would closely track 10,000 New Yorkers. Science 350: 493–494.

Snyder L. 1999. Renovating the *Novum Organum*: Bacon, Whewell, and induction. Studies in History and Philosophy of Science 30: 531–557.

Soranno PA, Cheruvelil KS, Elliott KC, Montgomery GM. 2015a. It's good to share: Why environmental scientists' ethics are out of date. BioScience 65: 69–73.

Soranno PA, et al. 2015b. Building a multi-scaled geospatial temporal ecology database from disparate data sources: Fostering open science and data reuse. GigaScience 4 (art. 28). doi:10.1186/s13742-015-0067-4.

Steadman I. 2013. Big data and the death of the theorist. Wired. (22 August 2016; *www.wired.co.uk/news/archive/2013-01/25/big-data-end-of-theory*)

Steinle F. 1997. Entering new fields: Exploratory uses of experimentation. Philosophy of Science 64: S65–S74.

Weinberg R. 2010. Point: Hypotheses first. Nature 464: 678.

White PJT, Heidemann MK, Smith JJ. 2013. A new integrative approach to evolution education. *BioScience* 63: 586–594.

Winsberg E. 2010. Science in the Age of Computer Simulations. University of Chicago Press.

Kevin C. Elliott (kce@msu.edu) is an associate professor in Lyman Briggs College, the Department of Fisheries and Wildlife, and the Department of Philosophy; Kendra S. Cheruvelil is an associate professor in Lyman Briggs College and the Department of Fisheries and Wildlife; Georgina M. Montgomery is an associate professor in Lyman Briggs College and the Department of History; and Patricia A. Soranno is a professor in the Department of Fisheries and Wildlife at Michigan State University, in East Lansing. All authors contributed equally to the conceptualization of the paper and the supporting research. KCE organized the collaboration and initiated the writing process. All authors contributed text, reviewed manuscript drafts, and approved the final version.