


CONTRIBUTED PAPER

Excavation in the Sky: Historical Inference in Astronomy

Siyu Yao 

Department of History and Philosophy of Science and Medicine, Indiana University, Bloomington, IN, USA

Email: siyuyao@iu.edu

(Received 23 January 2023; accepted 24 January 2023; first published online 16 February 2023)

Abstract

The philosophy of historical sciences investigates their distinct objects of study, epistemic challenges, and methodological solutions. Rethinking astronomy in this light offers a contribution. First, the methodology of historical sciences adds to a more adequate description of how astronomers study and utilize token events. Second, astronomy faces a typical difficulty in identifying traces of some past events and has developed a delicate solution. This enriches the idea of trace and suggests a methodology that relies on iterations between data-driven approaches and theory-driven approaches, together with the cross-validation between multiple relevant historical events or datasets.

1. Introduction

Historical sciences such as archaeology and paleobiology study objects and events in the past. Inquiries into the sky before the twentieth century appeared to have different objectives and methodologies from historical sciences. While cosmology contemplated the history of the universe as a whole, studies of its component celestial objects aimed mainly at uncovering general unchanging mathematical-physical regularities. In contrast, local knowledge and the reconstruction of token events play a central role in other historical sciences.

A closer inspection of astronomy in the last century shows more similarities with the latter. Aside from flourishing cosmological models, astrophysics and celestial mechanics have engaged more in impermanent phenomena that contingently occur and dynamically evolve, emphasizing token peculiarity beyond general laws. Both philosophers (Turner 2007; Fox 2021) and astrophysicists (Frebel 2015; Anderl 2016) have suggested epistemological and methodological analogies to archaeology: Like archaeologists, more and more astronomers “excavate” remnants of past events and infer backward.

The philosophy of historical sciences illuminates two general epistemic challenges common to their objects of study: the lack of manipulation and the reliance on sparse

traces. These challenges invigorate historical scientists to explore possible reasoning patterns and epistemological strategies that can mitigate them and make progress.

In this article, I synthesize the philosophical discussions surrounding historical sciences with recent practices of astronomy, suggesting contributions to both sides. First, it reveals some epistemological and methodological similarities between these sciences. Going beyond the metaphor of the “cosmic laboratory” that is typically invoked to describe astronomical practices, I highlight the historical reconstruction of token events in astronomy and its methodological importance. Second, certain domains of astronomy show a typical difficulty in identifying traces of some past events. This enriches our understanding of the general idea of “traces” and how they can be identified.

2. Astronomy and historical sciences

The category of historical sciences is not entirely uniform, but philosophers often stress several typical features of it. First, historical sciences make inferences about things and events in the deep past from their remnants. The deep time between the objects of study and their contemporary proxies is home to epistemic difficulties rendered by intractable decay, distortion, and confounders (Turner 2007; Currie 2018). Second, historical studies feature the type of knowledge at which they aim. Due attention is directed to the detailed *reconstruction* of *token* historical events or processes. This means answering “when” and “how” questions with details peculiar to that token, situating the event in its specific context, and fitting details coherently together (Cleland 2011; Wylie 2011). Moreover, historical studies often involve not only deep time but also *thick* time. The knowledge is not only about individual past states but also about how those states dynamically develop and influence later states over a long course of time.

The transition of observational astronomy from a physical-mathematical science describing static regularities to a science that incorporates dynamic change and token peculiarities germinated in the eighteenth century (Wilson 2003). Over the last 10 decades, this transition is further enabled by the surge in the amount, variety, and quality of observational data; the exponential growth of computational power; and the establishment and refinement of physical models of the evolution of the entire universe and its various components.

Many features of historical sciences are shared by present-day astronomy. Astronomical observations of the distant are always about the past (Turner 2007). Celestial objects also have a history that cannot be captured by the traditional taxonomy of stars based on synchronic kind-member conditions. This makes an individual star a “nomad” in taxonomy if its entire life is traced (Ruphy 2016). In addition, the historical reconstruction of token events constitutes a large portion of the aims of present-day astronomy, from the formation of the whole universe to the more contingent and specific ones, such as the assemblage of an individual galaxy.

Some subdomains of astronomy also show epistemological and methodological similarities to historical sciences. For example, knowledge about the early universe is obtained and justified through a combination of two approaches: far-field cosmology “observes” the temporal past by gazing at objects tens of billions of light years away from us, while the “near-field cosmology” *infers* the deep past from nearby stars with

the help of complex models of stellar and galactic evolution. This latter approach presents a similar epistemic situation typically encountered in historical sciences, as these near-field traces are often not straightforwardly informative about the past. Detailed historical reconstruction thus requires using various inferential strategies typically seen in historical sciences. Beyond this, even when the interest does not lie in the reconstruction of token events but in finding long-term regularities, many studies are of a historical nature. Our observation is only a “snapshot” of the cosmic processes that last for billions of years (Jacquart 2022). As it is unlikely that we could observe those processes or test a prediction in the far future, the relatively more accessible route is to observe fossils of such processes and make inferences backward.

3. Challenges and resources of historical sciences

Philosophers of historical sciences offer a general framework for the epistemology and methodology of sciences under similar epistemic situations.¹ Their analysis often starts from two stark challenges of historical sciences: the lack of manipulation and the reliance on sparse traces (Cleland 2002; Turner 2007). In experimental sciences, experimenters can replicate experiments and manipulate experimental conditions to study how the result changes with varying factors. They can thus rule out irrelevant factors and test competing hypotheses. Without the power to manipulate the past, historical sciences seem to lack a comparable variety and amount of evidence for these purposes.

Thus emerges the second challenge, the sparsity of traces with which historical scientists make inferences. Traces are material remnants that have been causally affected by certain past events or processes and, as a result, some of their properties can be a proxy for the past (Turner 2007). However, informative traces are often sparse. Trace formation is not the kind of process designed to record any information with which scientists are interested. What can be recorded depends on the natural processes that generate preservable traces. Traces also suffer from decay and distortion over time: rocks weather, fossils migrate, and dynamical patterns in galaxies dissipate, complicating their connection with past events.

Philosophers who are optimistic about what historical sciences can achieve argue that certain reasoning patterns and epistemological strategies can lead to relatively reliable historical inferences.

Justification in historical sciences often involves narrative and common cause explanations, both stressing the explanatory power of *token historical events*. Narrative explanation constructs a coherent causal story culminating in the phenomena being explained. It has a central subject, and the way stages of the story follow each other could be flexible, without strictly appealing to regularities. Championed by Cleland (2002, 2011), common cause explanation suggests that historical scientists are justified to favor a hypothesis that assumes a common cause for seemingly improbable coincidences. As a result, observation overdetermines hypotheses. Finding just a few traces may suffice as a “smoking gun” that tells apart competing

¹ Like Fox (2021), I believe the category of historical sciences does not show exclusive epistemology and methodology, but the study of their typical features can teach lessons to sciences sharing certain similarities.

hypotheses. Thereby, she suggests that the alleged epistemic difficulties rendered by the lack of manipulation are balanced by overdetermination in inferences.

The role of regularities has also been stressed. Jeffares (2008) argues that one has to establish “midrange theories” that reliably link contemporary observation with its implications for the past. An example of this is radiocarbon dating, a ubiquitous method that reliably infers the time elapsed from material debris. As midrange theories express stable regularities, they can be tested with experiments. Nevertheless, studies of token events are still substantial. They contribute to the development of midrange theories by validating them on a scale that cannot be reached by experiments and by testing their scopes of application with different objects (Wylie 2011).

Beyond these reasoning patterns, historical scientists also emphasize the utilization of various epistemological strategies. Epistemological strategies seek additional sources of evidence to construct or support a hypothesis. For example:

- *Productive speculation* generates a hypothesis about a token event that connects phenomena that initially appear irrelevant. This guides the search for more relevant facts (Currie 2018).
- *Consilience* is the convergence of multiple independent lines of evidence on one property of a past event. The independence between them enhances the confirmation of the hypothesis (Forber and Griffith 2011).
- *Variety of complementary evidence* utilizes different lines of evidence that complement each other, collectively constituting a comprehensive picture of a phenomenon (Wylie 2011).

What is worthy of a highlight in this brief survey is that the reconstruction of tokens is often not only the aim of historical research, but it also plays important justificatory and methodological roles. It forms a core to unite multiple lines of evidence, paving the way for possible explanations and new discoveries. Moreover, with the lack of manipulation, token events that were interlocked in space and time offer evidence that complements or cross-validates each other. As I will argue in the next section, the methodological role of tokens is also integral to astronomy.

4. Historical reconstruction in astronomy

4.1. The “cosmic laboratory”

According to Anderl (2016), when studying general regularities, astronomers make use of the “cosmic laboratory,” that is “the multitude of different phenomena and environments naturally provided by the universe” (652).

Embracing this idea, Anderl champions the use of observational counterparts to controlled experiments: natural experiments and quasiexperiments (661–62). As the “direct equivalent of randomized controlled experiments,” natural experiments refer to the situations where groups with and without a factor happen to be “as-if” randomly assigned such that all other factors are evenly distributed between two groups of situations; in quasiexperiments, the observed situations do not happen to be exactly randomized in all other confounding factors, and one needs to evaluate them separately.

Nevertheless, the applicability of these strategies is often limited. “Randomization” of situations can rarely be achieved due to the correlation between the factor under study and others. Quasiexperiments also require the acquaintance with contextual information to analyze all the confounding factors. In many circumstances, the distribution of the target and confounding factors among the two groups cannot be independently determined. For example, to study the change in physical properties of stars throughout their evolutionary tracks, one needs to statistically compare stars at one stage with those at another. However, due to the reduced dimension of time in our snapshot observation, this leads to a circularity: to determine how old and at which stage those stars are, an evolutionary model that matches observable properties with age should have already been assumed.

Under these circumstances, the statistical methods based on the “cosmic laboratory” metaphor do not accurately describe how astronomers make inferences. I suggest that the historical reconstruction of tokens plays a crucial methodological role here. A token could help to decipher regularities of a type, offer contextual information for other tokens, and serve as a center of coherent tests.

4.2. “Rosetta Stone” situations

Astrophysicists utilize certain special token situations that I call the “Rosetta Stone” situations. Analogous to the Rosetta Stone that enabled deciphering the Egyptian scripts with contingently attached Ancient Greek decree, Rosetta Stone situations involve token objects or processes with special properties that come from the historical contingencies of the target object or its surroundings. They provide external information to decipher the internal properties of those tokens. The internal properties can subsequently be generalized to the whole type of objects sharing the same mechanism and close physical parameters.

One example of “Rosetta Stones” is star clusters. While stars scattering in the Galaxy might have any age or metallicity, stars in clusters are assumed to be born around the same region and at a very close time, sharing similar ages and metallicity and varying only in mass. These assumptions make those clusters an additional source of temporal information. They constitute an isochrone of stars with different masses, from which age could be tightly constrained using established stellar evolutionary models. Another piece of external information comes from the spatial position of those clusters in the Milky Way. Given the growth history of galaxies, globular clusters in the halo area are supposed to be much older than the open clusters on the disk, enabling age comparison across stars. This external historical information supported several important discoveries, one among which is the general sequence of stellar evolution (Dick 2013).

Rosetta Stone situations also circumvent another limitation of statistical methods: Certain properties being investigated are not accessible in a large number of stars. For example, for most stars far away, access to their internal physical processes and chemical elements is limited to the observation of the stellar atmosphere. With divergent sources of evidence that can be found in its vicinity, the Sun thus constitutes a privileged object of study. For example, the internal structure and dynamics of the Sun can be studied with helioseismology, which requires accurate close observation of its oscillation. Chemical components of the Sun are also more

comprehensively collected with the study of meteorites formed out of the same gas cloud (Frebel 2015). The evolutionary history of the Sun thus reconstructed is crucial for building an evolutionary model for a whole type of star with similar mass, luminosity, and chemical components.

Wylie (2011) has pointed out that detailed knowledge about the history of tokens plays crucial epistemic and methodological roles in providing independent evidence for coherence tests or complements. These roles can also be found in Rosetta Stone situations in astronomy. Astronomers often speak of a problem of degeneracy: observable kind-membership conditions of a “type” of objects or phenomena may underdetermine the real type in which the generalization of theories and methods is legitimate. For example, many classes of stars delimited with synchronic properties like color and luminosity do not guarantee that they share the same evolutionary model. Coherence tests with multiple lines of evidence help to identify subdivisions within an apparent “type” and clarify conditions for generalization. These multiple lines of evidence often come from the peculiar celestial objects whose histories interlock with each other and offer constraints to each other. In a recent episode studying the cooling model of white dwarfs, Cheng, Cummings, and Ménard (2019) adopt two independent ways of determining the age of high-mass white dwarfs, one derived from the velocity dispersion of disk stars predicted by the dynamic model specific to the Milky Way history, the other from the stars’ photometric properties. Within the entire population of white dwarfs that share similar photometric properties, the dynamic method spots a novel subpopulation that is in fact much older than the rest, a distinction that cannot be made without knowing the peculiarity of the Milky Way dynamics.

Finally, Rosetta Stone situations also offer contextual information for other tokens with interlocking histories. For example, transient dynamical effects caused by some Milky Way structures in the past provide contextual knowledge about the radial migration of certain stars, which enables the inference of their source and explains their anomalies (Minchev and Famaey 2010).

The use of “Rosetta Stones” highlights an important aspect of astronomy: the codevelopment of multiple research subdomains. First, the studies of general regularities and token historical events iteratively support and fuel each other. Rosetta Stone situations with better-known histories offer contextual information for building general models, understanding other particular systems, and unveiling new “Rosetta Stones.” Second, relatively autonomous subdomains that focus on different celestial objects, such as the studies of stellar physics, Milky Way dynamics, and larger-scale cosmology, are often cross-referenced to inform, support, or correct each other. Therefore, a complex web of evidence from both tokens and regularities, as well as from different subdomains, is crucial for the progress of astronomy.

5. Historical evidence in astronomy

5.1. *Traces and historical inference*

Cleland (2002, 2011) proposes a pattern of evidential reasoning in historical sciences. First, scientists generate a number of rival hypotheses about a past event to explain available traces, and second, they search for a new trace (aka “smoking gun”) that could be best explained by one of the hypotheses, thus offering justification for it.

Characteristic of many historical studies nonetheless, many have pointed out that an important element is missing.

The missing element is concerned with what can be counted as a “smoking gun” or a trace in general. Forber and Griffith (2011), for example, point out that what data counts as a “smoking gun” should not be taken for granted, but historical scientists use caution when evaluating the quality of potentially decisive evidence. Proponents of midrange theories argue that the plausibility of a common-cause explanation cannot be secured simply because it explains certain contemporary phenomena, but this connection should also be backed by regularities that guarantee its univocality and reliability (Jeffares 2008). Turner (2013) further points out that in historical sciences that have already acquired a large amount of data, scientists do not design hypotheses and then stumble upon a “smoking gun,” but they build theories and models by finding general patterns in the portion of data that is potentially relevant to some past events.

What these philosophers stress is that the standards by which a contemporary phenomenon is identified as a trace of a past event cannot be settled directly by the immediate epistemic gain brought by taking it as a trace. There is a more complex process of recognizing something as a trace, drawing relevance to a past event, and evaluating the reliability of such connections.

5.2. Identification of traces

Similar to the notion of evidence, a trace is not only a causal concept but also epistemological. Ontologically, a trace should at least be causally affected by some aspects of an event (Turner 2007); epistemologically, a trace is what scientists could identify and use to make inferences about a past event given their background knowledge (Currie 2018). These widely agreed characterizations nonetheless, it is still complicated how something can be reliably identified as a trace of something in the past.

Identification and classification of traces have been a central topic in astronomy. In the age of massive sky survey projects, massive data is coupled with the difficulty of identifying the relevant portion. Not all past events of interest leave lasting and univocal signatures that help to differentiate a trace from the rest of the data and connect it to the past. First, our limited ways of observation do not have access to many properties of a star. Second, certain historically informative properties are not captured by the way stars are typically classified for synchronic purposes. Third, it is often not the properties of an individual star that reveal the plausible connection to the past, but the statistical properties of a group of stars sharing a similar history. Thus, making inferences from a historically informative group usually goes hand in hand with identifying the exact members of the group. Finally, the applicability of midrange theories is often limited. The long lifetime of celestial bodies may introduce multiple local confounders, transformations, and decay. As a result, the contemporary properties that bear a salient relation to the past often alter from one target event to another. With this, instead of choosing a fixed set of properties supported by some general midrange theories for trace identification, it is more plausible to develop a *method* that could *find* the most prominent properties given a specific event under study.

Astronomers have developed strategies to identify historically informative groups of stars. They first build catalogs of stars with all their observable properties, and then apply a method called “phase space clustering” (PSC). “Phase space” denotes the space in which all possible states of a system can be represented. Each dimension of the space stands for a physical quantity of the system. Celestial bodies are often characterized by quantities such as age, metallicity, position, speed, angular momentum, energy, and orbital shape properties. PSC finds clusters and patterns in any space defined with several of these properties.

PSC proceeds iteratively from two directions. From the theory-driven direction, scientists construct candidate models and simulations to study how different settings of a past event may leave their influences. This forward simulation forecasts what observable quantities may be more characteristic of a hypothetical model and what the relevant distribution of observables might be. The data-driven direction analyzes available data in a number of promising phase spaces to find apparent correlations or clusters that may either match the model forecasts or hint at unrealized historical causes. Algorithms of data clustering introduce a measurement of similarities and differences in data, which presents evidential power for hypotheses that explain those similarities and differences (Cat 2022). This mitigates the lack of stable and systematic midrange theories and possible uncertainties in the forward approach.

One application of this method is the recent reconstruction of the Milky Way formation history. This episode started at a point when scientists were aware that galaxies are formed by smaller mergers but did not know the details of how the Milky Way was assembled. As scientists did not have a preliminary grasp of what patterns to look for, PSC started by clustering kinematic and metallicity data of all the galactic halo stars. Belokurov et al. (2018) identified two apparent clusters of stars sharply separated in the eccentricity-metallicity space. This turns out to be evidence of the presence of one massive merger event. With the scenario being secured backwardly, scientists forwardly construct simulations of how dwarf galaxies may merge into the Milky Way with varying parameters and project the possible patterns of remnants. The match between the projection and contemporary data indicates the most likely epoch, mass, and angular momentum of the progenitor. This led to the reconstruction of the Gaia-Enceladus progenitor galaxy.

Once reconstructed, the physical properties of the progenitor are used to forecast more patterns of its remnants and to reduce confounders when scientists search for other progenitors in different datasets and phase spaces. For example, Massari, Koppelman, and Helmi (2019) analyze a new dataset differing in both objects and properties. The new dataset records globular clusters (GC) with more properties including age, metallicity, energy, angular momentum, and orbit shape. They extend models of several entering progenitors to GCs and generate potential patterns of them in the integrals of motion space. The GC assignment is then tested by their good fit to functions in another space, the age-metallicity relation space. By clustering data from both directions and in different spaces, scientists assigned most of the observable halo GCs to their origins without much conflict. With this method, observable ex-situ GCs have been mostly assigned to six major progenitors and a few small accretion events without inconsistency both in the reconstructed merging history and in the attribution of contemporary data.

5.3. Confirmation in PSC and its condition of success

The episode defies the simplified schema of historical inference. The identification of traces goes hand in hand with the gradual establishment of hypothetical models that identify and explain the traces. The correct recognition of traces is only gradually secured over the iterative course of assigning traces, specifying models, and fitting them together consistently.

I identify three sources of confirmatory power in the final evidential web consisting of models and data. First, the models explain data patterns not only in the dataset where they are first established but also in other datasets. The confirmatory power of this explanation comes from the increasing consilience of evidence that supports each model. Second, all the apparent clusters found in these datasets can finally be reasonably assigned to an event or explained away without much conflict. Third, the constructed models for the events are consistent with each other. An argument from miracle could be applied to the last two conditions: If several models or clusters drastically deviated from the real situation, it would have been impossible that the data finally turned out completely explained and the models were consistent. This web presents a threefold fitting-together between model and data, models, and clusters in different datasets.

This complex web undermines the possibility of finding an alternative web that accounts for the data to a similar degree, mitigating the underdetermination problem. It has been stressed by Wylie (2011) and Currie (2018) that fitting together details in a local context has the epistemic virtue of countering underdetermination. PSC further extends this insight to a broader context, that is, to a series of studies that endeavor to explain different portions of available data and reconstruct a story of historical events.

Another epistemic virtue of PSC is that the combination of the two directions of identifying traces alleviates the concern about theory ladenness. Patterns and clusters are not only recognized solely according to properties projected by models but also by clustering algorithms. Therefore, even if models are occasionally involved in the process of data processing, it does not necessarily lead to the confirmation of itself.

A possible failure in this process is misrecognizing certain confounder star clusters as relevant. The Hercules stream, for example, was misrecognized as remnants of a merger event for their shared dynamic properties. The misrecognition is noticed and corrected when additional datasets about age and metallicity are collected and taken into consideration (Bensby et al. 2008). Thus, one important condition of success in PSC is the collection of various datasets and the cross-validation between them. By various datasets, I mean datasets about different objects or physical properties. When reconstructing the Milky Way merger history, astrophysicists have been shifting between datasets of individual stars and globular clusters, as well as across phase spaces with different combinations of physical properties. With the cross-validation between different datasets, not only more background knowledge about different types of objects and physical processes is involved to inspect the validity of those clusters and models but also the invariance of clusters in data also enhances their reality.

6. Conclusion

Philosophical studies of special sciences not only reveal surprising epistemological and methodological similarities between disciplines but also deepens the understanding of general philosophical issues by absorbing the peculiarities of each

discipline. The connection I draw between modern astronomy and historical sciences contributes to the understanding of both. First, it highlights the importance of the historical reconstruction of token events in astronomy, enabling a better description of astronomical inferences. Second, the peculiar difficulty of partitioning stars in large datasets enriches our understanding of the notion of “trace” and how they can be identified.

Acknowledgments. I express my gratitude to Jordi Cat, Jutta Schickore, Jamie Shaw, Shannon Abelson, and Ann-Sophie Barwich for their constructive feedback on an earlier draft of this paper. I am grateful to Sihao Cheng for inspiring discussions and constant support.

References

- Andler, Sibylle. 2016. “Astronomy and Astrophysics.” In *The Oxford Handbook of Philosophy of Science*, edited by Paul Humphreys, 652–70. Oxford Handbooks. <https://doi.org/10.1093/oxfordhb/9780199368815.013.45>
- Belokurov, Vasily, Denis Erkal, N. W. Evans, S. E. Koposov, and Alis J. Deason. 2018. “Co-Formation of the Disc and the Stellar Halo.” *Monthly Notices of the Royal Astronomical Society* 478 (1):611–19.
- Bensby, Thomas, M. Oey, S. Feltzing, and B. Gustafsson. 2008. “Hercules-Stream Stars and the Metal-Rich Thick Disk.” In *The Metal-Rich Universe*, edited by Garik Israelian and Georges Meynet, 62–68. Cambridge: Cambridge University Press.
- Cat, Jordi. 2022. “Synthesis and Similarity in Science: Analogy in the Application of Mathematics and Application of Mathematics to Analogy.” In *Metaphors and Analogies in Sciences and Humanities: Words and Worlds*, edited by Shyam Wuppuluri and Anthony C. Grayling, 115–45. Cham: Springer.
- Cheng, Sihao, Jeffrey D. Cummings, and Brice Ménard. 2019. “A Cooling Anomaly of High-Mass White Dwarfs.” *The Astrophysical Journal* 886 (2):100.
- Cleland, Carol E. 2002. “Methodological and Epistemic Differences between Historical Science and Experimental Science.” *Philosophy of Science* 69 (3):474–96.
- Cleland, Carol E. 2011. “Prediction and Explanation in Historical Natural Science.” *The British Journal for the Philosophy of Science* 62 (3):551–82.
- Currie, Adrian. 2018. *Rock, Bone, and Ruin: An Optimist's Guide to the Historical Sciences*. Cambridge, MA: MIT Press.
- Dick, Steven J. 2013. *Discovery and Classification in Astronomy: Controversy and Consensus*. Cambridge: Cambridge University Press.
- Forber, Patrick, and Eric Griffith. 2011. “Historical Reconstruction: Gaining Epistemic Access to the Deep Past.” *Philosophy and Theory in Biology* 3 (201306):1–19.
- Fox, Craig W. 2021. “Back to the Beginning: An Empiricist Defense of Scientific Stories About the Past.” PhD diss., Western University.
- Frebel, Anna. 2015. *Searching for the Oldest Stars: Ancient Relics from the Early Universe*. Princeton University Press.
- Jacquot, Melissa. 2022. “Observations, Simulations, and Reasoning in Astrophysics.” *Philosophy of Science* 87 (5):1209–20.
- Jeffares, Ben. 2008. “Testing Times: Regularities in the Historical Sciences.” *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 39 (4):469–75.
- Massari, Davide, Helmer Koppelman, and Amina Helmi. 2019. “Origin of the System of Globular Clusters in the Milky Way.” *Astronomy & Astrophysics* 630:L4.
- Mincev, Ivan, and Benoit Famaey. 2010. “A New Mechanism for Radial Migration in Galactic Disks: Spiral-Bar Resonance Overlap.” *The Astrophysical Journal* 722:112.
- Ruphy, Stéphanie. 2016. *Scientific Pluralism Reconsidered: A New Approach to the (Dis)Unity of Science*. Pittsburgh: University of Pittsburgh Press.
- Turner, Derek. 2007. *Making Prehistory: Historical Science and the Scientific Realism Debate*. Cambridge: Cambridge University Press.

- Turner, Derek. 2013. "Historical Geology: Methodology and Metaphysics." In *Rethinking the Fabric of Geology: Geological Society of America Special Paper 502*, edited by Victor R. Baker, 11–18. Geological Society of America.
- Wilson, Curtis. 2003. "Astronomy and Cosmology." In *The Cambridge History of Science*, vol. 4, edited by Roy Porter, 328–53. Cambridge University Press.
- Wylie, Alison. 2011. "Critical Distance: Stabilising Evidential Claims in Archaeology." In *Evidence, Inference and Enquiry*, edited by Philip Dawid, William Twining, and Mimi Vasilaki, 371–94. Oxford: Oxford University Press.