

How Archaeological Evidence Bites Back: Strategies for Putting Old Data to Work in New Ways Alison Wylie

Science, Technology and Human Values, to appear in a special issue on "Data Shadows: Knowledge, Openness and Absence," edited by Sabina Leonelli, Gail Davies, and Brian Rappert.

Accepted 5 September 2016.

Published OnlineFirst (11 October 2016) as doi:10.1177/0162243916671200

Available through: <http://sth.sagepub.com/>

Abstract

Archaeological data are shadowy in a number of senses. Not only are they notoriously fragmentary but the conceptual and technical scaffolding on which archaeologists rely to constitute these data as evidence can be as constraining as it is enabling. A recurrent theme in internal archaeological debate is that reliance on sedimented layers of interpretative scaffolding carries the risk that "pre-understandings" configure what archaeologists recognize and record as primary data, and how they interpret it as evidence. The selective and destructive nature of data capture in archeology further suggests that there may be little scope for putting "legacy" data to work in new ways. And yet archaeologists have been strikingly successful in mining old datasets for new insights. I situate these concerns in the broader context of debate about the epistemic standing of the historical sciences, and then consider three strategies by which archaeologists address the challenges posed by legacy data. The first two – secondary retrieval and recontextualization – are a matter of reconfiguring the scaffolding that underpins evidential reasoning. The third turns on redeploying old data in the context of computational models that support the experimental simulation of the cultural systems and contexts under study.

I. The shadowy nature of archaeological data

Archaeological data are shadowy in a number of senses. One might say that archaeology as a discipline is defined by the challenges of working with gaps and absences in its primary data. The shadows archaeologists wrestle with are not just a function of the notoriously fragmentary and incomplete nature of the surviving "data imprints" of past lives but reflect, as well, the paucity and instability of the inferential resources they rely on to "bring data to light...to realize its value" as evidence (see the Introduction to this Special Issue). The worry here is that what archaeologists recognize as data, and what they infer to be its evidentiary significance, are necessarily functions of the "pre-understandings" they bring to bear (Bell 2015)¹; these include assumptions about the cultural/historical subjects they study, as well as an enormously diverse range of background knowledge and technical resources that inform their use of surviving material traces as evidence. I refer to these assumptions, knowledge and resources collectively as "scaffolding," conceptual and technical. A recurrent theme in debates within archaeology is that, given these limitations of primary data and scaffolding, vicious circularity is inescapable; archaeologists' data are "useable to prove claims or foster discoveries" (see the Introduction to this Special Issue) only insofar as they are legible as evidence, and they are legible only if they conform to expectations embedded in the scaffolding of pre-understandings that define the subject domain and set the research agenda. This epistemic anxiety is reinforced by the fact that the recovery of archaeological data is necessarily selective and often destructive; what the original excavators did not know to recover is gone forever and what they did record may be so thoroughly configured by pre-understandings that it cannot be expected to expose error inherent in them and is most likely not useable for purposes they did not anticipate.

And yet, time and again the material shadows of past lives have proven to be a rich evidential resource that archaeologists have put to good use, dramatically enlarging and repeatedly reconfiguring what we know about the past, often pushing inquiry in unexpected directions. Moreover, although it is the discovery of novel material traces that captures the public imagination and dominates headlines, as often as not these accomplishments are the result of painstaking reexamination of existing bodies of data, including field records and disciplinary histories as well as assemblages of cultural material, geophysical

¹ This is a term introduced by Hodder (1999); I use it in the sense discussed by Bell (2015: 55).

and biological samples. As often as not the process of repurposing legacy data calls into question the very pre-understandings that made it possible to "capture" these data in the first place (Chippindale 2000). Shadowy though they may be, archaeological data have considerable capacity to bite back in unexpected ways—a capacity, I argue, that lies not in some mark of empirical integrity that secures their status as epistemically foundational but in the strategies archaeologists have developed to continuously build and rebuild the scaffolding for evidential arguments that are understood to be provisional. I take this practice to be typical of any field in which researchers confront the challenge of "building scientific knowledge in the absence of infallible foundations," as Hasok Chang puts it (2004: 234).

The question about data shadows that I address here is: How do archaeologists press legacy data—data gathered by previous generations of archaeologists— into service to address new questions? I begin by noting that the epistemic anxieties that haunt archaeological debate are by no means unique to archaeology; I identify parallels in philosophical debate about the epistemic standing of the historical sciences. I then characterize three strategies by which archaeologists address the challenges posed by legacy data. This is a case-based analysis intended to tease out close-to-the-ground epistemic norms that are embodied in these strategies. These, I argue, bring into focus a broad spectrum of epistemic possibilities that is obscured when anxieties about the security of surviving material traces are amplified into general claims about the epistemic standing of the historical sciences.

II. Epistemic pessimism vs. optimism about the historical sciences: The role of scaffolding

The challenges and the seemingly improbable successes that arise from working with "trace evidence"² pose a conundrum that is not unique to archaeology. Indeed, it is characteristic of the historical sciences and has generated a philosophical literature that tends to extremes of epistemic pessimism and optimism. Bertrand Russell's five-minute hypothesis³ bounds the spectrum at the skeptical end (1921); appeals to intransigent underdetermination and a long history of unfavorable comparison with the experimental sciences make up a range of more realistically pessimistic views; and, at the other end of the spectrum, Carole Cleland argues that the continuous dispersal of traces from originating events underwrites an expansive optimism (2002). Adrian Currie usefully charts this debate in terms of an exchange about Cleland's overdetermination thesis in which Derek Turner (2007) raises concerns about the "epistemic retrievability" of material traces of past events and conditions, in this case, paleontology (Currie 2014: 50-84, 2016a: 17-20). Turner defends a principled epistemic pessimism based on a consideration of how the cascade of effects invoked by Cleland is subject to continual attrition and degradation, as well as the vagaries of dependence on limited and fallible background knowledge that may leave traces mute as evidence, even if they do survive. Currie makes a compelling case that neither view does justice to actual practice in the historical sciences. I endorse this conclusion: given how widely their epistemic resources and inferential strategies vary, we should be suspicious of any generalizations about the epistemic predicament of these disciplines.⁴

As a framework for case-by-case analysis Currie proposes what he calls a "ripple model" of historical evidence that incorporates the plausible elements of these contrasting stances (2014: 52-68). An originating event, or set of processes or states of affairs, can generate a widely dispersed and proliferating cascade of downstream effects, as Cleland argues, so that, despite the attrition that concerns Turner, the potential for there being traces that could bear historical witness to the past (cultural or otherwise) must be counted as open-ended. At the same time, the status of these traces as useable data and as evidence is a function of background knowledge in light of which their connection to the past

² I use this terminology as Currie has developed it in "What is a Trace?" (2014: 42-9).

³ Russell takes as his point of departure the observation that there is no "logical impossibility in the hypothesis that the world sprang into being five minutes ago exactly as it then was," complete with all the evidence on which we rely to make claims about a deeper past. He argues on this basis that we would have no way of learning that we are wrong (1921: 159-60). Chapman and Wylie discuss a similar argument developed in the context of British archaeology in the 1950s (2016: 15-18).

⁴ See also Currie (2016b) where he further develops this critique of pessimistic appraisals of the evidence available to the historical sciences that generalize from what he argues is an impoverished view of the methodologies employed in these fields.

events of interest to historical scientists can be recognized and reconstructed. While the "epistemic retrievability" of any past event or set of conditions depends on the vagaries of trace survival, as the pessimists emphasize, the effectiveness of the historical sciences in identifying and putting these traces to work as evidence depends on constructing or recruiting relevant background knowledge about how they could have been produced, and about conditions that affect their transmission and degradation—resources that constitute what I refer to here as conceptual and technical scaffolding.⁵ It is the ongoing process of building and rebuilding this scaffolding that secures the prospect of advances that sometimes dramatically shift the horizons of inquiry in the historical sciences. The ripple model makes systematic a "trace-centric" understanding of historical evidence in terms that "replace dichotomy with a graded set of possibilities" (Currie 2014: 83).

Beyond capturing the wisdom inherent in the practices of historical inquiry that is sometimes obscured by the dynamic of debate, Currie also challenges the myopia of trace theorists, arguing that in their preoccupation with strategies of "gap compensation" (2014: 194-8) they fail to recognize a number of other streams of evidence on which historical scientists rely: specifically, various forms of indirect and surrogate evidence. His most provocative suggestion is that simulations of past systems and events can quite literally "generate new evidence" (189), and that manipulations of these models can function like experimental interventions. They play not just the heuristic and explanatory roles typically attributed to them by philosophers of science, but also an evidential role: they are a source of data that fill gaps in the surviving repertoire of traces, making it possible to build and assess hypotheses about otherwise inaccessible aspects of a target system or process. Whether or not simulations are the source of a distinctive form of evidence overlooked by trace-centric theorists, they do represent a strategy by which historical scientists leverage multiple lines of evidence to build and assess hypotheses about aspects of the past that did not themselves leave "data imprints"—a strategy that warrants close attention as a source of insight about the past that has the capacity to destabilize and reconfigure pre-understandings.

Long-running epistemic debate within archaeology parallels the philosophical debate about the epistemic status of the historical sciences in which Currie intervenes. Five minute hypotheses and light-cone scatters of traces are not a concern for archaeologists, but there is a recurrent pattern of vacillation between optimism inspired by the open-ended possibilities of recruiting as yet untapped shadow data as an evidential resource and a profound epistemic pessimism that arises from confronting the limitations of these data (Wylie 2002: 25-41; Chapman and Wylie 2016: 18-31). In addition to the frustrations of working with an attenuated trace record and navigating the insufficiency (and instability) of the scaffolding necessary to put it to work as evidence—the sources of pessimism central to philosophical debate about the historical sciences—archaeologists add the concern I raised at the outset: the epistemic limitations they have to reckon with are a function not only of an incomplete database or the lack of relevant scaffolding but also of constraints imposed by scaffolding that has been put in place by previous generations of archaeologists.⁶

⁵ "Scaffolding" has a wide range of meanings in the philosophical and science studies literature. I draw on the account Wimsatt has developed in connection with the conceptual "architecture" he proposes as a framework for understanding the evolutionary dynamics of non-biological systems (2014), including the path-dependent processes that shape research programs and technologies. Wimsatt identifies various types of scaffolding that play a role in these processes, ranging from material resources, skilled practice, and institutional and social infrastructure, to the kinds of background knowledge and analytic techniques on which I focus here. Wimsatt's concept of "scaffolding" includes many of the elements that comprise "repertoires" (Leonelli and Ankeny 2015) and the "shared cognitive-emotional-interactional platforms" (Mansilla, Lamont and Sato 2016) that make interdisciplinary collaboration possible. He is also influenced by the more technical senses of scaffolding associated with theories of cognitive development. I am not drawing on this background so much as on Wimsatt's observations about the downstream consequences of entrenchment (2014: 82-92), and on complementary philosophical discussions of scaffolding for evidential reasoning, for example, Toulmin's account of the role of inferential "warrants" in domain-specific working logic (1958: 98) and Norton's discussion of "material postulates" (2003: 648; see Chapman and Wylie 2016: 33-40, 202-3).

⁶ There are obvious resonances here with Hanson's classic account of the theory-ladenness of observation (1958), Fleck's analysis of the ways in which the conceptual constructs of a "thought collective" come to be accepted as "facts" that are themselves "determined by their 'ancestors'" (1979/1935: 20), and Kuhn's historical account of paradigm-dependence (1970). These constitute the intellectual background of my analysis; I argue that attention to

The worry here is that entrenched pre-understandings about the cultural, historical subject and its material traces—a key component of the scaffolding that makes archaeology possible—configure the interpretive or explanatory claims that are the focus of inquiry and determine what will count as evidence in building or adjudicating them. Anxieties about vicious circularity, and the epistemic pessimism that follows from them, are motivated by the appreciation that these pre-understandings cannot be eliminated. They shape every aspect of archaeological practice, embodied as they are in the conventions of fieldwork and categories of analysis (formal, spatial, temporal, functional, stylistic, to name a few) that are essential to the creation of archaeological databases. However entrenched they have become as norms that define what it is to do archaeology, they are contingent, purpose-built creatures of context. They were invariably designed to address specific questions that presuppose a rich array of substantive assumptions about the nature of the subject domain, what's puzzling or interesting and, crucially, what can feasibly be asked about it. Reinforced by funding streams, publication practices, and communication networks, these pre-understandings often persist in conventions of training and practice long after the original questions that prompted them are forgotten and the assumptions that framed them have been lost from view.⁷ As scaffolding that enables inquiry, precisely because they are taken for granted, entrenched pre-understandings create repositories of legacy data that are shadowed by a vast store of data that are illegible as evidence. My focus here is on how archaeologists have successfully mobilized these data in a process that both depends upon and demands critical reconfiguration of this scaffolding of pre-understandings.

III. Archaeological strategies for eliciting useable data from the shadows

Consider, then, three strategies by which archaeologists put old data to work, which I illustrate in terms of five brief case studies.⁸ The first two strategies are a matter of reconfiguring different aspects of the scaffolding that underpins evidential reasoning in ways that make possible the secondary retrieval of old data and its recontextualization. The third turns on redeploying old data in the context of innovative modeling practices that approximate the experimental simulation strategies described by Currie.

1) *Secondary retrieval: extracting new data from old*⁹

Most straightforward, at least from a philosophical point of view, are strategies that involve literally extracting new data from old, often by bringing to bear new technical tools that enlarge the archaeological database. The multiple "radiocarbon revolutions" initiated by Libby in the 1950s are perhaps best known for this, but they are by no means unique. Consider two examples that illustrate how deeply these practices of secondary retrieval can destabilize the pre-understandings that inform long-running research programs.

The Roman Diaspora Project. This is a project in which archaeologists made sophisticated use of trace element and stable isotope analysis of samples of dental enamel to reconstruct the likely origins and lifetime travel of individuals buried in the late period Roman cemetery of Lankhills in Winchester, U.K. (Eckardt et al. 2009, Leach et al. 2009). An interpretation of burials in this cemetery dating to the 1970s had focused on material traces that were assumed to be markers of cultural affiliation, degree of "Romanization," and "incomer" status, for example: epigraphy, statuary, and grave goods; patterns of spatial orientation and distribution in cemetery layout; and skull morphology. Over four decades the conceptual framework of Romanist research had shifted substantially as critical reassessment called into

the specifics of archaeological practice brings into focus strategies by which the constraints of encompassing paradigms, "thought collectives," and semantic webs of theoretical association can be held accountable (Wylie 2002: 117-126; Chapman and Wylie 2016: 206-210).

⁷ For an illustration of how this works, see the discussion of fieldwork traditions in Chapman and Wylie (2016: 55-87).

⁸ This account originated in analyses of "ignorance" (Wylie 2008) and of "how facts travel" in archaeology (Wylie 2011a) and is developed in more detail, although in somewhat different terms, in Chapman and Wylie, "Working with Old Evidence" (2016: 93-143).

⁹ I refer here to the physical and analytic "extraction of new data from old," not to the reinterpretation of data as the phrase, "secondary retrieval", is understood in some contexts.

question assumptions about the uniformity of Romanization across the empire, and the projection of contemporary class and race-inflected interpretations of social status, hierarchy and mobility. The shift in these orienting pre-understandings reopened questions about the origins, affiliations and travels of individuals buried in the Lankhills cemetery, which the Diaspora project team addressed both by reanalyzing existing data and by using a suite of analytic tools that had been developed in the interim to extract new data on the composition of the human remains. The analysis of dental enamel for oxygen isotopes was the basis for positing the probable locus of birth and early growth, and strontium isotope profiles provided lifetime dietary profiles that indicate where individuals may have lived and travelled in the course of their lives.¹⁰ Making use of the results of these analyses for archaeological purposes was by no means straightforward, however; it required the Diaspora researchers to develop purpose-specific scaffolding in the form of geological and ecological/subsistence datasets for the regions under study. The results of the oxygen isotope analysis could only be interpreted as evidence of where an individual lived in their early childhood given baseline knowledge of cross-continent clines in the mineral composition of groundwater, and the strontium isotope profiles are informative of lifetime travel only given an understanding of the regional geology that could produce the isotopic signatures (by way of the food consumed) reported for individuals interred in the Lankhills Cemetery.

Although the Diaspora authors are explicit about the uncertainties of their analyses—they reject any assumption that isotope profiles can establish a secure “postcode” identification of locus of origin and range of travel—the cautiously framed conclusions they draw about the Winchester cemetery population are quite striking. Local-style grave goods and mortuary treatment proved to be associated with just half the individuals who had local isotopic profiles, as well as with a number of those who likely originated well outside the British isotopic range. In only one case did isotopic profile and burial treatment coincide in suggesting an origin in central Europe, the “Pannonian” affiliation posited for most of those originally identified as “incomers” (Eckardt et al. 2009: 2824). The results of other cemetery reanalyses found individuals of North African origin buried in elite graves, and several children who were almost certainly born at considerable distance from the region where they died. These results not only challenge specific interpretive claims that had been made about individual burials; they also substantiate the growing critique of received views about diversity and mobility in the Roman Empire and destabilize interpretive conventions that equate styles of material culture with ethnic identity (Eckardt et. al 2009; Leach et. al 2009). It was not only centurians who moved throughout the empire, but also women and children. By the late stages of the Roman Empire local populations in Britain were quite diverse, and high rank or social status, as marked by burial treatment and grave goods, did not consistently track racialized or ethnic identities of origin. Some “locals” who lived abroad retained cultural markers of their histories of travel when they returned home, while some “incomers” assimilated to local traditions. In short, the extraction of new data from old, motivated by new questions and made possible by new technical scaffolding, reinforced a systematic reassessment of pre-understandings that had defined the horizons of Romanist archaeology and informed the interpretation of archaeological data as evidence.

Archaeometallurgy. By contrast with the Diaspora project, this is an example of secondary retrieval in which critical insights that were largely unanticipated and, for some, profoundly unwelcome, arose from the cumulative results of a long tradition of chemical analysis rather than from the introduction of new technical scaffolding. Mark Pollard and Peter Bray (2015) trace the history of the analysis of metal artifacts going back to the 18th century. Early studies focused on identifying the composition of alloys, but by the mid-19th century attention had turned to questions about the geographical sources of the metals that made up these artifacts. The technical scaffolding necessary for trace element analysis was introduced in the 1930s, and since that time, Pollard and Bray argue, the “quest for provenance” and the assumptions that underpin it had persisted largely unchanged. They are clearly evident, they argue, in the rationale for an ill-fated research program that focused, in the 1980s and 1990s, on using lead isotope analysis (LIA) to build a comprehensive database of chemical profiles for geological ore sources and the copper artifacts that make up European Bronze Age assemblages. The hope was that LIA analysis would

¹⁰ These are described as two causally independent isotopic systems (Eckardt et al. 2009: 2818). The oxygen isotope profile is indicative of the composition of drinking water and, therefore, the temperature and climate of the region in which an individual spent their early life, while the strontium profiles reflect the underlying geology of the region where the food originated that an individual consumed over their lifetime (Chapman and Wylie 2016: 188).

yield a "chemical or isotopic 'fingerprint'" that could finally resolve longstanding questions about where these assemblages had originated. Its advocates proceeded on the assumption that "some characteristic of the source of the raw material passes through to the finished object and allows a relationship to be established between the two" (2015: 116).¹¹ The results, Pollard and Bray report, are disappointing.

What many still regard as puzzles to be resolved Pollard and Bray see as anomalies that call into question the "fundamental assumptions of provenance" (2015: 116). The geological sources proved to be more heterogeneous than initially recognized, but even with a refined mapping of Mediterranean ore fields the chemical complexity of metal objects still resist any neat attribution to a geological locus of origin. Pollard and Bray suggest that this is to be expected if you recognize that the more volatile elements, such as arsenic and antimony, are liable to loss through oxidation and to differential "washing out" in the course of smelting and casting. But rather than treat the impact of human intervention on the compositional profile of metal artifacts as a confound that distorts the signal of an originating ore source, Pollard and Bray argue for a fundamental rethinking of what it is that archaeometallurgists study. The point of departure should be a recognition that the chemical profiles of metal objects are dynamic, not static, and that, given a number of other lines of archaeological evidence, repeated melting, mixing and repurposing was a common practice "from the time of the earliest use of metal" (116). In short, the subject of inquiry is "not merely a set of material processes but also the human decisions and structures that surround them" (2015: 125).

On this conceptual reframing of the research program, the chemical mutability that undermines the quest for provenience becomes a resource that can be used to reconstruct the socio-technical biographies of metal objects. The key to doing this, Pollard and Bray argue, is to develop new technical scaffolding in the form of decay tables that reflect what the chemical profile of a unit of copper becomes when trace elements are differentially lost or transformed. Rather than seeking a static association of individual artifacts with an originating copper source, these chemical types, combined with information from the excavation of regional mines, ore petrology, artifact typologies and regional chronologies, serve as a basis for characterizing the dynamics by which assemblages of metal objects changed over time as well as being dispersed geographically.

In the case of the Diaspora project, the introduction of a new analytic technique put archaeologists in a position to reassess claims about late Roman burials, confirming and extending critiques that had already called into question, on other grounds, the pre-understandings that informed their original recovery and interpretation. By contrast, it was the unexpected complexity of the archaeological database brought to light by increasingly refined methods of chemical analysis that forced the point in the archaeometallurgy case. The cumulative weight of these unobliging material traces destabilized the pre-understandings that motivated the long-running "quest for provenience," bringing sharply into focus the need to reconceptualize the subject of inquiry, reframe the questions asked of it, and build scaffolding adequate to the task of addressing these questions. In both cases, strategies of secondary retrieval resulted in evidential claims that effectively counter the risks of circularity inherent in a reliance on scaffolding configured by entrenched assumptions about the target of inquiry. They do this not by identifying some uniquely secure, foundational line of evidence but by exploiting multiple empirical and conceptual constraints inherent in the scaffolding of background knowledge that makes these traces useable as evidence and in the material traces themselves.

2) *Recontextualizing data*

A related set of practices for putting old data to work in new ways turns on reanalyzing an existing body of primary data with the aim of documenting the "conjunctions" between its constituent elements or properties (Taylor 1948).¹² The result is a body of second-order data that captures the structure of the archaeological record: "facts about the relationships between types of material and their distribution,

¹¹ The complicated history of Lead Isotope Analysis in the UK is described in more detail in Chapman and Wylie (2016: 164-184).

¹² This use of secondary analysis is developed in more detail in an analysis of the Iron Age village of Glastonbury (Chapman and Wylie 2016: chapter 4).

comparisons, and patterning in the archaeological record" (Boozer 2015: 98). Recontextualizing data, in the sense of situating material traces in the context of newly recognized conjunctions, can be as powerful a means of putting shadow data to work as the extraction of new primary data.

The Karanis typology. A standard low-tech strategy, crucial for assessing and reconfiguring artifact and site typologies, turns on bringing new archaeological comparanda to bear. In a discussion of the "tyranny of typologies" Anna Boozer (2015) describes excavating a house structure that did not conform to the accepted typology of Roman Egyptian domestic architecture: it lacked a second floor and a central room with a roof, and showed evidence of food preparation and other activities in interior spaces rather than in an exterior courtyard. She retraced the steps by which the extant typology had been developed and found that it had been based on house forms identified at Karanis, a heavily referenced Romano-Egyptian site that had been excavated in the 1920s and 1930s that had become entrenched as the type-site for domestic architecture in the region. Despite its canonical status Boozer found that key questions about the stratigraphy of the site had never been resolved and its publication was incomplete with respect to exactly the details that underpin the typology. The site report focuses exclusively on the largest houses, ignoring the architectural layout of the most common and more modest houses, and specialist analyses of distinct classes of material are reported in aggregate terms for the site as a whole, rendering problematic any determination of room or area function. Boozer concludes that the empirical basis for the typology in terms of which subsequent Roman Egyptian domestic architecture had been described and analyzed proved to be an arbitrary selection of houses from a single, poorly understood site.

Invoking a different set of archaeological comparanda—house compounds excavated in the late 1990s on neighboring sites—Boozer argues that Roman Egyptian domestic architecture should be recognized to include a much broader "spectrum of house options" than represented in the Karanis-anchored typology (2015: 105). This sounds straightforward enough, but she reports stiff resistance to any proposal to break with the Karanis typology. Those urging its reassessment bear a heavy burden of proof to demonstrate, for example, that the anomalies they cite are not due to poor preservation and that the new comparanda they invoke are not marginal cases that should be set aside.

Although Boozer raises a number of concerns about the "tyranny" of typologies, she is clear that they are essential to archaeological practice. They are a paradigm example of conceptual scaffolding: a "repository of interpretive insight" and the necessary basis for building up conjunctions. They function as a medium of communication and a framework for systematizing archaeological data precisely because they reduce complexity. But when descriptive categories that are empirically problematic become entrenched they have "palpable downstream effects" (2015: 105), canalizing inquiry in ways that reproduce errors built into them at the outset. It is crucial, Boozer argues (citing Gero 2007), to "honor ambiguity" rather than smoothing, cleaning and otherwise suppressing the uncertainty of the jointly descriptive and interpretive claims that become the basis for subsequent reasoning with archaeological data. The recontextualizing strategies Boozer relies on to recapture some of the complexity of Romano-Egyptian domestic architecture include not only second-order analysis aimed at identifying conjunctions that had been read out of account by this "ill-conceived and empirically inadequate" typology (Boozer 2015: 98), but also a critical history of the formation of this typology that shows how and where the errors arose. Recontextualizing the anomalous house forms in relation to others that do not fit the Karanis typology makes the data derived from them usable—it brings them out from the shadows—and at the same time provides an empirical basis for contesting the scaffolding that had rendered them illegible.

The third radiocarbon revolution. Archaeological chronologies are similarly prone to founder effects of the kind Boozer describes for the Karanis-based house typology. Famously, radiocarbon dating got uptake after in the post-war period as a high-tech method of extracting new data that, it was hoped, could anchor archaeological chronologies to absolute dates, rendering obsolete any reliance on relative dating techniques such as typological seriation and stratigraphic sequences. The radiocarbon revolution set in motion by the first introduction of ^{14}C dates was, indeed, "sensational"; as Sturt Manning puts it, the results "entirely restructured the practice and understanding of prehistoric archaeology around the world" (2015: 128). By the late 1950s, however, questions were being raised about the reliability of ^{14}C results. Within a decade it was clear that the ratio of radioactive to stable carbon in the atmosphere could not be assumed to be stable or uniform. In addition, growing concern about the risks of contamination and the

inconsistency of results produced by different dating laboratories made it a priority to standardize practices of sample recovery and analysis. The last sixty years has seen what Manning refers to as a second radiocarbon revolution: a long, complex process of calibrating the atmospheric ^{14}C dates for samples of known age (e.g., based on tree-ring data) and reconciling discordant chronologies on this basis (2015: 131). The impact of this second revolution has been no less dramatic than that of the first, reconfiguring longstanding assumptions about cultural sequence and associations. But in the process a third radiocarbon revolution has taken shape that centers on contextualizing radiocarbon dates in relation to a wide range of other sources of chronological evidence, including the archaeological chronologies they were meant to render obsolete.

Two insights inform this most recent turn. One is the prosaic point that however precise radiocarbon analysis is, it dates a natural event—when organic material stopped exchanging carbon with the atmosphere, fixing the carbon content of a sample—so its use as archaeological evidence requires a series of inferences about how this natural event relates to the cultural contexts and events of interest. The other is that radiocarbon dates are probabilistic estimates of this natural event. Given wiggles in the calibration curves painstakingly developed through the second revolution it has become clear that calibration alone cannot resolve questions about where the actual date of an originating event lies in the spectrum of possible dates estimated from the ^{14}C signature recorded for a given sample. Taken together, these two points reinforce the now conventional wisdom that the results of any physical dating method must be interpreted in light of other contextual and chronological evidence. So rather than seeking the certainty of an imagined silver bullet—a physical dating technique that can displace reliance on archaeological chronologies and their “web” of background assumptions—the aim, Manning argues, is to “fully integrate archaeological information with ^{14}C dating in order to address archaeologically relevant (and therefore socially relevant) timescales and episodes” (Manning 2015: 151).

To this end, Alex Bayliss and Alistair Whittle advocate a “pragmatic Bayesian” approach designed to rigorously assess margins of error in physical dating (2015).¹³ This involves, first, a process of secondary retrieval: assessing the provenance and integrity of samples from which ^{14}C dates are drawn and discarding any for which there is a risk of contamination or for which archaeological associations are uncertain. It then requires a process of contextualizing the resulting ^{14}C dates in relation to multiple lines of evidence, for example, design sequence data, typological convergence and spatial distribution, the purpose of which is to delimit the range of dates that are plausible for the events of archaeological interest. Several different chronological models are generated for the target event and methods of sensitivity analysis are applied to test the stability of the model outputs and “identify which components of a model are most critical to the resultant chronology” (Bayliss and Whittle 2015: 234). The credibility of the chronologies that result depends, they argue, not on the singular authority of radiocarbon dates but on the whole range of substantive background and collateral knowledge that informs assessments of the integrity of the samples and reconstruction of the archaeological context in which they occur. Bayliss and Whittle characterize this as an “explicit methodology for weaving together different strands of evidence” (2015: 223); the credibility of the outcome is a function, not of any one line of evidence that is presumed to be foundational, but of what John Norton describes as “highly connected, massively tangled structure[s]” of inductive support (2014: 673).

These practices of secondary retrieval and recontextualization are well-established archaeological strategies of “gap compensation” (Currie 2014: 194) that embody several normative principles that practitioners make explicit. One is the point about honoring ambiguity made in particularly pointed terms by Boozer: keep in view the anomalies and the messiness inherent in the primary data that are necessarily regimented by conventions of description and analysis, and hold these scaffolding frameworks accountable to the data they are meant to systematize. A second is an insistence that archaeologists should not acquiesce to a “‘good enough’ attitude towards scientific evidence” (: 168 2015). Bayliss and Whittle advocate their pragmatic Bayesian approach on grounds that it forces archaeologists “to be explicit about...their reasoning” and, in this, counteracts the potential for provisional models to

¹³ See Steel for discussion of the role that Bayesian statistics play in the computational algorithms by which radiocarbon dates are calibrated, as well as inspiring this pragmatic Bayesian stance (2001: 153–4). And for a more detailed account the issues raised by this history of radiocarbon dating see Chapman and Wylie (2016: 147–163).

"fossilise" and "become received wisdom" (2015: 223). This commitment to critically scrutinize pre-understandings that have got entrenched depends not only on reanalysis of the primary data and the "conjunctives" built up through secondary analysis of these data, but also on empirical, conceptual and, crucially, genealogical interrogation of the scaffolding used to interpret them as evidence. Finally, a third norm of best practice made explicit by advocates of the third radiocarbon revolution is to actively exploit the very qualities of archaeological data that inspire epistemic pessimism: its incomplete, fragmentary nature. All the cases discussed turn on mobilizing multiple lines of evidence, and on scrutinizing these lines of evidence for their security and their independence from one another.. In this they are classic examples of "robustness reasoning" (Soler 2012; Chapman and Wylie 2016: 159-160).

3) *Experimental simulation*

Consider, finally, an example of archaeological simulation that makes use of legacy data but moves beyond the "gap compensation" approaches, as Currie would refer to them, that I have described in connection with strategies of secondary retrieval and recontextualization.

Gila Naquitz: The Early Mesoamerican Village (1986). This is a classic in the archaeological literature: an early modeling exercise undertaken by Kent Flannery and Robert Reynolds that makes use of a rich body of legacy data, alongside paleo-ecological datasets, to build a computational model of the evolution of the subsistence practices of a hypothetical microband that occupied a late Holocene cave site in the Oaxaca valley (8,700 to 6,600 BCE).¹⁴ The goal was to model how this band made the epoch-defining transition from "wide spectrum" foraging to agriculture, and the model itself was designed to allow Flannery and Reynolds to simulate the role of internal, social learning processes in this transition. Flannery's larger purpose was to assess the credibility of competing explanations for the development of agricultural practices, apparently independently, in many different locales around the world at roughly the same time (10,000-5,000 BCE). On his view, the accumulated body of archaeological and paleoecological evidence available to prehistorians by the mid-1980s had discredited accounts that emphasize a single exogenous forcing factor, like environmental stress and population pressure. He was intent on exploring the explanatory potential of accounts that posit more gradual processes, driven as much by internal cultural dynamics as by environmental factors.

As reported by Flannery and Reynolds (1986), the model they developed is an assemblage of subsidiary models of climate, ecology, subsistence strategy designed to be as realistic as possible. The repertoire of subsistence activities represented in the model was based on archaeological data that establish what resources were being exploited through the period when the cave at Gila Naquitz was occupied. The climate was modeled, based on paleoclimatic data, as randomly generating wet, dry, and average years. And the assignment of values to such variables as the availability, yield, labor requirements and dietary return for the dozen key sources of food that made up the hypothetical band's subsistence base was likewise informed by region-specific archaeological and paleoecological data. To probe otherwise inaccessible social aspects of the process, Flannery and Reynolds introduced subroutines that simulate information-sharing and decision-making operations by which the hypothetical foragers could learn from experience as they modified their repertoire of subsistence strategies and experimented with different resource collecting schedules.

With these components in place, Flannery and Reynolds modeled the transition to agriculture at Gila Naquitz in two stages; the first stage simulates the evolution of wide spectrum foraging, and the second models the emergence of incipient agriculture. When the simulation was run for foraging strategies alone it showed rapid improvement in efficiency until, after some 500 iterations, there was little improvement on the established pattern and positive feedback for change shifted to negative feedback that encouraged conservatism. At that point several incipient agricultural strategies were introduced to the repertoire—for example, clearing thorn forest to allow weedy plants to colonize (beans, and squash), and deliberately planting maize and squash seeds—and another learning process was initiated for the hypothetical band. This second-stage simulation showed a gradual shift in foraging strategies as the incipient agriculturalists incorporated the full suite of agricultural strategies documented archaeologically, reaching stable

¹⁴ I refer to these models as computational in the broad sense proposed by Weinberg (2013: 13).

performance in 550 iterations.

Flannery and Reynolds argue that the adequacy of their model should be evaluated in two ways. First, its representational adequacy should be assessed in terms of the correspondence of model outcomes with actual outcomes, as documented archaeologically. Here key measures of success were congruence in the relative emphasis on each plant species exploited and the time frame required for stabilization and, in the case of the model of incipient agriculture, the order in which changes in subsistence practice emerge. In addition, Flannery and Reynolds tested the robustness of the model as a simulation, manipulating model parameters and inputs related to social dynamics and environmental conditions. They disabled the information feedback loop and found that performance peaked early but then oscillated in a manner quite unlike anything suggested by the archaeological evidence. They also changed the environmental conditions and population density under which agricultural strategies were adopted and found that the random alternation of wet, dry, and average years is a crucial stimulus for the experimentation and learning processes that, in the simulation, gave rise to incipient agriculture. Under conditions of substantially greater climatic or population stress the hypothetical band proved to be more conservative, stalling the transition to farming, while under conditions of lower stress the band's subsistence strategies fluctuated without a directional intensification of horticultural practice of the kind observed archaeologically. This experimental component generated new data, in the form of information about the performance of the model, that bear on the assessment of competing hypotheses about the catalysts for and the tempo of the transition to agriculture that motivated the project as a whole.

This simulation functions both as a tool for investigating the archaeological subject, for which representational adequacy is key, and as an object of investigation in its own right, along lines described by Morgan for models in economics (2012). A program of computational modeling has since flourished in archaeology that likewise serves jointly representational and experimental purposes, now with an emphasis on agent-based modeling; contributors to *Model-Based Archaeology* demonstrate what can be learned by incorporating "cultural algorithms" (Kohler and van der Leeuw 2007: 89) into archaeologically and environmentally accurate models that then function as a platform for simulating the impact of various types of social organization, kinship systems, patterns of exchange and learning process on evolving settlement systems and land use practices. Although the aim of archaeological modelers is not primarily to mobilize legacy data, they make creative use of these resources, and they do this in a way that is not dependent on the extraction of new primary data or the introduction of new interpretive scaffolding that enables their reanalysis. As representational tools these simulations integrate widely diverse forms of archaeological data with a range of data from other sources that, together, tightly constrain the model with respect to those aspects of the past for which trace evidence survives. The experimental manipulation of the simulated cultural/ecological system leverages this representational adequacy to model mechanisms of cultural change that leave no direct trace evidence; it generates performance data that provide a basis for appraising hypotheses, at least as warranting further investigation, and that sometimes quite dramatically expand interpretive horizons. I suggest that the credibility of these model outcomes as evidence is a function of the capacity of the multiple independent lines of evidence deployed by archeological modelers to impose diffuse, mutually reinforcing constraints on these simulations; when they succeed it is not so much because they generate a new stream of evidence as that they are rich exercises in robustness reasoning.

IV. Methodological dynamism in the historical sciences

Ironically, it is precisely the shadowy nature of archaeological data—their unrealized potential—that makes possible the strategies I have described for putting them to work as evidence, sometimes long after they have been recovered and for purposes never envisioned by those who first identified, recorded and interpreted them. This is not to deny that the material traces with which archaeologists deal are often severely impoverished and suffer profound and continuous degradation. But these hardy survivors (Lucas 2015) can be a surprisingly rich evidential resource; time and again they give reason for an open-ended epistemic optimism.

The key to understanding how shadow data can be effectively mobilized as evidence is to recognize that the action is typically offstage; it is a matter of building and rebuilding the conceptual and technical scaffolding needed to effectively probe surviving material traces for data that have not yet been "brought to light." Increasingly it is also off-line; it involves simulating otherwise inaccessible dimensions of the cultural past. Both modes of practice depend on strategies by which archaeologists exploit the very fragmentation of their data that is sometimes cause for despair (Wylie 2011b). Insofar as these material traces are compositionally and causally complex they have the potential to sustain many different types of analysis that bring to bear different technical and conceptual scaffolding. It is the causal independence of (at least some of) the processes that converge in producing a trace, and the epistemic independence of the conceptual and technical scaffolding in terms of which these traces are constituted as evidence, that makes possible the strategies of secondary retrieval, recontextualization and simulation I have described. A preoccupation with the epistemic status of the traces themselves drives debate about the epistemic standing of historical sciences, directing attention away from the iterative process by which archaeologists (among other historical scientists) bootstrap-construct the scaffolding necessary to establish provisional evidential foundations for investigating the past. The stability and credibility of these provisional foundations is a function, not of locating epistemic bedrock, but of building ever more dense, reciprocally constraining tangles of evidence.

Acknowledgements

I am grateful to Sissel Schroeder for the invitation to comment on an archaeological discussion of uses of old evidence in 2003 that set in motion this line of inquiry, and to Mary Morgan and Sabina Leonelli for sharpening my appreciation of its broader significance. I also thank three anonymous referees for *HTSV* and Ed Hackett whose close reading from several different disciplinary vantage points gave me more to think about than I could do justice to in this short essay.

References

- Bayliss, A., & Whittle, A. (2015). Uncertain on Principle: Combining Lines of Archaeological Evidence to Create Chronologies. In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 213-242.
- Bell, M. (2015). Experimental Archaeology at the Crossroads: A Contribution to Interpretation or Evidence of 'Xeroxing'? In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 42-58.
- Boozer, A. L. (2015). The Tyranny of Typologies: Evidential Reasoning in Romano-Egyptian Domestic Archaeology. In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 92-109.
- Chang, H. (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.
- Chapman, R. & Wylie, A. (2016) *Evidential Reasoning in Archaeology*. London: Bloomsbury.
- Chapman, R. & Wylie, A. eds. (2015) *Material Evidence: Learning From Archaeological Practice*. London: Routledge.
- Chippindale, C. (2000). Capta and Data: On the True Nature of Archaeological Information. *American Antiquity*, 65(4), 605-612.
- Cleland, C. E. (2002). Methodological and Epistemic Differences between Historical Science and Experimental Science. *Philosophy of Science*, 69(3), 474-496.
- Currie, A. (2014). *Rock, Bone and Ruin: An Optimist's Guide to the Historical Sciences*. (Ph.D.), Australian National University.
- Currie, A. (2016a). Hot-blooded Gluttons: Dependency, Coherence and Method in the Historical Sciences. *British Journal of Philosophy of Science* 0: 1–24, DOI: 10.1093/bjps/axw005.
- Currie, A. (2016b). 'The Platypus' Tooth'. Manuscript available online: <https://www.dropbox.com/s/ly41va1ujoo4xk/1.%20The%20Platypus%27%20Tooth.pdf?dl=0>
- Eckardt, H., Chenery, C., Booth, P., Evans, J. A., Lamb, A., & Müldner, G. (2009). Oxygen and Strontium Isotope Evidence for Mobility in Roman Winchester. *Journal of Archaeological Science*, 36(12), 2816-2825.
- Flannery, K. V., & Reynolds, R. G. (1986). Simulating Foraging and Early Agriculture in Oaxaca. In K. V. Flannery (ed.), *Gila Naquitz: Archaic Foraging and Early Agriculture in Oaxaca, Mexico*, pp. 433-

508. New York: Academic Press.
- Gero, J. (2007). Honoring Ambiguity/Problematizing Certitude. *Journal of Archaeological Method and Theory*, 14(3), 311-327.
- Hanson, N. R. (1958). *Patterns of Discovery*. Cambridge: Cambridge University Press.
- Hodder, I. (1999). *The Archaeological Process: An Introduction*. Oxford: Blackwell Publishers.
- Killick, David (2015). Using Evidence from Natural Science in Archaeology. In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 159-172.
- Kohler, T. A. & van der Leeuw, S. E. eds. (2007). *The Model-Based Archaeology of Socionatural Systems*, NM: SAR Press.
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions* (2nd edition ed.). Chicago: University of Chicago Press.
- Leach, S., Lewis, M., Chenery, C., Müldner, G., & Eckhardt, H. (2009). Migration and Diversity in Roman Britain: A Multidisciplinary Approach to the Identification of Immigrants in Roman York, England. *American Journal of Physical Anthropology* 140(3): 546-561.
- Leonelli, S., & R. A. Ankeny (2015). Repertoires: How to Transform a Project into a Research Community. *BioScience* 65(7): 701-708.
- Lucas, G. (2015) Evidence of What? On the Possibilities of Archaeological Interpretation. In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 311-323. London: Routledge.
- Manning, S. W. (2015). Radiocarbon Dating and Archaeology: History, Progress and Present Status. In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 128-157.
- Mansilla, V. B., Lamont, M., & Sato, K. (2016) Shared Cognitive – Emotional – Interactional Platforms. *Science, Technology & Human Values* 41(4): 571-612.
- Morgan, M. S. (2012). *The World in the Model*. Cambridge: Cambridge University Press.
- Norton, J. D. (2003). A Material Theory of Induction. *Philosophy of Science* 70(4), 647-690.
- Norton, J. D. (2014). "A Material Dissolution of the Problem of Induction." *Synthese* 191(4): 671-690.
- Pollard, M., & Bray, P. (2015). The Archaeological Bazaar: Scientific Methods for Sale? In R. Chapman & A. Wylie (eds.), *Material Evidence*, pp. 113-127.
- Russell, B. (1921). *Analysis of Mind*. New York: Allen Unwin.
- Soler, L., Trizio, E., Nickles, T. & Wimsatt, W. C. (eds.) (2012). *Characterizing the Robustness of Science: After the Practice Turn in Philosophy of Science*. New York: Springer.
- Steel, D. (2001). Bayesian Statistics in Radiocarbon Calibration. *Philosophy of Science* 68 (Proceedings): S153-61.
- Taylor, W. W. (1948). *A Study of Archeology*. Carbondale Illinois: Southern Illinois University Press.
- Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.
- Turner, D. (2007). *Making Prehistory: Historical Science and the Scientific Realism Debate*. Cambridge: University of Cambridge Press.
- Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press.
- Wimsatt, W. C. (2014). Entrenchment and Scaffolding: An Architecture for a Theory of Cultural Change. In L. R. Caporael, J. R. Griesemer, & W. C. Wimsatt (eds.), *Developing Scaffolds in Evolution, Culture, and Cognition*, pp. 77-105. Cambridge MA: MIT Press.
- Wylie, A. (2002). *Thinking From Things: Essays in the Philosophy of Archaeology*. Berkeley CA: University of California Press.
- Wylie, A. (2008). Mapping Ignorance in Archaeology: The Advantages of Historical Hindsight. In R. N. Proctor & L. Schiebinger (eds.), *Agnotology: The Making and Unmaking of Ignorance*, pp. 183-205. Stanford: Stanford University Press.
- Wylie, A. (2011a). Archaeological Facts in Transit: The 'Eminent Mounds' of central North America. In P. Howlett & M. S. Morgan (eds.), *How Well Do Facts Travel? The Dissemination of Reliable Knowledge*, pp. 301-322. Cambridge: Cambridge University Press.
- Wylie, A. (2011b). Critical Distance: Stabilising Evidential Claims in Archaeology. In P. Dawid, W. Twining & M. Vasiliaki (eds.), *Evidence, Inference and Enquiry* (Vol. Proceedings of the British Academy 171), pp. 371-394. London: Oxford University Press.