

Radiocarbon Dating in Archaeology: Triangulation and Traceability



Alison Wylie

Abstract When radiocarbon dating techniques were applied to archaeological material in the 1950s they were hailed as a revolution. At last archaeologists could construct absolute chronologies anchored in temporal data backed by immutable laws of physics. This would make it possible to mobilize archaeological data across regions and time-periods on a global scale, rendering obsolete the local and relative chronologies on which archaeologists had long relied. As profound as the impact of ^{14}C dating has been, it has had a long and tortuous history now described as proceeding through three revolutions, each of which addresses distinct challenges of capturing, processing and packaging radiogenic data for use in resolving chronological puzzles with which archaeologists has long wrestled. In practice, mobilizing radiogenic data for archaeological use is a hard-won achievement; it involves multiple transformations that, at each step of the way, depend upon a diverse array of technical expertise and background knowledge. I focus on strategies of triangulation and traceability that establish the integrity of these data and their relevance as anchors for evidential reasoning in archaeology.

1 The Quest for an Absolute Chronology

If any data are “tragically local” (Latour 1999: 59), the fragmentary traces that make up the archaeological record would seem to fit the bill. Detached from the originating cultural events and contexts of interest to archaeologists, subject to the vagaries of preservation and the contingencies of recognition and recovery as a “record,” archaeological data are often seen as presenting insurmountable obstacles to their use as anchors for evidential claims about the past.¹ A major breakthrough, foundational to the formation of archaeology as a field, was development in the nineteenth century of methods for discerning the temporal structure of the material record

¹ See Currie’s reprise of and rebuttals to arguments that give rise to such pessimism (2018, chapter 4).

A. Wylie (✉)

Department of Philosophy, University of British Columbia, Vancouver, BC, Canada
e-mail: alison.wylie@ubc.ca

(Trigger 1989). The most influential of a number of chronological systems dating to this period – the “Three Age System” developed in the 1830s by Thomsen in Denmark and by Nilsson and Worsaae in Sweden – posited a cultural sequence of stone, bronze and iron ages based on the observation that, in undisturbed deposits, artifacts of these materials regularly occur together and in stratigraphic sequence (Trigger 1989: 124; Rowley-Conwy 2007: 32–47). Drawing on geological principles of superposition these assemblages were interpreted as chronological markers (Renfrew 1973: 24). To extend these sequences beyond the locales where they were established, archaeologists built fine-grained stylistic seriations that capture the orderly succession of form and design within classes of artefacts found in stratified deposits (e.g. Deetz and Dethlefsen 1967); artifacts of a similar material and design could be compared across sites and slotted into a design sequence presumed to hold for a cultural tradition or horizon. These attributions of “age” were, however, relative and of limited scope so, where possible, archaeologists made use of textual or epigraphic records to tie chronologies based on artifact typologies, seriation and stratigraphy to historically documented events and, thus, to one another. For example, coins and inscriptions testify to Roman presence in geographically distant regions, establishing (respectively) the earliest and latest dates at which the material associated with them could have been deposited. They also made use of dendrochronology and varve dating (annual sequences of tree-rings and glacial lake deposits) to anchor cultural to physical chronologies, but these too were of limited scope. The challenge was to link up chronologies of limited reach so that the trajectory of culture-transforming processes – the spread of farming, migrations and trade relations, the expansion and contraction of cultural spheres of influence – could be traced through space and time.

When radiocarbon dating was introduced in the early 1950s it was hailed as the solution to a range of chronological problems in archaeology; indeed, many expected that it would render obsolete these longstanding methods of constructing relative chronologies. The principle is straightforward. Radioactive carbon isotopes decay at a stable rate – their half-life is ~5730 years – so if you know the ratio of radioactive (^{14}C) to stable carbon (^{12}C and ^{13}C) in the environment in which a sample of organic material originated, you can use the difference between the proportion of radioactive carbon in the sample and this baseline ratio to estimate the time elapsed since “sample death” (Hamilton and Krus 2018: 198): the point at which the organic source of the sample stopped absorbing carbon and the decay process began. As Libby described the temporal data that could be captured by this means, the crucial warrant for its use as the anchor for an absolute chronology is the stability of the process of radioactive decay, a physical process that is not affected by other properties of the sample itself or its geological, much less its cultural, context.

The rate of disintegration of radioactive bodies is extraordinarily immutable, being independent of the nature of the chemical compound in which the radioactive body resides and of the temperature, pressure, and other physical characteristics of its environment. (Libby 1952: 9, as cited by Francis 2002: 297)

By contrast to the temporal data on which archaeologists had relied, this measurable ratio of time-dependent radioactive to stable carbon clearly seemed to qualify

as “mobile immutables” in Latour’s sense (Latour 1999; see also Morgan 2008, 2011). And, indeed, radiocarbon dating has had a profound impact on archaeology; in a recent retrospective Manning describes it as having “entirely restructured the practice and understanding of prehistoric archaeology around the world” (2015: 128). That said, the process of realizing its promise as a game-changing innovation has been a long, tortuous process. It is now described as proceeding through three radiocarbon revolutions (Manning 2015), each of which addresses distinct challenges posed by the multiple transformations involved in capturing, processing, packaging and interpreting radiogenic data for use in archaeological contexts.² In the process the Latourian ambitions that attended its initial introduction have been significantly rescaled. The radiogenic data made available by these successive revolutions is anything but “raw”; the ongoing process of refinement, calibration and interpretation affirms the robustly relational conception of data that frames this volume. I focus on two aspects of the transformations required to mobilize these data for archaeological purposes: the role of mediators, in the form of the inferential warrants and scaffolding of various kinds that make it possible to constitute material traces as temporal data; and the strategies archaeologists use to ensure the integrity of these data and, crucially, their credibility as anchors for evidential reasoning relevant to archaeological inquiry.

My aim here is to illustrate the irreducibly relational nature of data in this context where, at its inception, the radiocarbon revolution seemed poised to fulfill the most unqualified of foundationalist ambitions. I will identify a great many different kinds of objects and claims that function in archaeological contexts as data, extending an account I have developed elsewhere for a relational conception of evidence (Wylie 2011a; Chapman and Wylie 2016). On this view evidential claims are the terminus of practical arguments that, as characterized by Toulmin, originate with some “fact” or “datum” and are mediated by warrants that licence the inferential move from datum to conclusion (Toulmin 1958: 98, 218–221; Chapman and Wylie 2016: 34–36). Whether a claim counts as a mediating warrant or an evidential claim is a function of its role in an evidential argument; the material warrants that figure prominently in archaeological contexts are themselves the terminus of evidential arguments. The same is true of “data”; I concur with Leonelli that what counts as a “datum” is a function of its potential for use as evidence (Leonelli 2015, 2016) and that data are never simply “given”; they are themselves the terminus of an extended process of practice and inference that configure them as useable in a particular research context. In the case of radiocarbon data, literal journeys are involved; material traces are excavated, transported, curated, processed and incorporated into chronological models, and then put to work as archaeological evidence in a great many different contexts. But what I focus on here are primarily journeys across methodological frames. This account is itself a chronological; I chart the process by

²See Chapman and Wylie (2016: 147–151) for a more detailed account of the complex story of enthusiasm and ambivalence, institutional manoeuvring and competition that characterize the history of radiocarbon dating; we compare this with the reception and life history of other “external resources” imported to archaeology in recent decades.

which radically diverse types of expertise and bodies of background knowledge were brought together to refine the techniques and establish the standards that configure evolving practices for handling radiogenic data in archaeological contexts.

2 Capturing Radiogenic Data

Within a decade of the initial introduction of radiocarbon dating – the first radiocarbon revolution set in motion by Libby in the 1950s – it had become clear that its successful application to archaeological material would require a great deal of technical scaffolding. This first revolution is defined by two sets of issues: the need to establish archaeological field practices for recovering and handling samples that minimize contamination by younger or older organic material, and to refine the methods by which radiocarbon laboratories measure ^{14}C in archaeological samples.

In a Latourian analysis that provides a useful framework for considering the first of these challenges, Lucas characterizes archaeological fieldwork as an iterative process of intervention on field sites and materials – practices of disaggregation and assembly – by which an archive of material, both “found and made,” is assembled in ways that are configured by the anticipated requirements of data mobility. Invoking Latour (1999), he observes that “it is precisely what is portable or mobile that...defines the archive” (2012: 244). Field sites are, in a literal and documentary sense, standardized to approximate the material form of the archive, creating legible assemblages that can be “carried over” from the field to other sites of knowledge making (2012: 230, 234, 244). Lucas’s primary examples of these archive-producing processes are site survey, excavation and recording practices that are standardized within and across sites (at least, within research traditions), and designed to facilitate the translation of objects and observations “from one material form into another”: “the way we intervene with [a site] is set up precisely for the manner in which we [will] read it in translation” (2012: 238–239). So, for example, the practice of excavating in stratigraphic levels, cleaning exposed features and preparing the vertical walls of excavation units is keyed to photographic documentation, and to drawing plan views and stratigraphic profiles; the site is prepared, “sculpted,” so that it can be read “as if it were a drawing” or a photograph (2012: 239).

The archive in Lucas’ sense is, then, an active construct, designed to encode information about context and associations that will make it possible to retrace the steps by which the contents of the archive were produced, linking material samples and artefacts, drawings and photographs, field records and notes to one another and to features of their depositional context long after they have been removed, textually translated, and dispersed to distant museums, labs, offices and classrooms. Latour describes exactly this process in connection with the stratigraphic drawings created by the team of field biologists, ecologists, and soil scientists he observed in Brazil (1999: 57–58), and it figures in Bouman and Leonelli’s account of “data cleaning” – a practice documented within archaeology by Gero (2007). Traceability is crucial, especially when the field interventions are destructive, as in the case of excavation.

It is what makes possible the “iterative” analysis of an archaeological site that, on Lucas’ account, constitutes the mobility of the archive; it enables archaeologists to reassess, reposition and reinterpret the data that make up an archive, and sometimes to extract from it entirely new and unanticipated data (2012: 234; Wylie 2011b; 2016). Traceability is the key to establishing “sample-to-context” relationships that make the results of radiocarbon analysis useable in archaeological contexts – a fraught set of issues that have come into sharp focus in the last few decades and a point to which I return shortly.³ But first, consider in a bit more detail the translational processes by which radiocarbon data – the ratios of ^{14}C to ^{12}C and ^{13}C in archaeological samples – are generated.

Radiocarbon dating was initially applied to organic artefacts of known age held in well documented museum collections (e.g., Egyptian funerary furnishings). But as it became more widely available archaeologists reconfigured their field practices to anticipate the requirements of a new network of data-generating sites, specifically, radiocarbon dating laboratories. Material they had not routinely collected or that had been of marginal interest took on new significance – fragments of wood and bone, seeds and grains, the non-artefactual contents of storage and fire pits – and questions about sample collection, storage, and transport had to be addressed. As radiocarbon dating techniques evolved, the range of materials and the size of samples viable for dating changed, sometimes dramatically. With the use of Accelerator Mass Spectrometry (AMS) – based on direct detection of radiocarbon atoms – it is possible to work with samples as small as 20 milligrams, compared to the typical requirement of 10–100 grams for radiometric methods (Bronk Ramsey 2008: 258–259). At the same time the list of contaminants to be avoided has expanded from the obvious – cigarette ash and campfire charcoal – to include, for example, various types of glue, paper and cardboard that incorporate polyvinyl acetates; skin creams and lubricants in which polyethylene glycol is an ingredient; hydrocarbon-based fuels; and pesticides (biocides). Fieldworkers are advised to use glass or aluminium containers, but the specifics vary depending on type of sample, storage conditions and lab protocol; not surprisingly, given its greater precision, AMS dating is especially sensitive to contaminants.

Alongside the standardization of protocols for the recovery and handling of datable samples, radiocarbon laboratory techniques for processing them also had to be refined to control for a range of other confounds: the second set of issues that had to be resolved. These include, for example, the effects of electromagnetic impurities, ambient radiation, radon contamination and fractionation in reactions that do not go to completion, and the need to standardize count-time and conventions for calculating and reporting margins of error. By the early 1980s protocols ensuring inter-lab reliability had been instituted, but in a review of *Radiocarbon After Four Decades* (Taylor et al. 1992), Browman observed that, while “error magnitude is no longer linked clearly to lab type,” differences in the standards employed by different

³Shavit and Griesemer’s account of “locality in biodiversity surveys” (2009) illustrates just how complex traceability to “locality” can be, even when the original methodologies of data capture are not destructive (2009).

laboratories were still an issue (1994: 378). Fifteen years later Bronk Ramsey could report that “the measurement stage of the process is no longer the most critical element in determining precision and accuracy, except for the very smallest samples” (2008: 259), but problems persisted with the pre-treatment of samples. In short, fine-tuning laboratory protocols to ensure the reliable translation of radiocarbon samples from field to laboratory has been a long and ongoing process.

As these challenges were met, a growing number of anomalies were identified in the ^{14}C dates reported for archaeological material that could not be attributed to contamination or processing error. These drew attention to the complexity of the physical processes that radiocarbon dating exploits; much more background knowledge is required to estimate time elapsed since sample death than the “immutable” decay rate of radioactive carbon. In short, it was the interpretation of radiocarbon ratios as temporal data that came into sharp focus as needing attention. It was this recognition that set in motion the second radiocarbon revolution (Manning 2015: 129): a long process of calibrating radiocarbon dates that began in the mid-1960s.

3 Calibration: Refinement and Conversion

The second radiocarbon revolution was catalysed by two concerns: that, even if the half-life of radioactive carbon is stable, the ratio of ^{14}C to ^{12}C and ^{13}C in the atmosphere is not necessarily uniform over time or space; and that plants and animals take up carbon in different ways which affect its concentration in their tissues. Together these raise questions about what baseline should be used in determining how long the ^{14}C in a particular sample had been decaying. These were first addressed in connection with the “industrial” and “bomb” effects. By mid-century the widespread use of fossil fuels had dumped an enormous amount of “dead” carbon into the atmosphere, depressing the proportion of radioactive to stable carbon isotopes, while Cold War era nuclear bomb tests had “almost doubl[ed] the concentration of radiocarbon in the atmosphere” (Bronk Ramsey 2008: 251; Gillespie 1986: 20). In the event, the global standard, as “agreed internationally by the radiocarbon community,” was the average count rate for terrestrial wood dating to 1950, a choice of baseline described in the 1986 Oxford *Radiocarbon User’s Handbook* as “arbitrary”; “other values could have been used with perhaps more theoretical justification” (Gillespie 1986: 18).

Establishing a global convention for calculating elapsed radiocarbon years was just a beginning. What has ensued is a process of identifying and compensating for more localized effects of sample context and composition that has depended on recruiting an enormously wide range of substantive background knowledge about the “radiocarbon life cycle”: how carbon is produced, dispersed, and sequestered in diverse local environments, and how it is taken up and fixed by different types of organism (Bronk Ramsey 2008: 249–252). Where baseline carbon ratios are concerned, the complications now recognized include, for example, variation over time in the rates of radiocarbon carbon production in the upper atmosphere, which is an

effect of sunspot activity and dipole movement. This can have an impact on temperature which, in turn, affects the mixing and circulation of atmospheric ^{14}C as well as its rate of absorption into carbon reservoirs. The most significant reservoirs are marine; the rate at which radiocarbon is exchanged with the surface ocean is much slower than its dispersal in the atmosphere, and slower again in deep ocean reservoirs. A “marine offset” affects organisms sequestered in carbon sinks created by ocean currents where the proportion of radiocarbon may be considerably lower than in the atmosphere, the radioactive carbon in such an environment having decayed without being replenished. The atmospheric ratio also varies temporally and regionally. By the early 1980s it was recognized that there is a hemispheric difference in the concentration of ^{14}C , given proportionately more ocean surface in the southern than the northern hemisphere (Browman 1981: 249–67; Gillespie 1986: 26–7). In addition, as recently as 2001 two articles that appeared in *Science* reported that “a regional, time-varying ^{14}C offset can occur *within* a hemisphere” (Kromer et al. 2001; Manning et al. 2001; Reimer 2001). Wood samples from Anatolia and southern Germany, dated to the fifteenth through the seventeenth centuries AD on the basis of tree-ring sequences, had produced radiocarbon dates that diverged as much as 200 years. This discrepancy was attributed to a solar minimum that raised ^{14}C levels in the atmosphere, depressing radiocarbon relative to calendric ages, and an associated cooling effect that had seasonally different impact on trees with different growth periods (Kromer et al. 2001: 2529–30; Manning et al. 2001: 2533).

The challenges of determining baseline ratios of radiocarbon concentration for the environments in which organic samples originated is further complicated by an appreciation that processes of carbon uptake differ by type of organism. This has implications for how samples should be processed and how their measured carbon ratios should be interpreted. For example, plants that take up carbon directly from their environments have different concentrations of ^{14}C depending on whether they are terrestrial or marine, that is, whether they absorb carbon in the form of carbon dioxide or as bicarbonate. If they are terrestrial, uptake depends on the photosynthetic pathway by which they fix carbon, which differs between arid, succulent, and temperate zone plants. Radiocarbon concentrations also differ between herbivores that ingest photosynthesized carbon directly, and carnivores that get their carbon by “a more circuitous route through the food chain” (Bronk Ramsey 2008: 253). In addition, their metabolic processes may discriminate against heavy isotopes (e.g., in bone collagen) or affect the absorption of carbon by specific types of tissue (e.g., horns and nails do not continue to absorb carbon once formed).

Far from providing an autonomous and incontrovertible empirical foundation for archaeological chronologies, radiocarbon data are the conclusions of extended practical arguments that depend upon a great deal of contingent and, I will argue, local scaffolding. To be sure, the data that anchor these arguments are measurements of the carbon content of archaeological samples. However, as the process of second revolution calibration makes clear, they are only usable as a source of temporal data – that is, an estimate of time elapsed since sample death – given a complex of chain of inferences that take into account the conditions of their recovery, transport, storage, processing, and the technical details of radiometric or AMS analysis. The

inferential moves by which these measurements are converted into temporal data depend, in turn, on an immense array of mediating warrants: substantive background knowledge drawn from organic as well as physical chemistry, atmospheric science, geology, marine and terrestrial biology, to name just a few of the fields that were enlisted in the process of standardizing analytic procedures, controlling for confounds, and establishing computational and reporting conventions for radiocarbon data.⁴ I use the terminology of “warrants” in the sense suggested by Toulmin (1958), to refer to all the background knowledge and assumptions that license inferences from an originating observation or measurement, mark or inscription (a “datum” on his account), to a conclusion that, in this case, takes the form of a claim about the estimated time elapsed since “sample death”.

This emphasis on the substantive nature of these warrants resonates with Norton’s arguments for recognizing, more generally, that inductive inference is mediated by domain-specific “material postulates” (Norton 2003: 648). In a similar spirit Woodward insists that the assumptions “required to license...reliable inference from data [to phenomena]” are “empirical,” not “matters of stipulation” (2011: 172, 175). Alongside examples drawn from chemistry (determining the melting point of lead) and neuroscience (smoothing fMRI readings), he cites the assumptions on which archaeologists depend to infer temporal data (the date of a fossil) from radiocarbon decay counts: for example, “the way in which soil conditions and atmospheric exposure may affect the presence of carbon” (2011: 172). As he argues, it does not follow from the fact that such assumptions “go beyond the data” that they are “arbitrary, empirically unfounded, untestable, or matters of stipulation or convention” (2011: 173). The credibility of the data claim – that a given observation or measurement tracks a phenomenon of interest – depends upon the credibility of these mediating warrants. As Woodward also notes, the epistemic goals of inquiry and “attitudes toward risk” are also constitutive of these arguments (2011: 172, 174). So, for example, the claim that the radiocarbon content of an organic sample should be recognized as archaeological data depends not only on the background knowledge about confounds and offsets but also, prospectively, on its potential to serve as the point of departure for further inferences that support evidential claims about the age of a cultural feature, deposit, or site – the phenomena of interest to archaeologists.

For radiocarbon data to fulfil this function – to anchor a chronological claim that can serve as archaeological evidence – the crucial contribution of the second radiocarbon revolution has been the development of finely tuned calibration programs based on datasets that integrate the most sophisticated knowledge available about offsets and confounds of the sort I have described. To identify sources of error and correct for them archaeologists routinely rely on strategies of triangulation. They may, for example, compare carbon isotope ratios measured in material of archaeological interest against samples of “known age” that come from the same

⁴In “Circulating Reference” Latour remarks that “one science always hides behind another,” registering some disappointment that the Brazilian fieldwork he observed did not, in fact, represent “the birth of a science *ex nihilo*” (1999: 32). What I foreground here is this networked interdependence among fields that comprise the trading zone in which archaeology operates (Chapman and Wylie 2016: “Archaeology as a Trading Zone,” chapter 4).

(or relevantly similar) environments. The determination of “known age” may also depend on historical chronologies and on dendrochronology as well as, in some cases, stratigraphic sequences and typological chronologies – precisely the sources of temporal data that radiocarbon dating was meant to displace. The Southern German/Anatolian case mentioned above illustrates how this works in the case of dendrochronology. The annual accretion of tree-rings yields patterned sequences that can be stitched together across species and regions, so that radiocarbon dating of these samples can provide a temporal (usually decadal) profile of regional fluctuations in atmospheric radiocarbon. Varved lake sediments can support similar analyses that extend beyond the temporal reach of dendrochronological sequences. These local baseline data make it possible to refine the radiocarbon-based calculations of the time elapsed since sample death, but by no means do they resolve all the anomalies that signalled the need for calibration. At this point several ^{14}C calibration systems are available (e.g. CALIB 7.1, Stuiver et al. 2018; OxCal 4.3, Bronk Ramsey 2018). As these have been refined, “wiggles effects” have been identified such that, for some periods of archaeological interest, samples with different true ages correspond to the same radiocarbon ages, or the spread in their true ages is exaggerated, compressed, or even reversed. Here is Bronk Ramsey’s appraisal of the achievements and limitations of second revolution calibration techniques:

The problems of variable radiocarbon content in the atmosphere distort and defocus our view of the passage of time. The statistical methods now available to deal with calibrated dates act like a corrective lens to overcome these problems. However, with this clearer image other problems are also thrown into sharper focus: the statistical methods do not overcome methodological shortcomings in the radiocarbon method itself. (2008: 265)

The upshot is that, to use radiocarbon data as the basis for an “absolute” chronology – a temporal framework that, in the ideal, extends to the whole of the global archaeological archive – it has been necessary to rely on a system of warrants that effectively add contextual data back in and are valuable precisely because they are local and limited in their mobility. This predicament of locality – that secure anchoring to the local is a condition of mobility – is by no means unique to archaeology. Norton makes the point in general terms. The ‘portability’ of the material postulates that mediate inductive inference is invariably limited; they underwrite inference only within fields where the regularities and causal dynamics they capture can be shown to obtain (2003: 663).

4 Traceability and Triangulation

The catalyst for a third radiocarbon revolution, associated with a program of “Bayesian” chronological modelling (Bayliss and Whittle 2015),⁵ is the further realization that various forms of “tragically local” data (Latour 1999: 59) are indispensable not

⁵ Bayliss and Whittle describe this as a “pragmatic” Bayesian approach to archaeological problems (personal communication, 2014). Their central point, which resonates with Manning’s appraisal (2015), is that any assessment of the evidential bearing of radiocarbon data on questions about

only to ensure accuracy in the translation of radiogenic into temporal data, but also when it comes to transforming temporal into chronological data that can be used to address archaeological questions. The challenge here is to determine how a measure of time elapsed since the physical event of sample death relates the timing of cultural activities that are responsible for the production, use and deposition of the organic material from which samples are drawn. This is a problem that no amount of technical refinement – in standardizing sample collection and measurement practices, or in calibrating the translation of radiocarbon ratios into time scales – can resolve. As Manning describes this current and on-going revolution, it marks a decisive shift away from the quest for temporal data that approximate an ideal of absolute immutability and unconstrained mobility. Advocates of Bayesian approaches give up the epistemic ambitions that animated earlier revolutions; rather than expecting ^{14}C dating to deliver foundational, physics-backed temporal data that can displace reliance on context-specific resources, they embrace a commitment to “fully integrate archaeological information with ^{14}C dating,” including the “web” of background assumptions underlying relative chronologies (Manning 2015: 151). The emphasis in this third revolution is on integrating radiogenic data into chronological models that are archaeologically meaningful.

Whether the target of inquiry is an individual artefact or feature, a single short episode of use or occupation, a sequence of occupational layers in a densely stratified site, or a regional cultural formation that extends over millennia, the first step in the process of transforming temporal into chronological data is to assemble and appraise a set of ^{14}C dates that are potentially relevant to archaeological questions about age and chronological sequence. Traceability is crucial here. The determination of which samples to date when an archaeological archive is being created, and the choice of ^{14}C dates to include in a chronological model, depends on an appraisal of their provenance and integrity. Hamilton and Krus emphasize the need for a “holistic understanding” of the archaeological and geological context in which a sample originated that requires “at the very least...a description of the dated sample, the specific laboratory methods, and the sample’s provenience in relation to the archaeological features” (2018: 193). In the case of legacy data this appraisal sometimes involves quite literally retracing the steps of those who originally recovered a sample back into the field or to the repositories and laboratories to which finds and records were dispersed, reconstructing a record of the context from which it was drawn and the processes by which it was transformed into radiogenic data (Wylie 2011b: 312). Unless these data journeys can be reconstructed – unless the chains of recovery, transport, transformation, inscription, and relocation are “reversible,” as Latour puts it (1999: 61) – the samples have little value as a source of temporal data that can anchor archaeologically relevant evidential claims. Done well, this is a process of source criticism that exploits traceability as a means of making explicit

archaeological chronology must take into account how well supported a chronological model is on other grounds (its prior probability), as well as the degree to which these data support lines of evidential reasoning that are discriminating with respect to the plausibility and accuracy of the model (an appraisal of the prior and posterior likelihood of the data that anchors evidential claims).

and appraising the warrants that underpin attributions of integrity to individual samples and trustworthiness to the data claims based on them (Wylie 2011b).

In addition to traceability, triangulation provides a further check on accuracy and allows for a closer specification of the date ranges generated by ^{14}C analysis. For example, when archaeologists aggregate ^{14}C dates, rather than just calculating a mean or median date for the data set, they sometimes model the range of dates a hypothetical sample would generate, given standard margins of error, if the actual date of sample death was this calculated mean, a strategy of internal triangulation that can delimit the dispersion of pooled or averaged dates (Chapman and Wylie 2016: 152; Shott 1992: 221–223). Typically, however, triangulation strategies make use of radiogenic data drawn from different sources to cross-check the accuracy of individual ^{14}C dates and the credibility of the assumptions that inform the construction of chronological models. This may involve testing multiple samples from a single artefact or feature, sometimes submitting them to different laboratories, to control for contamination and laboratory error, or testing different types of samples drawn from a single depositional context to control for biases that can arise from relying on one type of material. It may also involve dating non-cultural, botanical and ecological samples that originated in the same environment as cultural samples in order to cross-check assumptions about baseline carbon ratios (Hamilton and Krus 2018: 195). Latour seems to have such strategies in mind when he mentions, in passing, a field practice of cross-field triangulation whereby the geomorphologist on the Brazilian field crew “adds her two cents to all the conversations, allowing her expatriate colleagues to ‘triangulate’ their judgments through hers” (1999: 47). Here credibility is a function of the capacity of these different types of radiogenic data to constrain one another, exposing sources of error that may not be identifiable by tracing data journeys and assessing the security of warrants for individual (calibrated) ^{14}C dates.

More expansive strategies of triangulation typical of this third ^{14}C dating revolution depend on mobilizing a range of different, non-radiogenic types of temporal data. Given practices of reuse, curation, trade and other forms of circulation that complicate the life histories of organic material in cultural contexts, datable samples may come from organisms that were cut, butchered, burned or otherwise taken out of the carbon cycle long before they were deposited in the archaeological contexts from which they are recovered. To establish a connection between the ^{14}C -datable natural event of their death and the cultural target of interest to archaeologists, a premium is put on drawing samples from organic remains that can be assumed to be “functionally related to their deposit” (Hamilton and Krus 2018: 194), to have originated in a short timeframe, or to derive from a temporally ordered sequence of deposits. Articulated animal bone or undisturbed human burials are examples of the former; geologically sealed cave deposits and the association of human remains or artefacts with extinct mega-fauna are a classic example of the latter (Chapman and Wylie 2016: 35), as are stratigraphic associations more generally: the location of a sample in relation to stratified occupational levels may set temporal bounds on its age in relation to other datable samples. The stylistic homogeneity of the artefact assemblages with which a sample is associated, and comparanda from related sites that support the seriation of particular types of artefact or feature, can also be used to establish contemporaneity or temporal sequence (Chapman and Wylie 2016: 151–155).

The point of recruiting these diverse types of data is to re-embed the much-manipulated ^{14}C mobiles in a local context of inquiry, delimiting the range of physically possible dates and margins of error generated by radiocarbon analysis and, crucially, integrating discrete traces, features, and sites into an archaeologically plausible chronology. As the number of distinct types of data built into these models is expanded, so too is the range of background knowledge – the substantive warrants – that are required to secure the inferences that link the temporal claims they support to an archaeological target. This vastly complicates the construction of chronological models, but it is also a source of epistemic credibility. The principle at work here is that the likelihood of spurious convergence on a specified range or sequence of dates is much reduced when mediating warrants are drawn from diverse sources and the material they configure as data are themselves generated by different causal processes. The credibility of the resulting chronological models is not just a function of the aggregation of individual data points or sets assessed as trustworthy; it arises from the collective capacity of these data to reinforce and to constrain one another.

5 Robustness Reasoning About Temporal Data

The strategies central to these practices of chronological modelling are recognizably a genre of “robustness” reasoning, as Wimsatt has described the diverse methods of “multiple determination” that he finds ubiquitous across the sciences (Wimsatt 1981: 123–4, 2012; Soler 2012: 3). In this case they are applied to the kind of problem Hacking explores in connection with microscopy (Hacking 1981, 1983: 186–203). They are meant to ensure that the heavily scaffolded temporal data archaeologists rely on do, in fact, track the cultural phenomena of interest, counteracting the risk that they are artefacts of, or otherwise distorted by, practices of extraction and measurement, processing and packaging for travel as elements of an archaeological archive. I have argued elsewhere that, in constructing evidential claims, archaeologists routinely exploit the causal and epistemic independence between distinct lines of evidence that originate in a common target of inquiry (Wylie 2011a: 387–389). The strategies of triangulation characteristic of the second and third radiocarbon revolution suggest that this is true, as well, of ^{14}C data. To use a metaphor of Norton’s (2014: 673), the empirical objects and claims that comprise the data recruited in support of various components of a chronological model are reciprocally strengthened by being bound into “highly connected, massively tangled” and self-stabilizing systems of data-cum-evidence.

Strategies of multiple determination, coupled with traceability, can certainly mitigate the risk that convergence is spurious when diverse types of data and the evidential claims they anchor come together in support of a coherent chronological model. Nonetheless the worry remains that, absent “absolutes” in the form of immutable temporal data that can function as a decisive, wholly autonomous arbiter of chronological questions, there is an inherent nepotism in the process of mutual adjustment required to calibrate temporal data and integrate them into archaeological chronologies. The strategies for identifying, controlling and correcting for error

developed in the course of successive radiocarbon revolutions suggest four conditions that data-evidence tangles must meet if the risk of vicious, rather than virtuous, patterns of self-stabilization is to be avoided.

The first condition is a requirement that the source data and the warrants backing data claims be “secure”: each, taken on its own, must be well substantiated in terms of the best technical and theoretical expertise available in the fields that make possible their capture and mobilization. This was the central preoccupation of the first radiocarbon revolution in which techniques for reliably measuring radiocarbon ratios in organic samples were the focus of attention. It figures, as well, in the long process of calibrating radiogenic data against a much expanded range of background knowledge about the nature of the samples, their contexts of origin, possible confounds that affect the measurement of carbon ratios, and their translation into the temporal scale of elapsed calendar years.

The second condition is a requirement of causal and conceptual independence between the various types of temporal data that are used to calibrate one another and to build chronological models. In the ideal, any given tangle of interlinked chains of data-cum-evidence should incorporate data that have causally distinct “life histories,” and the warrants mediating the various transformations these data and their use as evidence should derive from conceptually independent research traditions. At their most effective, the triangulation strategies that figure prominently in the second and third radiocarbon revolution meet exactly this requirement.

By extension of this second condition, when one type of data is used to calibrate another, the tuning of measurement systems and the refinement of the warrants that underpin them should be justified on substantive grounds, not just because they ensure convergence. Manning describes several cases in which this was a central consideration in the process of reconciling early Cycladic and late Bronze Age Aegean chronologies with sequences of radiocarbon dates (2015: 142–150), as do Bayliss and Whittle with reference to chronological models of artefact and occupational sequences of different scales (2015: 222–230). A striking example of such reasoning recently analysed by Bokulich ([forthcoming](#)) is the decision, in 2012, to base a significant revision of the Geological Time Scale on an independent, non-radiometric measure of geologic time – a choice explicitly informed by a concern to preserve the independence of the two radiometric methods that are typically used to cross-check one another in geochronological dating.

The trajectory of the multiple radiocarbon revolutions makes it clear that traceability as well as triangulation is required. The usefulness of ^{14}C data depends on their mutability, which means that they are vulnerable to error and distortion in the course of their journeys. Detailed documentation and ongoing critical scrutiny of the transformations that comprise these journeys is crucial and, in fact, an explicit demand for traceability is a recurrent theme in the literature on chronological modelling. Hamilton and Krus emphasize the need for “transparency” with respect to model structure and the “choices and assumptions” that inform its construction (2018: 195); the hypothesized relationship between sample dates and the dates of a target event should be clearly specified, and the basis for these assumptions – background knowledge about the archaeological context and mediating warrants – should be made explicit.

A final condition might be described as a requirement of epistemic democratization: a normative implication of the relational account of data. Assessments of security and strategies of triangulation can justify privileging some types of temporal data over others as inherently more trustworthy, accurate and/or precise. However, none should be presumed to be empirically foundational “immutable,” exempt from re-examination when discrepancies arise in data-evidence tangles, or when the process of retracing data journeys brings to light previously unrecognized confounds or as yet unaddressed uncertainties. This is the central motivation for the third radiocarbon revolution: that however compelling their physics-backing may be, ^{14}C data must be accountable not only to standards of credibility in their field of origin but also those that are specific to their contexts of use. This norm underwrites a commitment to treat even the most promising “silver bullet” techniques of data extraction and mobilization as tentative, the starting point for a process of epistemic iteration in which it is expected that they will be subject to continuous refinement and sometimes replacement as an evolving empirical scaffolding for inquiry (Chang 2004: 43).

These are demanding ideals, rarely fully met in practice. Nonetheless, I suggest that they are orienting norms of practice exemplified by, and responsible for, the considerable achievements of the successive radiocarbon revolutions and that have unfolded since the 1950s.

6 Conclusion

What exactly are the data in this sprawling story of extraction, processing and packaging, calibration and circulation by which radiogenic data are captured and integrated into chronological models in archaeological contexts? There are the organic artefacts and residues that survive *in situ* or are curated in the archaeological archive from which datable samples are retrieved; the carbon extracted from these samples; the isotope ratios produced by means of AMS or decay counts; the calibrated estimates of radiocarbon years elapsed since sample death; the translation of these radiocarbon dates into calendar years; and then their interpretation as dates when organic elements of the archaeological archive were created, used, and deposited. All of these constitute the “data,” now repeatedly transformed, that figure as the starting point for the evidential reasoning that grounds cultural/historical chronologies. There are also all the ancillary data that back the substantive assumptions – the warrants – that mediate each step in these tangled chains of reasoning from and about the material samples, test results, and records that comprise the archaeological archive.

I submit that all of these count as data. Their status as data is a function of their role in anchoring practical arguments for a range of different types of evidential claim, not an intrinsic quality of ‘givenness’, closeness or similarity to the target of inquiry, much less their status as self-warranting or empirically foundational. The framing argument for recognizing that data are relational in this sense has been made by Leonelli (2015: 817, 2016: 79), and the recognition that they are as hard-won an achievement as the evidential claims they support figures centrally in the philosophical and science studies sources on which I have drawn, diverse as they are. It is also a recurrent theme in internal discussion of the vagaries of archaeologi-

cal inquiry. In addition to Lucas' account of "the archaeological record," Chippindale urges his fellow archaeologists to adopt the term "capta" rather than "data" (2000), a sentiment that resonates with Latour's admonition that one should "never speak of 'data' – what is given – but rather of 'sublata', that is, of achievements" (1999: 42).

An implication of this relational view is that the data that anchor evidential arguments are themselves the terminus of further practical arguments that depend upon their own warrants; as such, their points of origin, and each of the steps involved in capturing and transforming them into useable data are also subject to critical scrutiny, and open to demands for further backing. In the case of archaeology, building these tangles of practical argument is an achievement that depends on a genre of robustness reasoning; it is a matter of enlisting not only the data generated by physical dating techniques but also a wide range of less transportable, context-specific data. The epistemic integrity and credibility of the resulting temporal data is a function of the traceability of these transformations, a point that Latour acknowledges when he considers their "reversibility" (1999: 59, 74), not the immutability of these mobiles that he otherwise emphasizes. Bronk Ramsey captures this point when he observes that, as radiocarbon dating has become "markedly more precise (and hopefully not less accurate) we need to be even more careful...about the chain of reasoning that allows us to go from a radiocarbon measurement to an understanding of chronology" (2008: 266):

Radiocarbon dating should not be viewed as a black box, which occasionally has to be shaken because it does not give the right answer. (Bronk Ramsey 2008: 270)

References

- Bayliss, Alex, and Alasdair Whittle. 2015. Uncertain on Principle: Combining Lines of Archaeological Evidence to Create Chronologies. In *Material Evidence: Learning from Archaeological Practice*, ed. R. Chapman and A. Wylie, 213–242. London: Routledge.
- Bokulich, A. forthcoming. Calibration, Coherence, and Consilience in Radiometric Measures of Geologic Time. *Philosophy of Science*.
- Boumans, Marcel, and Sabina Leonelli. this volume. From Dirty Data to Tidy Facts: Clustering Practices in Plant Phenomics and Business Cycle Analysis. In *Data Journeys in the Sciences*, ed. Sabina Leonelli and Niccolò Tempini. Cham: Springer.
- Bronk Ramsey, C. 2008. Radiocarbon Dating: Revolutions in Understanding. *Archaeometry* 50: 249–275.
- . 2018. *OxCal version 4.3*. Oxford Radiocarbon Accelerator Unit. https://c14.arch.ox.ac.uk/oxcalhelp/hlp_contents.html. Accessed 10 May 2018.
- Browman, David L. 1981. Isotopic Discrimination in Correction Factors in Radiocarbon Dating. *Advances in Archaeological Method and Theory* 4: 241–295.
- . 1994. Review of *Radiocarbon After Four Decades*, eds. R. E. Taylor, A. Long, and R. S. Kra, and *Proceedings of the 14th International Radiocarbon Conference*, ed. A. Long. *American Antiquity* 59: 377–378.
- Chang, Hasok. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.
- Chapman, Robert, and Alison Wylie. 2016. *Evidential Reasoning in Archaeology*. London: Bloomsbury.
- Chippindale, Christopher. 2000. Capta and Data. *American Antiquity* 65: 605–612.
- Currie, Adrian. 2018. *Rock, Bone and Ruin*. Cambridge, MA: MIT Press.

- Deetz, James F., and Edwin S. Dethlefsen. 1967. Death's Head, Cherub, Urn and Eillow. *Natural History* 76: 29–37.
- Francis, Kevin. 2002. "Death Enveloped All Nature in a Shroud": The Extinction of Pleistocene Mammals and the Persistence of Scientific Generalists. PhD, Program in History of Science and Technology, University of Minnesota.
- Gero, Joan. 2007. Honoring Ambiguity/Problematizing Certitude. *Journal of Archaeological Method and Theory* 14: 311–327.
- Gillespie, Richard. 1986. *Radiocarbon User's Handbook*, Monograph # 3. Oxford: Oxford University Committee for Archaeology.
- Hacking, Ian. 1981. Do We See Through a Microscope? *Pacific Philosophical Quarterly* 18: 305–322.
- . 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.
- . 1992. The Self-Vindication of the Laboratory Sciences. In *Science as Practice and Culture*, ed. A. Pickering, 29–64. Chicago: University of Chicago Press.
- Hamilton, Derek W., and Anthony M. Krus. 2018. The Myths and Realities of Bayesian Chronological Modeling Revealed. *American Antiquity* 83: 187–203.
- Kromer, Bernd, Sturt W. Manning, Peter Ian Kuniholm, Maryanne W. Newton, Marco Spurk, and Ingeborg Levin. 2001. Regional 14CO₂ Offsets in the Oroposphere: Magnitude, Mechanisms, and Consequences. *Science* 294: 2529–2532.
- Latour, Bruno. 1999. Circulating Reference: Sampling the Soil in the Amazon Forest. In *Pandora's Hope*, 24–79. Cambridge, MA: Harvard University Press.
- Leonelli, Sabina. 2015. What Counts as Scientific Data? A Relational Framework. *Philosophy of Science* 82: 810–821.
- . 2016. *Data-Centric Biology: A Philosophical Study*. Chicago: University of Chicago Press.
- Libby, Willard. 1952. *Radiocarbon Dating*. Chicago: University of Chicago Press.
- Lucas, Gavin. 2012. *Understanding the Archaeological Record*. Cambridge: Cambridge University Press.
- Manning, Sturt W. 2015. Radiocarbon Dating and Archaeology: History, Progress and Present Status. In *Material Evidence*, ed. R. Chapman and A. Wylie, 113–127. London: Routledge.
- Manning, Sturt W., Bernd Kromer, Peter Ian Kuniholm, and Maryanne W. Newton. 2001. Anatolian Tree Rings and a New Chronology for the East Mediterranean Bronze-Iron Ages. *Science* 294: 2532–2535.
- Morgan, Mary S. 2008. 'On a Mission' with Mutable Mobiles. LSE Working Papers on the Nature of Evidence No 34/08.
- . 2011. Traveling Facts. In *How Well Do Facts Travel? The Dissemination of Reliable Knowledge*, ed. P. Howlett and M.S. Morgan, 3–41. Cambridge: Cambridge University Press.
- Norton, John D. 2003. A Material Theory of Induction. *Philosophy of Science* 70: 647–670.
- . 2014. A Material Dissolution of the Problem of Induction. *Synthese* 191: 671–690.
- Reimer, Paula J. 2001. A New Twist in the Radiocarbon Tale. *Science* 294: 2494–2495.
- Renfrew, Colin. 1973. *Before Civilization: The Radio Carbon Revolution and Prehistoric Europe*. New York: Penguin Books.
- Rowley-Conwy, Peter. 2007. *From Genesis to Prehistory: The Archaeological Three Age System and its Contested Reception in Denmark, Britain, and Ireland*. Oxford: Oxford University Press.
- Shavit, Ayelet, and James Griesemer. 2009. There and Back Again, or the Problem of Locality in Biodiversity Studies. *Philosophy of Science* 76: 273–294.
- Shott, Michael J. 1992. Radiocarbon Dating as a Probabilistic Technique: The Childers Site and Late Woodland Occupation in the Ohio Valley. *American Antiquity* 57: 202–230.
- Soler, L  na. 2012. The Solidity of Scientific Achievements: Structure of the Problem, Difficulties, Philosophical Implications. In *Characterizing the Robustness of Science: After the Practice Turn in Philosophy of Science*, ed. L. Soler, E. Trizio, T. Nickles, and W.C. Wimsatt, 1–60. New York: Springer.
- Stuiver, M., P. J. Reimer, and R. Reimer. 2018. *CALIB Radiocarbon Calibration Version 7.1*. Accessed 10 May 2018. <http://calib.org/calib/>.

- Taylor, R.E., A. Long, and R.S. Kra, eds. 1992. *Radiocarbon After Four Decades: An Interdisciplinary Perspective*. New York: Springer.
- Toulmin, Stephen E. 1958. *The Uses of Argument*. Cambridge: Cambridge University Press.
- Trigger, Bruce G. 1989. *A History of Archaeological Thought*. Cambridge: Cambridge University Press.
- Wimsatt, William C. 1981. Robustness, Reliability, and Overdetermination. In *Scientific Inquiry and the Social Sciences*, ed. M.B. Brewer and B.E. Collins, 124–163. San Francisco: Josey-Bass.
- . 2012. Robustness: Material and Inferential, in the Natural and Human Sciences. In *Characterizing the Robustness of Science: After the Practice Turn in Philosophy of Science*, ed. L. Soler, E. Trizio, T. Nickles, and W.C. Wimsatt, 89–104. New York: Springer.
- Woodward, James F. 2011. Data and Phenomena: A Restatement and Defense. *Synthese* 182: 165–179.
- Wylie, Alison. 2011a. Archaeological Facts in Transit: The “Eminent Mounds” of Central North America. In *How Well Do Facts Travel? The Dissemination of Reliable Knowledge*, ed. P. Howlett and M.S. Morgan, 301–322. Cambridge: Cambridge University Press.
- . 2011b. Critical Distance: Stabilising Evidential Claims in Archaeology. In *Evidence, Inference and Enquiry*, ed. P. Dawid, W. Twining, and M. Vasiliaki, 371–394. London: Oxford University Press.
- . 2016. How Archaeological Evidence Bites Back: Strategies for Putting Old Evidence to Work in New Ways. *Science, Technology & Human Values* 42: 203–225.

Alison Wylie is Professor of Philosophy and Canada Research Chair in Philosophy of the Social and Historical Sciences at the University of British Columbia. Her philosophical analyses are case-based, chiefly concerned with archaeological practice and feminist research in the social sciences, and address such questions as follows: What counts as evidence? How should we understand ideals of objectivity given the role of values and interests in inquiry? How do we make research accountable to the diverse communities it affects? Her recent publications include *Evidential Reasoning in Archaeology* (2016) and *Material Evidence* (2015), with archaeologist Bob Chapman; “How Archaeological Evidence Bites Back: Strategies for Putting Old Data to Work in New Ways” (STHV 2017); and contributions to the *Springer Handbook of Model-Based Science* (2017), *Objectivity in Science* (2015), *How Well do Facts Travel? (2010) and Agnotology* (2008). Her work on feminist standpoint theory includes “What Knowers Know Well” (*Scientiae Studia*, 2017) and her 2012 APA Presidential Address, and her essays on research accountability appear in “A Global Dialogue on Collaborative Archaeology” (*Archaeologies*, 2019), *The Ethics of Cultural Appropriation* (2009) and *Embedding Ethics* (2005). She is a Past President of the American Philosophical Association (Pacific Division) and Current President of the Philosophy of Science Association.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

