Archaeology and archaeometallurgy: some unresolved areas in the interpretation of analytical data

Mark Pearce

To cite this article: Mark Pearce (2016) Archaeology and archaeometallurgy: some unresolved areas in the interpretation of analytical data, STAR: Science & Technology of Archaeological Research, 2:1, 46-53

To link to this article: http://dx.doi.org/10.1080/20548923.2016.1160593

© 2016 The Author(s). Published by Taylor & Francis

Published online: 06 Apr 2016.
Archaeology and archaeometallurgy: some unresolved areas in the interpretation of analytical data

Mark Pearce*

Department of Archaeology, University of Nottingham, Nottingham NG7 2RD, United Kingdom; tel.: ☎+44 (0)115 951 4839

Abstract
This paper uses examples from Mediterranean and in particular Italian prehistory to explore the interface between prehistoric archaeology and metals analysis by examining three areas: the usefulness of data from past analyses (‘what is it made of?’), lead isotope analysis and the problem of unpublished data (‘where is it from?’), and the interpretation of analytical data (‘what does it mean?’). Issues discussed include big data, the integration of datasets from different analytical programmes (especially where analytical results are in disagreement), and open access and the withholding of data through incomplete publication, which means that conclusions cannot be verified. It offers some suggestions as to how communication between archaeologists and archaeometallurgists can be improved.

Keywords
Big data; open access; composition; provenance; interpretation; lead isotopes; Italy

Received 24 April 2015; accepted 22 January 2016

Statement of significance
Investigation is science when it asks specific questions and when its conclusions can be repeated and verified. This paper applies the important themes of big data (the integration and mining of disparate datasets) and open access (the free availability of scientific data) to the study of ancient metallurgy. It argues that communication between archaeologists and archaeometallurgists needs to be improved, and offers practical suggestions as to how that may be achieved, in order to integrate analytical findings more effectively into archaeological discourse.

Data availability
All data used in this article are fully published and referenced.

Introduction
This brief discussion paper looks at the interface between prehistoric archaeology and metals analysis, from the point of view of Mediterranean and in particular Italian prehistory, by examining three areas where problems in the interpretation of analytical results are unresolved. These comprise the usefulness of data from past analyses (‘what is it made of?’), lead isotope analysis and the problem of unpublished data (‘where is it from?’), and the interpretation of analytical data (‘what does it mean?’).

My approach is that of a prehistoric archaeologist rather than that of a materials scientist, but I make no apologies for that, as analytical data from materials science is only of use to archaeology if it answers specific questions, and as will become clear, I believe that there is a serious deficit in the use of analytical data by prehistoric archaeologists.

Using data from past analytical programmes
Compositional data has been generated for prehistoric artefacts since at least the end of the eighteenth century (Pollard 2013), and ‘what is it made of?’ was the first question that was asked of materials scientists (e.g. in Britain, Pearson 1796). Compositional data is typically used for two purposes, to infer provenance and to track technological innovation. In an evolutionary paradigm where prehistoric copper metallurgy ‘progressed’ from the working of native copper to smelting and then from arsenical copper to tin bronze (a paradigm which may now be questioned, at least for the Americas – Lechtman 1996), the question of an artefact’s composition was long perceived to have great heuristic value, not least as an indication of relative chronology. However as the accuracy and...
precision of analytical techniques has improved through time, and the range of component elements detected has widened from major to minor, rare earth and trace elements, a large number of data of varying reliability have been produced. Moreover, and it would be invidious to cite particular examples, especially in Mediterranean archaeology many analyses have been carried out using inappropriate techniques (perhaps because a specialist in the local University offered their services for free) and without necessarily a clear research question; indeed the lack of a clear research question is perhaps often the reason why inappropriate techniques are chosen. In the specific case of Italian prehistory, analytical data are often published in conference proceedings, local journals and other publications which may rarely circulate outside Italy or indeed even the region in which they were produced.

The amount of data available should not be underestimated: for example, the Stuttgart optical emission spectroscopy analytical programme published 22,000 determinations (Junghans, Sangmeister and Schröder 1960; 1968; 1974), and 801 of these relate to artefacts with a provenance in present-day Italy (Pearce 1998, 51). Such a large dataset, for which most of the data was generated using the same analytical techniques and equipment and is therefore presumably internally consistent, allows large-scale statistical investigation. The Stuttgart team carried out univariate statistical analysis on five of the 11 elements determined (Arsenic, Antimony, Silver, Nickel and Bismuth - Junghans, Klein and Scheufele 1954; Junghans, Sangmeister and Schröder 1960, 57–90). However we might doubt the metallurgical groups which they identified. One of the elements included in the statistical analysis was Arsenic, which is potentially the result of alloying or, at the very least, the result of deliberate selection of ores for particular artefact classes (Pearce 2007, 84–86) and which can be lost through oxidation when an artefact is re-melted so that Arsenic content can be used as a proxy to detect recycling (Bray and Pollard 2012; Pollard, Bray and Gosden 2014). Another was Bismuth, which (like Lead) segregates during solidification and so will vary in different parts of an artefact (Slater and Charles 1970). The Stuttgart group included the results of other analytical programmes in their statistical analysis: the 1374 spectroscopic analyses carried out from 1931 by Otto and Witter (1952), plus 98 analyses of artefacts from Britain and Ireland (Coghlan and Case 1958), 37 from France (Briard and Giot 1956) and 21 from Slovakia (Novotná 1955). More recently, Rüdiger Krause (2003; cf. Pernicka 1995, 79–99) has reassessed the compositional groups proposed by the Stuttgart team, and his cluster analysis suggests that they are largely valid. The challenge is however the integration of this data with that from other analytical programmes, such as (to stick with my Italian focus) Barker and Slater’s dataset from metalwork in the Rome Pigorini Museum, amounting to 106 analyses (Barker 1971; Slater 1971). This latter dataset was produced by atomic absorption spectroscopy and arc emission spectroscopy and determined the same range of elements as the Stuttgart programme, so integration should arguably be relatively easy (but cf. Table 1).

However, a vast range of analytical techniques have been applied to Italian prehistoric metalwork, ranging from wet chemistry to atomic absorption spectrometry and instrumental neutron activation analysis (e.g. Berzero et al. 1991).

Anyone who has tried to collect and simply collate such data will know that there are a number of macroscopic problems which need resolution, such as the different ranges of elements determined and the limits of detection of the technique and equipment used, but perhaps the most obvious problem arises where the same artefact has been analysed more than once and the results are not in agreement. Tables 1 and 2 (Pearce 1994, 54–55, tabs 7 & 9) provide an example: according to the Stuttgart analyses, the six axes from the Pieve Albignola hoard conserved at Rome have an Arsenic content of between 0.15% and 0.02% (average 0.78%), whereas Slater did not detect the element (Table 1); the Stuttgart programme analysed just one axe from the same hoard conserved at Pavia and found it to have an Arsenic content of 0.03% but neutron activation analysis by Berzero et al. (1991, n.31) gave a determination of 0.002971% (Table 2). In all, Berzero and colleagues (1991) analysed 26 of the axes from the Pieve Albignola hoard conserved at Pavia, and their Arsenic content was found to vary between 0.003361% and 0.001371% (average 0.002209%; Figure 1). As we have seen, Arsenic was one of the elements used by the Stuttgart team to establish their compositional groups and can be used as a proxy for recycling, but the discrepancy between the analytical programmes is so great (an order of magnitude) as to throw any considerations based on its content in these axes into doubt. The analyses by Berzero et al. also suggest that Slater’s rather than the Stuttgart figures for Arsenic are correct, despite Pernicka’s (1995, 85, Abb.33) dismissal of them.

Müller and Pernicka (2009) acknowledged that there are differences between the results obtained on Iberian material by different analytical programmes, especially as regards Silver and Antimony, but argued that they were broadly comparable and that there was no effect on the compositional groups identified. A number of explanations for such discrepancies may be adduced, such as analytical error, but the doubt remains that they may be due to compositional variation within individual artefacts. Compositional variation was acknowledged by Junghans, Klein and Scheufele (1954, 102, Abb.18, Tab.4) but they dismissed its impact on their compositional groups. It is however worth noting that most of the copper-based metalwork analysed in Europe relates to early periods of metalworking when we may imagine that metallurgy was relatively unsophisticated. If compositional variation within an artefact is greater than
the range of experimental error then compositional data may be of limited value, however reliable the analytical techniques used (indeed it should be stressed that the more precise such techniques are, the greater the potential impact of compositional variability).

Table 1. Comparison between two different programmes of spectroscopic analysis, for six early Bronze Age flanged axes from Pieve Albignola (PV, Italy) at Rome, Pigorini Museum (Pearce 1994, tab.9). Data is reported as published: SAM = Junghans, Sangmeister and Schröder 1974, 318–9, n.20422-20427; Slater 1971, n.4i, 4j, 7c-g; n.d. = not determined; Tr. = trace; + = presence

<table>
<thead>
<tr>
<th></th>
<th>Axe n.23190</th>
<th>Axe n.23191</th>
<th>Axe n.23192</th>
<th>Axe n.23193</th>
<th>Axe n.23194</th>
<th>Axe n.23195</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>SAM</td>
<td>Slater 1971</td>
<td>SAM</td>
<td>Slater 1971</td>
<td>SAM</td>
<td>Slater 1971</td>
</tr>
<tr>
<td>Sn</td>
<td>5.49</td>
<td>1.82</td>
<td>3.92</td>
<td>3.92</td>
<td>0.77</td>
<td>1.62</td>
</tr>
<tr>
<td>Pb</td>
<td>0</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>As</td>
<td>0.15</td>
<td>0.03</td>
<td>0.07</td>
<td>0.07</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>Sb</td>
<td>0.0427</td>
<td>0.096</td>
<td>0.039</td>
<td>0</td>
<td>0.042</td>
<td>0.019</td>
</tr>
<tr>
<td>Ag</td>
<td>0.0017</td>
<td>0.0024</td>
<td>Tr</td>
<td>0.002</td>
<td>Tr</td>
<td>0.005</td>
</tr>
<tr>
<td>Ni</td>
<td>Tr</td>
<td>Tr</td>
<td>0</td>
<td>0</td>
<td>Tr</td>
<td>Tr</td>
</tr>
<tr>
<td>Bi</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.d.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Au</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.d.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zn</td>
<td>0.15</td>
<td>0.03</td>
<td>0</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Co</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n.d.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fe</td>
<td>+++</td>
<td>0.042</td>
<td>+</td>
<td>0.023</td>
<td>+++</td>
<td>0.042</td>
</tr>
</tbody>
</table>

Table 2. Comparison between different analytical techniques, for an early Bronze Age flanged axe from Pieve Albignola (PV, Italy) at Pavia, Musei Civici (Pearce 1994: tab.7). Data sources: Optical emission spectroscopy – Junghans, Sangmeister and Schröder 1960, 108, 127, n.635; neutron activation analysis and atomic absorption spectroscopy – Berzero et al. 1991, n.31; n.d. = not determined.

<table>
<thead>
<tr>
<th>Element</th>
<th>Optical emission spectroscopy</th>
<th>Neutron activation analysis (except Pb, determined by atomic absorption spectroscopy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sn</td>
<td>0.78</td>
<td>0.851</td>
</tr>
<tr>
<td>Pb</td>
<td>0</td>
<td>0.008930</td>
</tr>
<tr>
<td>As</td>
<td>0.03</td>
<td>0.002971</td>
</tr>
<tr>
<td>Sb</td>
<td>0</td>
<td>0.001731</td>
</tr>
<tr>
<td>Ag</td>
<td>0</td>
<td>0.010936</td>
</tr>
<tr>
<td>Ni</td>
<td>0.02</td>
<td>0.4963</td>
</tr>
<tr>
<td>Bi</td>
<td>0</td>
<td>n.d.</td>
</tr>
<tr>
<td>Au</td>
<td>0</td>
<td>0.000451</td>
</tr>
<tr>
<td>Zn</td>
<td>0</td>
<td>0.000493</td>
</tr>
<tr>
<td>Co</td>
<td>0</td>
<td>0.001831</td>
</tr>
<tr>
<td>Fe</td>
<td>0.04</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Figure 1. Plot of As versus Ag in early Bronze Age flanged axes from Pieve Albignola (PV, Italy) at Pavia, Musei Civici. Data source: Berzero et al. 1991, tabs II and III.
A further problem relates to the validation of analyses: how do we verify published results when there are discrepancies between analytical programmes? Not all publications detail the exact procedures followed, their experimental error or even whether they have used reference standards. This is particularly a problem where data is published in archaeological journals or appendices to archaeological works rather than in publications where such details are required as part of the normal editorial standards, as in mainstream archaeological science journals.

There are therefore a number of specific challenges concerning this large body of data, but it simply cannot be ignored, as it unfortunately tends to be by most prehistoric archaeologists (Pearce 1998, 51); I explore some of the reasons for this below. There are many reasons why we need to consider this data, but they include the simple fact that once an artefact has been sampled (and early analyses tended to require large samples), most museum curators or heritage managers are loathe to permit further damage to the objects in their care. ‘Big data’, or the integration and mining of large and disparate datasets, has become a key priority in a wide range of scientific disciplines (see e.g. for Archaeology, Bevan 2012; e.g. for other disciplines, Reichman, Jones and Schildhauer 2011; Ratib et al. 2014), not least because data which exists should be exploited as far as possible, and it is my contention that the exploration and integration of existing analytical datasets is a major priority for archaeological research.

How can this be done, given the range of analytical techniques used, the differing elements determined and the problems of varying precision and accuracy? One approach, applied by Krause, has been exploratory comparison of two variables at a time (i.e. compositional elements) using graphs with logarithmic scales (e.g. Krause 1988, Abb.76–8; cf. my Figure 1). Others, such as Liverasage (2000) and de Marinis (1979; 2006), have used Waterbolk and Butler’s (1965) frequency distribution histograms. Bray and Pollard have recently suggested using the presence/absence (based generally on a cut-off at 0.1%, renormalizing to account for alloying) of Arsenic, Antimony, Silver and Nickel to define groups (Bray and Pollard 2012; Pollard and Bray 2014; Bray et al. 2015). These approaches are of course limited to the lowest common denominator of the elements determined by the greatest number of analytical programmes, and it is by no means sure that these elements are those with the greatest heuristic value, but they have arguably yielded important results.

The problem of unpublished data: lead isotope analysis

It has become clear that lead isotope analysis is a powerful tool for investigating the provenance of prehistoric metalwork (‘where is it from?’), as was shown by the debate concerning the ox-hide ingots of the late Bronze Age Mediterranean basin (usefully summarised in Lo Schiavo et al. 2009). Indeed, and counter-intuitively, it is now generally agreed that the ox-hide ingots found in copper-rich Sardinia were imported all the way from copper-rich Cyprus, rather than being made from local ores, though the reason why ‘coal was taken to Newcastle’ in this way is still not understood (Hauptmann 2009).

Lead isotope ratios are generally presented in two-dimensional plots (e.g. Figure 2) which may be easily generated from spreadsheets, but in reality there is rarely an unequivocal correlation between the field relating to the ore body and that relating to the artefact(s) analysed (Pollard 2009, 184, 187), even where the artefacts were found in the neighbourhood of the outcrop and so their provenance would seem prima facie easy to establish (Artioli et al. 2012, 171–3, figs 3 and 4). There may be a number of reasons for this discrepancy, ranging from variations in lead isotope ratios within the ore body itself (the result of a complex geological history), as yet not understood effects of the smelting and alloying processes, or simply that the actual ore-body exploited in prehistory has not yet been sampled (a rather dangerous argumentum e silentio), but the discrepancy needs elucidating and investigating, rather than denying.

A more serious problem relates to the way that lead isotope analyses are published, or perhaps it would be better to say, sometimes not published. The problem seems to be the great investment needed to create a reference database of relevant ore bodies, which gives those who have created it an advantage over those who have perhaps analytical data pertaining to artefacts that they have analysed, but lack the full range comparative data with which to compare their results, and so cannot reach useful conclusions. In the early period of the application of lead isotope data to problems of archaeological provenance, the Gales were (perhaps justly) accused of not always publishing their data fully, limiting their data presentation to two-dimensional plots of isotope ratios, and inconsistencies in the reporting of analyses (Budd et al. 1995, 5). This meant that others could not easily verify the correlations that they claimed or reproduce their results. They answered their critics by publishing databases for ore-bodies of the western Mediterranean (Stos-Gale et al. 1995), the Aegean (Stos-Gale, Gale and Annetts 1996), Cyprus (Gale et al. 1997), Bulgaria (Stos-Gale et al. 1998) and the British Isles (Rohl 1996) and more recently their dataset of analyses has slowly begun to be put on line (OXALID, the Oxford Archaeological Lead Isotope Database available at http://oaxalid.arch.ox.ac.uk/, accessed 10 February 2015).

It is of course axiomatic that science should be verifiable and repeatable, and so we might expect that lead isotope data should now be fully published and available to the scientific community, but this is not always the case. Thus for example Jung, Mehofer and Pernicka (2011) present lead isotope ratio plots in a...
scientific publication for artefacts whose provenance is of great archaeological importance, assuring us that the full data ‘... will be published elsewhere ...’ (Jung, Mehofer and Pernicka 2011, 236; promised again in Jung and Mehofer 2013, 187 note 33, but still unpublished at the time of submission of this paper, April 2015!), or analytical data for the artefacts is published, but the reference ore-body dataset from which the provenances are deduced remains unpublished or unreferenced (e.g. Stos 2009).

In fairness to these authors, many scientific journals (and other publications) do not like to publish large amounts of raw data, but this lack of publication means that the conclusions drawn are neither verifiable nor repeatable (using the datasets used by the authors). I would argue therefore that we might question whether these publications constitute true science. There is also another consideration, which relates to the ‘open access debate’ (see for example Darley, Reynolds and Wickham 2014) and it is appropriate to raise this issue in Science and Technology of Archaeological Research, which is of course an open-access journal with a data availability policy. Most archaeometrical research in Europe at least is publicly funded, either through direct project grants or indirectly through the salaries of public employees working in publicly funded laboratories. As such it is arguable that the results of such research should be available to the public; certainly the data is not private property.

It is therefore my contention that editors of journals, conference proceedings and other publications should refuse to publish the conclusions of lead isotope analysis programmes where the data necessary for the verification of those conclusions are not available to readers, whether in the same publication, another publication, or an on-line data repository. Those colleagues called upon to referee such works should make the same point.

The effect of such a policy would be to allow an immediate and major advance in the interpretation of the results of the many programmes of analysis undertaken to date. It would also allow new analytical programmes to enter the arena, competing on a level playing field with those that have existed for a long time but unfortunately have not published their geological data.

Lack of engagement with mainstream archaeology

Analytical data from materials science is only of use to archaeology if it answers specific questions (‘what does it mean?’); analysis is of course not an end in itself. It is, however, my contention that there continues to exist a serious deficit in the use of analytical data by prehistoric archaeologists. There are a number of reasons for this.

In an insightful contribution to the lead isotope debate, James Muhly commented on the difficult relationship between analytical scientists and archaeologists, noting how ‘...far-reaching claims were made early on, claims that could not be supported on the basis of the available evidence. Practitioners [55] came to be seen as scholars who were making up the rules as they went along’ (Muhly 1995, 54–5). He further noted ‘Most archaeologists welcome scientific evidence but abhor scientific controversy. They

Figure 2. 208Pb/206Pb versus 207Pb/206Pb Lead isotope plot, suggesting that some of the Bronze Age bronzes from the Lake Garda lake villages were made from south Alpine copper. Data sources: south Alpine ores – Nimis et al. 2011: tab.2 and for Val Mala, Köppel and Schroll 1985, tab.4; Garda bronzes – Pernicka and Salzani 2011, tab.4.
want answers: if those answers derive from some of the most sophisticated techniques known to science, so much the better, but it is the answer, not the technique, that interests the archaeologist’ (Muhly 1995, 55). I do not want to argue that the technique is not important: indeed its refinement, its limitations and its applicability are all important aspects of scientific research. The problem is that archaeological scientists often seem more interested in the technique than in the specific answers that it can give to archaeological problems.

Archaeometallurgists seem to have formed a ghetto (Killick 2015, 298), with their own conferences (such as the very successful ‘Archaeometallurgy in Europe’ series) and journals (whether specific to archaeometallurgy like Historical Metallurgy or more general archaeological science journals, like Archaeometry or Journal of Archaeological Science). It is rare that articles appear in mainstream journals discussing archaeometallurgical topics (though see, for example, Pearce 1998 or the 1995 Special section on ‘Lead isotope analysis and the Mediterranean metals trade’ in Journal of Mediterranean Archaeology 8 (1), 1995, 1–75), and when sessions are organised at more general archaeological congresses, like the Annual Meetings of the European Association of Archaeologists, they tend to be dominated by discussion of technique rather than the contribution that these can offer to the resolution of archaeological problems, and so are deserted by the generalist archaeologists who are uninterested in the minutiae of technical problems. If archaeometallurgists want to escape the appendices of archaeological reports they must leave their ghetto and engage with the mainstream archaeological debate. This is not to argue that the generalist archaeologists are not also responsible for this lack of communication. Too many (especially in the USA – Muhly 1980, 102) lack specific training in archaeological science to equip them to use archaeometallurgical data, and improving the training of archaeologists, for example in statistics, may alleviate this problem, but as Muhly (1980, 102) remarked, ‘As archaeology grows ever more technical and scientific the problems of control and comprehension facing the humanist grow greater’. My point is that if archaeometallurgists want their data to be used they have to communicate them to the generalists, and in a way that can be understood without misrepresenting their complexity. I would add that if the data are not used, if they do not answer specific questions, then they are probably not true science.

A further reason for the lack of communication with mainstream archaeology is the fact that many senior figures in archaeometallurgy were not trained as archaeologists and thus do not feel the need to engage with the debate about social, symbolic or other aspects of interest to archaeologists (Killick 2015, 298). They also do not always understand the point that ‘archaeologists reish the use of scientific evidence and the conclusions drawn therefrom, if they fit within the limits of a general sense of “historic probability”’ (Muhly 1995, 57); that is to say, archaeological data has an important role in the framing of the interpretations and hypotheses resulting from analytical data.

Conclusions

Answering the important questions ‘what is it made of?’ and ‘where is it from?’ is primarily specific to the domain of archaeometallurgists (though of course the latter also involves typological and other studies), but their interpretation (‘what does it mean?’) is pertinent to mainstream archaeology. Answering this latter question requires better communication between archaeometallurgists and archaeologists, and in this short polemic I have tried to offer some suggestions for the improvement of dialogue between them, which I hope will lead to more use of analytical data by archaeologists and thus ultimately to better science.

Acknowledgements

I am grateful to David Killick, Carl Heron, and two anonymous referees for reading and commenting on a draft of this paper. They are of course not responsible for the opinions expressed.

Author Biography

Mark Pearce is Professor of Mediterranean Prehistory at the University of Nottingham (UK). He studied at the Universities of Cambridge, Lancaster and Pavia, and has taught at the Universities of Pavia and Nottingham. His research focuses on northern Italy and early mining and metallurgy. His most recent books are Rethinking the north Italian early Neolithic (London 2013); The Death of Archaeological Theory?, edited by John Bintliff and Mark Pearce (Oxford 2011); and Bright Blades and Red Metal: essays on north Italian prehistoric metalwork (London 2007).

ORCID

Mark Pearce © http://orcid.org/0000-0003-4209-7923

References


Darley, Rebecca, Daniel Reynolds, and Chris Wickham. 2014. Open access journals in Humanities and Social Science 3. Berlin: Mann.


