

# Butser Ancient Farm

---

Project Consultant : Christine Shaw

## The Nature of Experiment in Archaeology

In the same year in which this paper was published, as a chapter in a book honouring John Coles, the [Conference Paper](#) from 1996, on which this is based, was also published.

### Dr. Peter J. Reynolds.

The object of this paper is to explore the nature of experiment in archaeology today and to assess its potential role insofar as it may confirm or deny interpretations of data from excavations.

In addition, there is an urgent need to define the meaning of experiment and further to disassociate archaeological experiments from both education and experience in archaeology.

At the outset, it is a fundamental tenet that experiment has absolutely nothing to do with the exercises of 'living in the past', 'dressing in period costume', 're-enactment of past events' or, indeed, the teaching of well understood techniques - which may well have been originally established by the experimental process - like, for example, lithic technology, pottery manufacture or laying mosaics. The former are at best theatre, at worst the satisfaction of character deficiencies; the latter are simple skills which, should they wish to be acquired, require learning. It is extremely unfortunate that these activities have become generally subsumed under the overall title of experimental archaeology since their inclusion militates against the real value of experiment and its acceptance professionally. The labelling of an activity like shaving with a flint flake or even a Roman bronze razor as an experiment rather than exploration is clearly absurd. It advances our knowledge not one iota and serves generally to increase our prejudices of history and pre-history.

The misunderstanding of experiment in archaeology has been brought about by the confusion of three separate issues: experiment, experience and education. Experiment will be dealt with at length below. Experience is a completely different issue and invariably involves people doing things and discovering for themselves the nature and application of a range of technologies. To manufacture a flint arrowhead for example is to experience, learn and/or execute a technology.

Similarly to coppice a hazel woodland, to till a field, to mix daub, to manufacture a pot is to come to terms with material on the one hand, and on the other to appreciate the nature of hard physical work. That all of these and a myriad other activities are of value is undeniable. Indeed they are all the more laudable in that understanding of the requirements of these activities is gained and thus an increased sympathy if not empathy with the past is occasioned. There is nevertheless a great gulf between the experimental and the experiential.

Education necessarily is integral to both experiment and experience. The original remit of Butser Ancient Farm was a programme for research and education. Essentially, unless its results are communicated and are therefore educative, research is relatively valueless. Further, the methodology of research itself is a core element of education. Experience is perhaps the greatest and best teacher of all. Ancient technologies are a fundamental building block of the human state. However, experiment is the ultimate arbiter in that it supplies the confirmed material of and for both education and experience.

An experiment is by definition a method of establishing a reasoned conclusion, against an initial hypothesis, by trial or test. There is no doubt that experiment is a scientific term and, therefore, engenders in the lay mind an almost pathological fear of non-comprehension. Experiment is seen to be from the complex worlds of physics, chemistry, mathematics and biology and is, therefore, arcane and, without lengthy training, incomprehensible. Sadly this is undoubtedly the result of inadequate and uninspired education in most people's formative years. Consequently it is much easier to abandon such complexities to specialists and to rely upon their reports. These reports and findings may or may not influence politics, which is invariably driven by expediency, or the humanities which are similarly driven by fashion and/or religious conviction. Yet the methodology of experiment is extremely simple to understand. Its execution, on the other hand, is often likely to be detailed and demanding.

Archaeology has traditionally been an 'arts' subject and it is only relatively recently that science has begun to have a significant role. The normal pattern of archaeological activity has been, especially in the field of prehistory which is largely uncluttered by documentary evidence, to excavate a site and, within the perceptive knowledge and experience of the excavator, to interpret and publish the findings. It was O.G.S.Crawford who described this process as the 'disciplined use of the imagination'. Inevitably it has become more complex as the data base has increased and as other subjects like ethnography, ethnology and the physical and biological sciences have made their contributions. Consequently interpretations have become more soundly based but, nevertheless, remain no more than interpretations. Their limitations are necessarily defined by the data, the excavated evidence, its quality and frequency and, not least, the manner of its recovery. The current practice of avoiding total excavation of a site and leaving a sector for future excavation, against a time when techniques are further refined, is a very real

recognition of present day technological inadequacy, which in turn is superior to the methodology of even the recent past. It is against this background that any way of examining an interpretation has to be an improvement upon its blind or unquestioning acceptance. Consequently the experimental process actually enhances interpretation.

Its application can be readily appreciated by the following formula. A site is excavated and its product, described as the prime data, is subjected to analysis and interpretation. Rather than use the term 'interpretation' which implies full comprehension, the term 'hypothesis' is substituted. Hypothesis implies a deduced or reasoned conclusion which can and should be further subjected to test or trial to confirm or deny that conclusion. The method of testing is called an experiment. This is built to the specification of the reasoned conclusion using the prime data as the given evidence.

The experiment, therefore, is not an exercise imagined or concocted on an unconstrained basis by the experimenter. It is quite specific to a particular hypothesis and data resource. Partiality, therefore, is removed in principle. However, bias many still enter especially where sampling contains an element of human choice or estimation. Notwithstanding, the ambition of the experiment is not only to explore the hypothesis to its extremities but even to its destruction.

The requirements of an experiment are also specific. The experiment must satisfy the tenets of the academic or technological discipline within whose remit it falls. For example, an agricultural experiment must be acceptable within the disciplines of agriculture and agronomy. An experiment must be replicable and replicated. An experiment should be designed so that the results may be assessed statistically, otherwise the outcome is again little more than subjective or partial.

Once the experiment has been conducted within the above parameters, the outcome is compared directly with the prime data upon which the hypothesis was postulated. If the comparison is positive, the hypothesis can be accepted as valid. If negative, the hypothesis is to be rejected as disproved and, therefore, wrong. Given the focus upon the prime data during the construction and execution of the experiment, in the case of the hypothesis being dismissed as in error, it is often the case that an alternative hypothesis can be formed and subsequently tested. In this context during the execution of an experiment, it is occasionally realised that the experiment will be a failure unless it is altered. However, to do so in the midst of the trial would be to deny the point of the experiment itself. It is necessary to conclude the experiment and thus disprove the hypothesis before embarking upon a changed and, therefore, new experiment which is, in fact, testing a changed and new hypothesis. This new hypothesis naturally enough is the result of an enhanced perception of the prime data, taking account of the new information derived within the experiment.

Thus the formal experimental approach can be seen to be cyclical in form. However, in common with many formulae there is an important corollary. It is perfectly possible, especially given the limited nature of archaeological evidence, to derive more than a single hypothesis from a set of prime data. Nor is it unlikely that a number of such hypotheses could be reasonably validated. This factor itself underlines the significant difference between validity and truth or reliability.

There are further *caveats* to be aware of in the conduct of experiments, the most important of which is to dismiss the human element. It may seem rather odd to emphasise this point since archaeology is essentially the study of man in the landscape through or at a given time but it is critical that an experiment is inanimate. No experiment can be designed to enhance our understanding of human motive or emotion in the recent or remote past.

It is also signally valueless to record the time taken to achieve an end product. Time taken may be of mild interest, even wonder, on the part of the experimenter but that interest, that wonder, is entirely the result of the temporal state of the experimenter. It may have nothing to do with the actuality of the past. This, however, does not deny the significance of the time needed for a natural physical process to be completed. In the simple case of firing pottery, the time taken to achieve the ceramic change from clay to pottery within the context of the variables of kiln type, fuel type and clay fabric, time is of undoubted interest in that it is independent of human motive and emotion. Almost by definition, experimenters who record human input are recording their own prejudices, efficiency or inefficiency and are, therefore, not conducting experiments on those factors. Similarly those who record their feelings or emotions are recording modern and minor irrelevancies.

Once an experiment has been satisfactorily conducted within all the above limitations and the hypothesis has been validated, the result becomes an accepted 'given'. In the case of disproof there is clearly need for further analysis and experimentation. Thus experiment is locked into the interpretational process. It is a direct check against absurdity and wild flights of fancy and removes the sometimes ludicrous claims of fashion. However, experiment is significantly associated with the basic hardware of archaeology in that it tests the understanding of the products of excavation. For example, an experiment can only be focused upon the nature of the primary evidence in so far as that represents structure, process and function. It is restricted to the building blocks of interpretation and can only have an influence upon the broad generalisation and period overview. That it can significantly affect these is without doubt. A peculiarity of Roman Britain, for example, has been the so-called Romano-British grain dryer (Morris) which experiment proved conclusively would not dry grain efficiently or economically (Reynolds). A second hypothesis, that such structures could have been malting floors for the production of beer, was tested and validated. The primary experiment denying the original hypothesis immediately removes these structures from the agricultural process

because, on the one hand it is a straightforward negative conclusion, on the other an important positive component in the consideration of the overall Romano-British agricultural economy. In the coin of interpretation, if the obverse is the hypothesis then the reverse is experiment.

The types of experiment are as naturally diverse as the material evidence they seek to examine. It is as well to realise and underline the fact that the data recovered by excavation, despite it being representative of less than one percent of the original material, is indicative of human activity in all its forms. Therefore, experiment will necessarily draw upon virtually all the sciences in its exploration of hypotheses. In order to simplify the complexity thus implied, it is possible to group experiments into broad categories provided it is clearly understood that these categories are complementary and inter-dependent rather than exclusive. In general terms experiments can be grouped into five categories: the construct, process and function, simulation, eventuality and technical innovation.

The construct experiment is perhaps the easiest of these categories to understand. It is defined as the exploration, at a 1:1 scale, of the third dimension. It is exemplified by the examination of prehistoric, Roman and proto-historic houses and structures for which the evidence is only patterns of post-holes or simple foundations identified by the excavator as buildings. At this point, however, it is critical to focus upon the purpose of the experiment. If the experiment is to explore an hypothesised structure, then by definition it uses a specific set of data from a specific site. **The experiment is thus totally restricted to those data and cannot import further convenient data from other sites.** The experiment in its execution explores the adequacy and inadequacy of the prime data and ultimately has the potential to feed back into the prime data features, necessary for the construct, which may have been unrecognised as associated with the hypothesised structure or wrongly attributed or even unseen but photographically recorded and subsequently recognised as critical elements. This site specific aspect of exploring a construct cannot be over-stressed.

It is ironic that over the last thirty years building pre- and proto-historic structures throughout Europe has been virtually a growth industry. By the same token it is extremely regrettable that the motivation for a large number of these buildings has not been a genuine desire to explore the archaeological data from a specific site but rather to erect a generic museological and/or educational resource. It seems rather illogical, if not irresponsible, on the one hand to display and teach a fashionable image, on the other to lose the opportunity of testing a specific hypothesis.

The plethora of European long houses are generally identified by their overall similarity, by their having no specific ground plan as excavated, by their being thatched with river reed whether the plant is available locally or not, and by their use of secreted 150mm nails (fulfilling perceived health and safety requirements and the lack of faith of the builder). They are distinguished one from another by

their level of internal decoration, by whether their attendants are in costume and playing a role or not, and by whether members of the public are voyeurs or potential participants. It is, no doubt, a reflection of modern society that the primary purpose is not the quest for enhanced understanding of the past but rather an enhanced bank balance. The real tragedy lies in the fact that with greater forethought and planning the causes of both specific experiment and museological resource could be equally served for similar expenditure. In recent years it has become the turn of the Romans to be exploited. It is now possible to bathe Roman style in Holland and Germany albeit in bathhouses built of modern materials. Had the Romans had access to modern building materials they would have doubtless used them. Since they did not, it is somewhat perverse to gull the public when it would have been perfectly possible at little increased cost to replicate the original materials and provide a research/academic model as well as a museological theme park. The supporting argument of 'n' thousand visitors and 'x' ecus is quite simply fallacious and deceptive.

In this context of experiment, the term 'construct' has been intentionally used throughout. Normally in any discussion or description of pre- and proto- historic or even Roman buildings, the word employed is 'reconstruction' which, in itself, implies for the lay reader if not the professional a spurious degree of certainty. From famous reconstruction drawings wherein wisps of smoke, elemental rainfall and convenient clouds shroud uncertainties (Sorrell) to the certainty of reconstructed roundhouses which inconveniently collapse, the word is misapplied. Ideally its use should be associated with buildings or objects for which sufficient material evidence survives for accurate reconstruction to be possible. In effect, the building of a reconstruction is generally restricted to those open air museums throughout Europe which seek to rescue exemplary period structures and subsequently present them to the public as specific time capsules. Sixteenth century farmhouses like the Bayleaf Farmhouse at Singleton Museum of Buildings, in England, or nineteenth century farm complexes at Sentendra Museum in Hungary are cases in point. When only a ground plan survives, any structure based upon it can only be conjectured and is, therefore, best described as a construct. Similarly such a construct should be quite specifically designed to explore the nature of the ground plan and any adjacent evidence which may not initially be recognised as being part of the ground plan.

However, the above does not deny any role for reconstruction within the experimental process. Indeed reconstruction is a quite vital element. For example, many wooden objects, particularly agricultural implements like prehistoric ards have been recovered from waterlogged deposits. To build accurate new replicas or reconstructions of such implements allows functional experiments to be conducted into the efficacy and efficiency of these tools. In fact the results of such trials with such ards have led to a complete re-appraisal of prehistoric agriculture. But this type of experiment falls into the second category, that of process and function. Since this kind of experiment by definition involves a passage of time, it is less

susceptible to museological or thematic perversity. Ironically although the term "process and function" clearly indicates the passage of time, as observed above, one of the least valuable results is that which measures human input. How long it takes to thatch a roof, plough a field, make a joint, are questions dependent upon human motivation and skill with the tool to hand. Clearly the time taken to achieve an outcome is also subject to the perception of time within an historical context. Modern perception of time is undoubtedly different, embracing as it does a range of contemporary economic and political connotations and denotations contrasting strongly to that of even a century ago.

Process and function experiments seek to examine how things were achieved. One particular example involved the proposition that large pits found on many Iron Age sites in Britain and Europe were used for the long term storage of grain. A long and complex series of experiments were carried out by the author to explore this hypothesis. The experiments were particularly significant since the unsubstantiated interpretation proposed had been used as a major pillar in an argument to compute population estimates, based upon grain consumption per capita and against pit capacity, assuming that a pit had a functional life span. The experiments not only established the methodology of storing grain in underground silos and its efficiency or otherwise, depending upon a range of variables, they also proved that it is possible to store seed grain - that is, after storage in a pit the average germinability of the grain was in excess of ninety percent - and that the pit had an unlimited life. Thus the pit proved to be no more than an innocent container and storage failure was due not to it becoming contaminated in any way but due either to acceleration of the life-cycle of micro-organisms endemic to the grain being stored, or to water penetration, or a combination of both. Thus the experimental sequence, which spanned a period of fifteen consecutive years, validated the hypothesis (interpretation) that certain types of pits could have been used for the bulk storage of (seed) grain but invalidated any argument which sought to establish population estimates. In effect, the results of the experiments broadened considerably the understanding of the potential agricultural economy of the Iron Age.

There is a continuing need for further process and function experiments, especially with regard to implements which currently attract definitive interpretations, as well as with structures which are customarily designated as buildings or features having only a single purpose. In this latter case one only has to consider the standard four post structure found ubiquitously on prehistoric sites. To ascribe the standard or traditionally accepted interpretation to such structures that they are overhead granaries, is not only to seize upon a convenient label and thus deliberately avoid potential concomitant data but also denies the fact that virtually any structure from watch-tower to multi-functional shed can be based upon a rectilinear arrangement of four post-holes. In essence, unless there is a wealth of accompanying evidence of function, experimental studies could validate a whole range of hypotheses whilst confirming none of them.

Indeed, the third category of experiment is specifically designed to address problems of this nature. The overall category is called simulation. In simple terms, the objective is to understand elements of archaeological evidence by projecting backwards from the excavated state to the original or new state and then monitoring the deterioration through time until the archaeological state is reached. While such experiments are necessarily long term, they are unlikely to require the passage of millennia. A construct has locked within it a simulation trial, should the construct be used and its deterioration observed and monitored. However, the best example perhaps is the experimental earthwork. Several experimental earthworks have been built in Europe this century but undoubtedly the most celebrated are the Overton and Wareham Down Earthworks in Britain. These were designed as linear earthworks with a bank and box section ditch. Different materials like leather, bone and pottery were located in the bank in order to study their movement, degradation and potential survival. The overall purpose was to observe the erosion of both bank and ditch, through time, with an excavation programme scheduled on a binomial progression of years passed. The thirty-two year excavation of both earthworks has recently been completed. Subsequent to these monumental earthworks, a new series of earthwork experiments was implemented in the early 1980s. These experiments were designed specifically to examine the nature of the typical domestic enclosure ditch of the Iron Age period. In detail, the nature of the ditch is a 'V' shaped section 1.50 metres deep and 1.50 metres across the top. The bank, of dump construction, contains variables of turf retaining walls and turf cores, with and without a berm. The standardised plan of the earthwork is octagonal with a twenty metre length of ditch and bank opposed to each main and intermediate point of the compass. The design purpose is far less complex than the Overton and Wareham Down earthworks, seeking only to monitor erosion and revegetation. The life span will be determined once revegetation is complete and, therefore, erosion has ceased; currently this is estimated at between ten and twenty years. At this point, a full history of the revegetation will have been recorded against a daily meteorological record. Sections will then be cut across each treatment variable, which will provide a working example, where the full history, including extremes, is known, for the field archaeologist to use to interpret excavated ditches.

The pilot scheme to the octagonal earthwork programme was designed only to study erosion episodes and layer depositions unexpectedly reversed the accepted method of interpretation of ditch sections, in that the skewed layers are most affected by the open (non- bank) side of the ditch.

Because enclosures are found on a range of soil and rock types, experimental octagonal earthworks have been set up on, respectively, upper and lower chalk and the aeolian drift of the coastal plain of central southern England. One further such earthwork has been created in Catalonia, Spain, examining the same questions but against a hugely different climatic pattern and on 'marga' rock (a sedimentary limestone) and on its derivative soil.

The simulation experiment thus seeks to unravel the problems of how material evidence arrived in the state found. In practice it provides a paradigm which the archaeologist can compare to the actual data and, given correlation, can elucidate for the archaeologist the physical processes involved in creating the data. In addition, in the case of the experimental earthworks, there is an abundance of vegetational and meteorological data which are directly and fundamentally involved in, if they are not actually the cause of, the physical processes.

Present day archaeologists are largely oblivious to the above processes and the natural world from which they derive. Thus it is essential to appreciate the need to study the nature of plant communities and their interdependence. For example, the results from the experimental earthworks demonstrate the following plant succession, the early plant colonisers, like the mosses, which cling to bare rock and initiate exploitable niches for other opportunist plants by trapping soil particles, the mid-term colonisers like nettles and thistles, the long term occupiers, the grasses which, in their turn, will be dominated by brambles and thorns and ultimately provide the habitat for tree seedlings to take root and flourish. The impact of man in using any of these as raw materials must be separately tested.

The vegetation sequence above argues for interference management especially when an enclosure is used for living and working. Brambles, for example, recreate bare earth conditions which will initiate a new phase of erosion, which in turn might be observable in the layers deposited in the ditch. Similarly the snail populations vary against the nature and abundance of the vegetation cover. These and many other questions should have been addressed within the context of the earthwork programme but it is the very nature of this type of experiment which enhances and broadens the perception and indicates future programmes which need to be implemented, in order to understand more fully the archaeological data.

Such observations beg many supplementary questions. Is it possible to extract further working paradigms which will prove eventually what is seen to be happening? To take another example, incident pollen rain unfortunately depends for its survival upon soil acidity levels which are hostile to agricultural exploitation but nonetheless a search for pollen is worth consideration.

The fourth category of experiment is virtually the combination of the first three categories. Described as an eventuality trial, it seeks to explore the potential product. For example, one of the greatest problems in understanding a prehistoric, classical or historical society is to be able to assess the underlying agricultural economy. Public buildings, fine cities, complex societies all depend upon successful exploitation of the landscape. It is a truism to record that climate drives landscape drives man. Until recently, man's activities in exploiting landscape were circumscribed by the nature of the local climate and its variability, the underlying geology and the soil itself. Beyond the probably misunderstood technology of decreasing soil alkalinity by the application of animal dung, the farmer of the past

was entirely constrained by his landscape and the flora it would bear. Against this background, given that climate has changed remarkably little in the last three millennia, with the exception of minor and relatively short lived episodes and that the soil types are also exactly similar, the best example of an eventuality trial is that which seeks to explore the agricultural potential of the past.

Our knowledge of ard/plough technology is considerable and capable of replication. Similarly from carbonised seed evidence the crops grown, including some of the weeds of those crops, are also known. The landscape exploited is necessarily adjacent to the settlements and to a very large extent is unaltered by the passage of time or even the treatments of modern agriculture. Virtually all the cereals exploited in the remote past have survived to this day and are available if difficult to obtain.

For the past twenty years, a series of eventuality trials has been carried out at Butser Ancient Farm seeking to examine the potential of the late Iron Age agricultural economy. The archaeological evidence indicates a full agricultural facility in terms of implements, embracing as it does the rip ard for cultivating virgin or fallow land, the tilth ard which is remarkably successful, even in comparison with modern ploughs, and the seed drill ard which argues for sophisticated plant management. Almost by definition this includes maximising the seed germinability, reducing input and, thus, creating an increased return as expressed by input:output ratio. The presence of hand tools like hoes and the scale of Iron Age fields imply regular plant maintenance within the context of artificial time-reward management, in that most agricultural requirements can be met within one working day. Similarly there is evidence for manuring and, by implication, non-manuring practice, autumn and spring sowing, and even crop rotation of nitrogen fixing and nitrogen using plants.

However, in setting up an eventuality trial there is still insufficient evidence for a precise research programme of specific replication. The deficiencies lie particularly in quantity choices - how much manure is applied per hectare and what weight of seed is planted per hectare. In consequence, the construction of the eventuality trial requires the establishment of a series of limits against which the variables can be examined. The output variable comprises simply the product, the yield per hectare against treatment. The weather pattern from planting to harvest is an input not under direct control. Other input variables are respectively the inputs and treatments, seed and manure, fallowing and rotation, planting times and the soil type.

Of all the variables the weather is doubtless the most significant. Infinitely variable in itself, despite the statistical comfort of averages, the weather has always been and continues to be the primary factor between agricultural success and failure. The farmer's perennial pre-occupation with the weather is entirely justified.

This type of experiment is extremely complex, because of the number of variables involved, and so needs to be repeated over a considerable number of years, a suggested minimum being a decade, not only to achieve statistical validity but also to ensure that all or as many as possible variations of the weather have been experienced within the trial period. Such a trial, for example, would be essentially valueless if run over one or two years. One further important consideration is sheer scale. The field areas involved must be sufficiently large to allow for typicality to be experienced. A field edge, for example, is subject to other important variables than the bulk of the field area, in that greater rooting facilities and inward nutrients may or may not be available. A research plot of a square metre, therefore, can have no validity whatsoever in assessment of yield.

The results of such an eventuality trial need to be treated with the greatest care simply because they are the product of specific combinations of variables. While averaging results through time, especially against climatic variability, may give an overall figure, the average is still specific only to the selected treatments, the weather and most particularly to the soil type. While it might reasonably represent the potential product of prehistory and history, it can only reflect the potential of a specific soil type and landscape. It is perfectly possible to manipulate the figures against different soils and landscapes but not at all sensible to transfer the figures indiscriminately.

In fact, this particular type of trial has been a core research programme of the Ancient Farm since its inception in 1972. To date, two soil types in different landscapes, on one underlying geology, have been examined for eighteen years. A further soil type on a different geology in a different landscape has been under examination for five years. In addition, research outstations in Catalonia, Spain, and Hungary have been in operation for respectively seven and two years. Ultimately comparisons will be possible with appropriate adjustments. In reality none of the soil types, with the exception of Hungary, would be regarded as of the highest quality for cereal growing. Ironically, for the longest trial period, the reverse is the case with the trials operated under the worst possible option. Nonetheless if one averages the results for the worst option over all treatment variables the product or yield is some 2.5 tonnes per hectare, a figure which, remarkably, equates to the national average yield in Britain in A.D. 1950. If the experimental results are uprated to incorporate better soil types in less hostile landscapes, the expected yield must surely be larger still.

However, the experimental programme underlined the nature of prehistoric agricultural practice, as perceived, in that it was more comparable to market gardening than the cultivation of 'broad acres'. Input time would have been far in excess of that two millennia later. This does not deny, of course, that the countryside was far more densely populated in the Iron Age and Caesar's reference to the export of grain in the first century B.C. is amply substantiated by the experimental results. Perhaps the most significant intermediate result of this

ongoing research programme is the complete rebuttal of the proposition that agricultural production in the prehistoric period was at subsistence level and that it was not until the arrival of the Roman influence that commercial viability was achieved. Why this proposition gained credibility is difficult to understand since the prehistoric agricultural technology of north-west Europe was the same as, if not actually superior to, Roman technology but with the added great advantage of a far better climate.

The fifth category of experiment, entitled technological innovation, is quite obvious if generally unrecognized and unappreciated as an experimental procedure. This kind of experiment describes the testing of new scientific equipment, though not necessarily new within its own designated area, to improve archaeological data acquisition. It also embraces evaluating equipment specifically designed for archaeological purposes. In the former case, the classic example is the assessment of the resistivity meter for use in prospection surveys. The initial application, with a device called the Mega Earth Tester, was an experiment. An example of the latter is the magnetic susceptibility meter which was specifically developed to examine the topsoil for traces of human activity. Similarly, recent experiments have been carried out with ground radar, X-rays, thermal sensing and many other techniques which may or may not prove to be of value practically or economically to archaeology. The whole thrust of this type of experiment is inspired by increased awareness of the potential within archaeological data. Especially is this so with the application of methodology from the physical sciences, not least of which is soil chemistry.

The recognition that the topsoil is an archaeological resource has heralded a new range of approaches, all of which are initially, by definition, experimental. Indeed, an extremely simple experiment in this category was set up to monitor the manner and extent of movement of artificial artefacts comprising standardised pieces of plastic, containing a tiny magnet, placed systematically at 50mm depth in the topsoil and subjected to modern and prehistoric cultivation practices. Far from the accepted hypothesis that artefacts were infinitely separated from their point of deposition, it was discovered that virtually ninety percent of the material remained within two metres of its start point.

From the above brief descriptions of the proposed categories of experiment, it can be readily appreciated that each and every category is not exclusive of all the others. It may well be simpler and clearer to divide them out for explanatory reasons but it would be quite wrong to regard each category as a stand-alone exercise.

The purpose of this paper has been to define experiment in archaeology and to argue that experiment is an inescapable element of interpretation. Where interpretation is capable of being tested, it should be tested. The testing process itself must be rigorous and should not admit the variables of human motivations.

On completion, the test or experiment will provide a positive or negative result. A positive result will validate the interpretation or hypothesis. A negative result will disprove the interpretation requiring another to be raised in its place. It should not be surprising that the contribution from experiment is most frequently negative. Experiment is necessarily restricted to those hypotheses which are capable of direct examination and have an adequate data base, not only to allow the hypothesis initially but also to formulate the experiment itself. In addition, an experiment must be repeatable, including repetition by other agencies.

*This paper appeared as a chapter in "Experiment and Design in Archaeology" in honour of John Coles - published by Oxbow Books 1999, pp 156-162, edited by A F Harding.*