Chapter 2

GENERAL CONSIDERATIONS

1. PHILOSOPHY

When a product fails unexpectedly, experience has shown that in almost every case the problem can be traced back to lack of, or inadequate, testing, which in turn resulted from an attempt to save money. It has to be admitted that testing can be very expensive in both time and money; so why is it essential? Put simply, men and women make mistakes, machines go wrong and we don't know enough.

If people are going to make mistakes we have to check their production. Similarly, if machines can vary in their performance we have to check their output. Generally, we don't have enough knowledge to make a product and be sure that it will work. The customer is unlikely to believe us if we said we did and expects us to test to prove fitness for purpose. Demands for greater quality assurance and consumer protection, together with improved performance, are likely to result in more testing rather than less.

We certainly don't know enough to design a new product without making use of material property data, and as new materials are continually developed there is an ongoing need to test their properties. As design methods become more sophisticated and expectations of performance increase we need better data even for established materials. In this context, because rubbers are such complex materials the demands on testing are perhaps particularly onerous.

When things go wrong we often do not know why. If we did we probably would have stopped it happening. Hence, we may also test to fathom out the reasons for failure.

From this reasoning as to why testing is necessary, the purposes of testing can be summarised:-

- Quality control
- Predicting service performance
- Design data
- Investigating failures

Before considering which properties to measure and which methods to use it is essential to clearly identify the purpose of testing because the requirements for each of the purposes are different. This may be an obvious point, but failure to appreciate what purpose the results must satisfy easily leads to unfortunate choice of method and conditions. Also, lack of consideration of why another person is testing and what they need to get from their tests frequently leads to poor appreciation of the merits and limitations of a particular test and inhibits communication between, for example, the university researcher and the factory floor quality controller.

There are a number of general requirements for a test method; it must have adequate precision, reproducibility etc. There are, however, particular attributes related to the purpose of testing:-

For quality control: the test should preferably be as simple, rapid and inexpensive as possible. Non-destructive methods and automation may be particularly attractive. The best tests will additionally relate to product performance.

For predicting product performance: The essence of the test must be that it relates to product performance - the more relevant the test to service conditions the more satisfactory it is likely to be. Extreme speed and cheapness are less likely to be important but there is a need for test routines which are not excessively complex. Non-destructive methods may be acceptable.

For producing design data: The need is for tests which give material property data in such a form that they can be applied with confidence to a variety of configurations. This implies very considerable understanding of the way material properties vary with geometry, time etc. Extreme speed and cheapness are of relatively minor importance, there is little interest in non-destructive methods. For complex and long running test automation may be desirable.

For investigating failures: Having established what to look for, the need more than anything is for a test which discriminates well. There is often little need for absolute accuracy or in some cases even relevance to service.

There is of course nothing black and white about attributing these requirements to the purposes of testing but they indicate the emphasis which usually applies in each case.

Tests are usually classified by the parameters to be measured - mechanical, thermal, electrical etc. These can be sub-divided to list the actual properties so that under mechanical, for example, there are strength, stiffness, creep etc. This form of classification will be used in this book because of its convenience. However, in terms of the purposes of testing discussed above and considering what is required of the results from a particular method, this classification is not particularly useful. A more generalised way of classifying tests is to consider:-

- Fundamental properties or tests
- Apparent properties or tests
- Functional properties or tests

Regardless of the type of property or particular parameter chosen, this classification can be helpful when considering what is needed from the result and, hence, which test method should be used. Taking the example of strength, the fundamental strength of a material is that measured in such a way that the result can be reduced to a form independent of test conditions. The apparent strength of a material is that obtained by a method which has completely arbitrary conditions and the data cannot be simply related to other conditions. The functional strength is that measured under the mechanical conditions of service, probably on the complete product.

This classification can be loosely linked to the purposes of testing. For quality control, fundamental properties are not needed, apparent properties will usually be acceptable, although functional properties would certainly be desirable. For predicting service performance, the most suitable properties would be functional ones. For design data, fundamental properties are really needed, although considerable help can often be got from functional properties. For investigating failures, the most useful test will depend on the individual circumstances but it is unlikely that fundamental methods would be necessary.

When looked at in this way, the gaps in readily available methods become obvious: most measures of mechanical properties yield apparent properties and there is a need for fundamental methods, whereas most dimensional methods and many thermal and chemical tests give fundamental properties. Overall, there is a dearth of fundamental tests. It is worth noting that when measuring the effects of environment, weathering for example, for use as design data, it may not be necessary to use a method giving absolute results to monitor changes with time. An apparent method may suffice because the change in property need only be comparative.

It becomes clear that there can never be one direction for the development of test methods and apparatus. The perceived deficiencies in the existing methods are viewed differently according to the particular purpose under consideration and, hence, development effort is targeted appropriately. However, collectively, the advances which are generally sought have remained constant over many years - quicker tests, cheaper tests, more reproducible tests, better design data and tests which are more relevant to service performance.

Many of the tests described in this book are standard methods, often with their roots in tradition, which cannot meet everyone's needs. Most frequently they are best suited to quality control and only in relatively few cases are they ideal for design data. For as long as one can remember the most often voiced criticism of existing test methods, particularly standard ones, has been that they are arbitrary and do not measure the fundamental properties needed for design purposes. This complaint has been so consistent that it is perhaps surprising that sufficient improvement has not been made. For many properties, it is extremely difficult to devise a fundamental test and where one can be devised it is likely to be relatively difficult or expensive and only required by a minority. For mechanical properties, the methods for generating input data for finite element analysis and the fracture mechanics approach to failure can be noted. More useful information can generally be obtained at the expense of measuring a physical property as a function of test piece geometry etc. Dynamic properties are an example of where methods useful for design exist but even now only a relatively small proportion of industry uses them.

In contrast, those needing tests for quality control are more satisfied with existing methods, but nevertheless an equally consistent complaint is that the tests should be quicker and cheaper to perform. Vast strides have been made in efficiency through the automation of apparatus and the manipulation of data, but commercial pressures are such that improvements continue to be sought.

In the same way that tests based on arbitrary conditions are deficient for design data purposes, so they may tend to lack in their direct relevance to service conditions and, hence, their value for predicting service performance. The two situations are not identical, in particular a test may simulate service use to enable predictions to be made but not yield data which can be used in design calculations. Not only for product proving but also for quality control, there is increasingly demand for tests which are better in this respect. A prime difficulty is that as effort is made to make the method reflect service so it tends to become more complicated and more expensive. There are many instances in specifications where a more relevant test exists but is not used because it is more time consuming or complex.

Until the 1980s most people remained unaware of how reproducible their rubber test methods were. Then, interlaboratory test programmes revealed the true scale of the problem. Long established test methods were found to have far poorer reproducibility than previously realised, in some cases to the extent that it could be questioned whether the tests were are worth doing and whether specifications based on them are valid. Hence, there was greatly increased interest in improving reproducibility, although this has more often been focused on better standardisation of the test and calibration of the test parameters than on new methods.

A somewhat paradoxical situation has arisen in more recent years because of the increased interest in comparative data and databases. Some of the pressure for better design data and the criticism of the standardised methods has been turned to a demand for very tightly standardised data. To be comparable, data in a database needs to be all produced in exactly the same way and the development of standards to offer a choice of method and several choices of conditions for the same property is not compatible with this need. Consequently, there has been a lobby for what might be termed extreme standards which are specifically intended to yield completely comparable data very efficiently, but possibly at the expense of other attributes.

2. TEST PIECE HISTORY

The properties of a material and, hence, the test results obtained will depend on the processing used to produce the test material or product. Frequently, this is beyond the control of the tester and he or she is only required to characterise the samples received. Nevertheless, if any controlled comparison is to be made it is essential that the test material is produced in exactly the same way. Similarly, any preparation which is performed on the material to produce the test pieces is likely to influence the results. Consequently, it is highly desirable that preparation is standardised and comparisons only made between test pieces produced in the same way, including the direction within the sheet that test pieces were cut. Standard procedures are discussed in chapter 4.

The history of a material or product between manufacture and testing can clearly affect its properties, although the history may not be known to the tester. Normal practice is to adopt standard conditioning procedures to bring the test pieces as far as possible to an equilibrium state, although this will not generally compensate for any degradative influences to which the material may have been exposed. In rubber testing, conditioning usually only involves temperature but if the material or property is sensitive to moisture then the conditioning atmosphere should include a standard humidity. Occasionally, mechanical conditioning is used in an attempt to reach equilibrium of the transient structure of the material. Conditioning is dealt with fully in chapter 5.

These comments make it clear that any result is not material specific but relates to the particular sample of material, the manner in which it was processed and what has happened to since it was formed. In an ideal world the test result would be accompanied by statements covering test piece history but frequently this is not possible.

3. TEST CONDITIONS

Whilst the fact that changing the test conditions will almost certainly change the result obtained is generally appreciated, it is not always strictly taken notice of. There are plenty of good reasons for using different conditions - to better simulate service, to use geometries which can yield design data, to obtain data as a function of temperature, to allow tests on irregular shaped products and so on. There are also plenty of opportunities to vary the conditions unintentionally.

Many test results are sensitive to the geometry of the test piece and many of the geometries used are arbitrary so that a specified geometry should, where possible, be adhered to and it must be appreciated that it may not be simple to convert results to a different geometry. The classical example is assuming that a property is proportional to thickness when in a great many cases, for various reasons, it is not. Generally, it is necessary to have data as a function of geometry or to know the relationship between the two before conversion is attempted.

Even standard methods of test often allow alternative test pieces or procedures and these may not yield equivalent results. Hence, it is important to clearly define which procedure has been followed and, when a standard has been used, to identify any deviations made from the set procedure. Test procedure requires careful attention to detail as small, apparently innocent changes can produce significant deviation in results. Equally, it is essential that that test conditions are accurately set and maintained, which is really just one aspect of quality control considered in more detail in Section 7.

These comments on factors affecting test results may seem extremely obvious but it is a simple fact that failure to pay sufficient attention to them is the main cause of the poor reproducibility that has been found when comparisons between laboratories have been made, and the reason for most disputes over test results.

4. STATISTICS

Earlier editions of this book had a complete chapter devoted to this subject which opened by commenting that it was tempting to claim that it was the most important chapter. The reasoning was that, whatever property we measure, whatever test method we use we end up with results and the question "What do the figures really mean?" Results are useless unless we know their significance; significance means statistics. However, at that time it was very unusual for statistical methods to be applied to rubber testing.

For many reasons, not least the influence of the quality movement and the widespread availability of personal computers, statistical methods are now much more widely appreciated and more frequently applied to the results of rubber tests. Also, a practical reason for not now needing a statistics chapter is the existence of the comprehensive British Standard Application of Statistics to Rubber Testing[1] which at the time of writing is being considered for adoption as an ISO standard. It contains references to standards on statistical methods and also has a small bibliography.

Despite the existence of this standard, and indeed many other standards and good text books on statistics, it is worth emphasising that statistics has an important role to play both in the analysis of results and in designing the experiment. All the clever analysis in the world will not compensate for poor experiment design and planning. In particular, it is no use screaming for the statistician to sort out the mess after the testing has been done. If help is needed it should be called in at the very beginning. With regard to experimental design, a very useful and review and comparison of different designs has been given by Hill et al[2].

One sign of statistics being applied is seen in the precision statements which have been added to many test method standards. These give measures of the within and between laboratory variability which were obtained from an interlaboratory trial conducted under specific conditions. Although it is true that a different set of figures might have been obtained from another trial with a different group of laboratories, they are representative of the variability which can be expected. Generally, those taking part would be judged as being among the most experienced in the industry and less good figures would not be too surprising from a broader range of less knowledgeable testers. Hence, when the quoted precision figures are relatively poor it is necessary to subdue the inclination to believe that if they were all like one's own laboratory this would not happen. The general standards for accuracy of measurement methods and results are the ISO 5725 series[3] but ISO TC 45 has its own procedure[4] that differs in part from the general ISO method. The ASTM method[5] for rubber is very similar but the UK has not adopted the TC 45 standard, believing that the general ISO method should be used.

Another statistical measure increasingly being used in connection with results is the estimate of uncertainty. No measurement is exact; there is always some uncertainty as to the trueness of the figure obtained. It is possible to make estimates of the likely uncertainty by considering the uncertainties introduced by each factor involved in the measurement. This includes, for example, the accuracy of calibration of each instrument used and the variation in applying the procedure. Note that uncertainty and accuracy are not the same thing – accuracy of an instrument is just one factor in the uncertainty of the measurement result. Accredited calibration laboratories have for a long time been required to make uncertainty estimates for all their measurements and the same practice is now applied to test results. The generally accepted procedures are given in Guide to the expression of uncertainty in measurement (GUM)[6] and there is also an ISO technical specification that deals with uncertainty estimates from precision results[7]. Estimating uncertainty is not an easy matter and some assistance is available in the form of a practical guide to application of the GUM methodology[8]. One of the problems with estimating uncertainty from combining the contributing factors is that it is extremely difficult to get a measure for some of the factors. ISO TC 45 is currently considering the production of a guide to how to deal with this situation which, if successful, should be extremely useful.

TC 45 has also recently produced a standard for the evaluation of the sensitivity of test methods[9], sensitivity being defined as a derived quantity that indicates the level of technical merit of a test method from the ratio of the test discrimination power or signal to the noise or standard deviation of the measured property. There is a very similar ASTM method[10] but, again, the UK argued that such a method should be produced by ISO TC 69 for test methods in general as there is nothing in the standard that is specific to rubber.

5. SAMPLING

The significance of test results depends to a considerable extent on how the physical sample was obtained. Whatever the purpose of testing, it is necessary to question whether the samples tested adequately represent the population being investigated. In many cases, one is limited by the amount of material available, there may be only one product or batch to be evaluated, but in routine quality control there is the added dimension of needing to sample repetitively in time. This means that a good measure of the population mean and variance is obtained eventually but there is need for a long term sampling plan and a continuous method for assessing the results.

The nature and size of a sample and the frequency of sampling obviously depend on the circumstances. First, the number of test pieces or repeat tests per unit item sampled must be decided. Our current standard methods are not consistent, ranging from one to ten or more, and it is usually argued, although open to challenge, that the more variable a test the more repeats should be made. There is no doubt that financial considerations have played a large part in the deliberations, witnessed by certain very variable but long-winded methods calling for one test piece only. There is no doubt that to use one test piece only is rarely satisfactory but testing very large numbers will not yield a proportional increase in precision. There is a trend towards five as the preferred number and this has a lot to recommend it for the more reproducible tests, being just about large enough to make reasonable statistical assessments of variability. An odd number of tests is advantageous if the median is to be extracted. In a continuous quality control scheme the number of test pieces at each point is usually rather less important than the frequency of sampling, i.e. it might be better to use one test piece but check five times more often.

Efficient sampling really boils down to selecting small quantities such that they are truly representative of the much larger whole. The necessity for sheets to be representative of batches and for batches to be representative of the formulation is self evident. The direction of test pieces relative to the axes of the sheet and randomisation of their position in the sheet are also important if the sheet cannot be guaranteed homogeneous and isotropic.

When powders are sampled, devices must be used to take representative from the sack, drum or other container, bearing in mind that coarse particles tend to separate out.

In the rubber factory, sampling is very much influenced by the fact that rubber production is a batch process and that for moulded products each heat (or lift) constitutes a batch (in a different sense). A common procedure is to sample each batch of compound mixed, but by the time the finished product is rolling off the lines several batches may well be intermixed. Good quality control schemes will enable batch traceability to be achieved. The selection of discrete products should preferably be randomised and certainly care must be taken that the sampling procedure is not biased, for example, by sampling at set times which might coincide with a shift change or other external influence. A book of random numbers (a set of tables designed to to pick numbers at random without the risk of unconscious bias) is invaluable. Sampling is very much part of quality control and information, particularly from a statistical point of view, can be found in BS 903-2[1] and in quality control text books.

6. LIMITATIONS OF RESULTS

It becomes clear from the discussion of the previous sections that any test result is not absolute but is limited by a variety of factors. Before testing starts there are limitations arising from how the material was produced, how it was sampled and how the test pieces were formed. The results are further limited by the form of test piece, the selection of the test method and the exact test conditions adopted. The actual results obtained are then subject to uncertainty limits that arise from such factors as natural material variation, tolerances on the accuracy of test instruments and tolerances on test conditions.

When an estimated uncertainty of a result is quoted it refers to the uncertainty associated with that particular measurement. It can be used to demonstrate significance of the result, for example whether it is significantly different from another result or whether it is significantly above a specification level. However, it tells us nothing about the significance of the result in terms of whether it is typical of the day's production or how different it might be from a result obtained in another laboratory. These further uncertainties can only be estimated by carrying out tests on a number of batches of production, using different test machines, etc. Although these uncertainties are real, in practice they are often overlooked because assumptions are made, such as assuming that the sample tested was typical of the whole population. This may be expedient but don't bank too heavily on getting the same results next time.

Significance in the statistical sense refers to what reliance can be placed on a result taking account of experimental error, or the extent to which the result is typical for the material. Significance can also refer to the relevance of a result in terms of material or product performance. A result might be proven to be highly significant statistically, typical of the material and exhibiting low uncertainty, but if it is of minor importance to the product performance it would not be significant in practical terms.

7. QUALITY SYSTEMS

Quality assurance is concerned with maintaining the quality of products to set standards. This embraces the control of incoming materials, the control of compounds produced, the control of manufacturing processes and guaranteeing as far as possible the quality of the final product. Quality assurance schemes utilise physical testing methods as a most important part of their system. In fact most of the standardised test methods are principally intended for quality control use and probably, in terms of quantity, the majority of testing carried out is for quality assurance purposes.

Taking quality assurance in a wide sense, it is necessary to consider specifications, the relevance of test methods, the accuracy of test methods and the statistically based control schemes which make up the discipline of the quality engineer. This is a specialised subject that happens to involve testing and it is not appropriate to consider here quality assurance of the production of rubber products.

However, in the same way that we expect factory production to be subject to a quality assurance system, so the test laboratory needs its own quality procedures. To keep apparatus, procedures and people in the best condition to produce reliable results requires systems and control. Almost certainly, the best way of achieving this in a testing laboratory is to be subjected to the disciplines of a recognised accreditation scheme. The ISO9001[11] standard is commonly applied in industry and the laboratory will be included in that system. However, more rigorous and focused schemes for test and calibration laboratories have been standardised in ISO/IEC 17025[12] which requires procedures for everything from the training of staff and the control of test pieces to, most importantly, the calibration of equipment. To maintain the requirements, which are given in deceptively short form in the standards, is both time consuming and difficult but anything less than these standards is not ensuring the highest possible quality in the output of the laboratory - the results. Many countries have a national body entrusted with accreditation of laboratories to this standard (for example The United Kingdom Accreditation Service, UKAS) and they interact through such bodies as the International Laboratory Accreditation Conference (ILAC). Some of these bodies have mutual recognition agreements.

Whilst all aspects of a laboratory's operation require systematic control, it is the calibration of test equipment which gives rise to most problems and which is also the most expensive. All test equipment and every parameter of each instrument requires formal calibration. For example, it is not good enough to calibrate the force scale of a tensile machine, there are also requirements for speed of traverse, etc. plus associated cutting dies and dial gauges.

Calibration is based on the principle of traceability from a primary standard through intermediate standards to the test equipment, with estimates of the uncertainty which increases at each step in the chain. Wherever possible, bought in calibrations should be carried out by an accredited laboratory. It is perfectly acceptable for the test laboratory to do its own calibration but then they must maintain appropriate calibration standards and operate a measurement management system in accordance with ISO10012[13]. One factor which has hindered full appreciation of the detailed needs of adequate calibration is the lack of definitive guidance. The position ha been greatly improved with the publication of ISO 18899[14], Rubber – Guide to the calibration of test equipment, and the adoption of the practice of adding a calibration schedule to all ISO rubber test methods. The guide outlines the requirements for calibration and procedure to be used for each parameter and is intended to assist test laboratories who are not experienced in calibration. The schedules list all the parameters and the associated tolerances for the test method in question and are intended as advice to the calibration laboratory.

Another area which has tended to be overlooked is the validity of manipulations made on the test data. It is probably reasonable to trust a calculator to perform a simple arithmetic operation - although that may not always be the case with the operator. However, increasingly data is being manipulated by a computer to automatically produce the test result involving quite sophisticated operations. This includes such things as area compensation, modulus calculation and curve fitting. If you carry out these tasks by hand any abnormalities are likely to be apparent but a computer will happily carry on regardless. As they say, rubbish in, rubbish out. It is essential to verify any software used to ascertain that it will produce valid results under all circumstances. A particularly obvious example is to account for offset zero points but others can be quite subtle. A computer will apply a strict formula to deriving figures from a stress-strain curve whereas a human will make judgments based on knowledge and experience. However, there appears to be little international standardization of guidance on software verification.

The object of quality control procedures in the laboratory is to produce correct and reproducible results. Up until the 1980s, although good reproducibility was desired and it was known that some tests were better than others, it was assumed that for most properties the level of agreement between laboratories was reasonable. There was not a wealth of published data to support or contradict this complacent state but the scattered accounts which could be found almost always revealed large discrepancies. One must surmise that that these did not raise great concern because of a general attitude that when there was disagreement the other chap had done something wrong!

When ASTM, followed by ISO and others, started conducting systematic interlaboratory trials to obtain precision data for test methods, the true state of affairs became apparent[15]. For many standards the variability was worse than realised and in some cases was so bad as to question whether the tests were worth doing at all or whether specifications based on them could be considered valid. The general advance of the quality movement prompted these investigations and have ensured that reproducibility has continued to occupy one of the top spots for attention in recent years.

There are a number of reasons for excessive scatter of results found between laboratories - wrong calibration, incorrect apparatus, misinterpretation of the standard, deviation from the procedure, operator mistakes etc. They reduce in the end to either the standard being too lax in its specification and tolerances or somebody is doing something wrong. An interlaboratory comparison tells you the magnitude of the scatter but not which of the possible causes is responsible. That requires further and probably very expensive investigation.

As mentioned in Section 4, interlaboratory comparisons organised by ISO committees are conducted with what are reckoned to be good quality laboratories so that they might be expected to represent an optimistic situation. However, there is some unpublished evidence that a comparison within a closer group, for example all UKAS accredited, produces better results. This would tend to indicate that more fault lies with mal-practice than with the quality of standards. On the other hand, the few investigations of the uncertainty of standard methods have found areas where the tolerances need to be tightened. The third factor, the variability of the material tested, needs to be kept in mind because there is a limit to the useful tightening of test equipment tolerances. In fact, for most tests the calculation of an uncertainty budget reveals that by far the largest factor is the material variability.

There have been various initiatives to investigate the causes of variability and make improvements but financial restrictions have kept the scale of these modest in relation to the size of the problem. One of the earlier ones was initiated by the UK Ministry of Defence[16]. There are essentially three approaches:- a) interlaboratory trials with the organiser visiting each laboratory and probing into the apparatus and procedure used; b) normalising the consistent bias of each laboratory against an arbitrary "standard" laboratory and on-going monitoring of changes in the level of bias; c) systematic investigation and quantification of the possible effect of each parameter and, hence, identifying those that require closer tolerance and deriving the theoretical level of variance to be expected.

The first approach is that classically employed and in relatively small groups has had notable success. There has been a concerted effort made in Sweden for several physical tests along these lines[17,18]. A very carefully designed and researched proposal for the second approach known as Intercal was made and a prototype run by the USA[19] but, unfortunately, it has not been further developed. An example of the third is given by a Rapra analysis of hardness[20]. The functions of interlaboratory testing as a quality assurance tool in the rubber industry have been examined by Leete[21] whilst Koopman considered the idea of comparing test methods by their sensitivity[22].

Interlaboratory trials with the organiser making detailed assessments of the laboratories is clearly particularly suited to helping individual laboratories and will at least qualitatively indicate the parameters requiring attention. This approach is, however, very expensive in total effort. The Intercal approach does not identify the causes of variability immediately but certainly alleviates the effect and, because trials are on-going, allows improvement to be monitored. Systematic quantification of the effect of individual parameters is probably the most cost effective approach and is the most useful for aiding standards committees to improve the specification of methods, but is of less direct help to individual laboratories.

Any shortcoming in a standard can only be put right after analysis has pinpointed the problems. Hence, standards committees cannot act quickly if an interlaboratory trial reveals excessive variability. It is highly unlikely that faults in standards account for the majority of variance, although clearly it is important that any that do exist are identified and action taken.

The most powerful tool to minimise the component of variance due to error in the laboratory is the discipline which recognised accreditation schemes bring. They encompass all the likely areas which produce mistakes, documented procedures, training, checking procedures, control of samples, monitoring conditions, formal audits and perhaps above all calibration. The general quality movement has produced pressures to make laboratory accreditation commonplace and as more laboratories reach this status it must be expected that reproducibility will improve. In the current economic climate, a problem is finding sufficient laboratories able to devote sufficient time to precision trials.

8. TEST EQUIPMENT

The basic requirement for test equipment is that it is adequate for its purpose - it needs to comply exactly with any standard test method being used, be in good working order and be properly calibrated. However, there is then scope for a considerable range of level of sophistication, ease of use etc.

Going back in time some 50 years, laboratory equipment was almost all manually operated and often very dependent on the skill of the operator. The greatest change in test laboratories since then, and the rubber laboratory is no exception, is the improvements made to apparatus by the introduction of automation and, in particular, the application of computers to control tests and handle the data produced. These developments can and do influence the test techniques which are used, for example by allowing a difficult procedure to become routine and, hence, increase its field of application. However, advances in instrumentation and data handling are primarily noticed as improvements in efficiency or accuracy rather than intrinsically improving the relevance of tests to product performance. That is, the technical developments more often change the way the test is performed rather than change the basic concept. It is not practical to include a chapter on instrument hardware and software but, wherever appropriate, comment is made in later chapters on the form of apparatus now available for any particular test.

It is worthwhile to bear in mind the ways in which instrumentation advances have been advantageous, and also their less desirable aspects. Automation in particular is first thought of as saving time and, hence, money. If the test can be left to measure itself and an operator's time is saved, there is a particularly attractive cost benefit. However, automation is also frequently very important in improving accuracy, reproducibility or making a procedure possible.

Some processes are taken for granted, for example no one is on record as having sat up all night adjusting the controls of an ageing oven, and to manually maintain a temperature ramp on a temperature retraction test, although attempted, is the next thing to impossible. Thermal analysis techniques such as DSC only became feasible with developments in instrumentation, tailored dynamic loading cycles needed the introduction of servo-hydraulic machines and many other examples could be cited where we could not have the test without the instrumentation.

Automation frequently aids accuracy and/or reproducibility by being more consistent than humans. Non-contact extensometers ensure no unwanted stresses on the test piece and any automatic extensometer will be less subjective than a technician with a ruler. Digital thermometers, load cell balances and many other apparatus introductions have made measurements easier and less prone to operator error.

Time and cost saving has been most notable in the logging and processing of results where computerisation has amounted to nothing less than a revolution. Around 1970 it was estimated that a rubber testing laboratory could spend half its time processing results and presumably quite a bit more in recording them in the first place. That time is probably now only a few percent. It is also significant how views have changed. Then, it was widely held that direct links between test machine and computer were only justified in a few cases. Now, any major equipment is likely to be operated via the keyboard.

The automation of sample handling has not taken off as some predicted in the nineteen sixties when the first automatic systems were developed for tensile machines and hardness and density apparatus. Robots are rare alongside the test rig and the reason is doubtless to do with volume, as such automation only becomes worthwhile when a very large number of identical tests have to be made.

Advances in instrumentation have not been without their disadvantages. On a pure time saving basis, tests would now be remarkably cheap but the cost advantage has been counteracted by the fact that more sophisticated apparatus costs more money and is likely to be outmoded more quickly, leading to much higher capital costs. Although development should make equipment more reliable it can be generalised that more complicated and advanced equipment requires more maintenance by highly skilled and highly paid people. The cost side of the equation has also been added to by rising standards of calibration and laboratory quality control generally. In this context, it should be noted that expensive, sophisticated equipment is all very well when a large volume of testing is needed but cannot be justified for occasional use.

The calibration of more sophisticated apparatus has also been fated with additional problems arising from the difficulty of directly reaching the actual measured values. The software which so efficiently transforms the data can give rise to concern as to what has happened between the transducer and the final output. As mentioned earlier, the software itself requires verification which is often not an easy task.

When technology allows it, there is a natural tendency to specify lower and lower tolerances on equipment parameters but this does not necessarily bring significant advantage because, for many properties, the contribution to uncertainty from material variability far outweighs that from machine accuracy. When reduced tolerances cannot be fully justified there is an unreasonable cost burden to be borne by the laboratory.

9. THERMOPLASTIC ELASTOMERS

Thermoplastic elastomers are, by definition, not a conventional rubber nor a typical thermoplastic. Consequently, there has been a long and unfinished debate as to how they should be tested. A long time ago a paper considering their particular requirements[4] concluded that for most physical properties the methods used for vulcanised rubbers were suitable, and as a generality that is probably still true. Very significantly, ISO TC 45 decided to add thermoplastic elastomers to the scope of rubber test methods wherever the method was thought suitable, so that now the majority of ISO physical test methods for rubber include thermoplastic in the title. Consequently, although thermoplastic elastomers are not necessarily specifically mentioned in each chapter of this book, with few exceptions it is assumed that both vulcanised and thermoplastic materials are covered in the accounts of physical tests.

This is not of course the complete story. Most often, thermoplastic elastomers are processed on plastics machinery and it will be convenient, and sensible, that test pieces are produced in the same way. The thermoplastics processability tests are also likely to be more relevant and, certainly, curemeter tests are irrelevant.

The most suitable physical properties are likely to depend on the particular material, with plastics test methods being used for the harder elastomers (where the title elastomer may not even seem appropriate) and rubber methods for the less hard and more elastic materials. Where thermoplastic elastomers are to compete with conventional rubbers then clearly rubber test methods will be expected. On the other hand, where they are being compared to normal thermoplastics it would seem reasonable to use appropriate plastics test methods.

It is unfortunate that test methods for soft plastics and for rubbers, although very similar, are not identical, for example differences in tensile stress strain, tear and hardness methods. If they were aligned, much of debate about which method to use would be eliminated. For some properties, there is a distinct difference in approach. For example, glass transition temperature is frequently determined for plastics whilst various low temperature tests have been specifically developed for rubbers.

Some of the conditions used in rubber test methods may need modifying for application to thermoplastic elastomers because of their intrinsic thermoplastic nature. If the temperatures generally used in ageing and compression set tests on thermosetting rubbers were applied to thermoplastic materials they could appear to perform extremely badly. Whether this was significant would depend on the service temperature. Data sheets need to be checked as those for thermoplastic elastomers may have used much lower temperatures that would be found for conventional rubbers, and it is only too easy to get a misleading impression of performance.

At the time of writing, there is a proposal in ISO TC 61 for a standard on Acquisition and presentation of comparable data for thermoplastic elastomer materials along the lines of those already in existence for thermoplastics. The first draft is rather different from documents on the same theme proposed in TC 45 for rubbers generally and it is to be hoped that either the two committees can cooperate on the production of a thermoplastic rubber document or the idea is dropped.

10. PRODUCT TESTING

If our knowledge of the properties and behaviour of rubbers, and hence our design rules, were such that we could predict the performance of the product accurately from tests on laboratory test pieces, then perhaps product testing would be rarely needed. The serious problem of the changes which the manufacturing process introduces can be overcome by obtaining test pieces from the product as discussed in chapter 4. However, the fact is that our understanding of the properties of rubber is simply not good enough to make performance predictions reliably in a great many cases, even if the test pieces come from the product. Hence, there will often be need to test the whole product to be sure that it will perform satisfactorily.

In the case of a new design it can be more expedient, and certainly effective, to subject prototypes to real service rather than to develop simulation tests. However, there are many cases when this is simply not sensible for time, cost or safety reasons. So, when real service trials have to be ruled out and prediction from laboratory material tests cannot be relied upon then there must be whole product testing.

It can be extremely difficult and/or expensive to devise tests to simulate service adequately and the justification for investment will be in proportion to the importance of the product in risk and/or sales terms. There is clearly much skill in designing rigs and test schedules which maximise information gained at minimum cost. In practice, there is danger of spending very large amounts and still not getting the simulation accurate enough, but most commonly the pressure is to under design the apparatus and to curtail the test programme to cut costs. By far the most difficult factor is when assessing durability and there are a number of degradative agents and some form of acceleration is required to reduce the time scale.

The same principle applies to quality control testing, but here there is much greater probability that the experience gained from proving the product initially will allow the quality of subsequent production to be reliably judged on the basis of tests on test pieces or the product test procedure can be simplified.

Sometimes a product test will give more valuable assessment of quality for the same testing cost as needed for test pieces. This would be true, for example, for compression testing of a simple engine mounting (Figure 2.1) because the cost of moulding test pieces would be little different from the value of the mounting and the testing costs would be equal. It would be pretty pointless to go to the trouble of cutting test pieces from the mounting. When the value of the product is high, it is again a matter of judging whether control on test pieces gives sufficient confidence to reject the costly alternative of product tests. In this situation, non-destructive tests become particularly attractive.

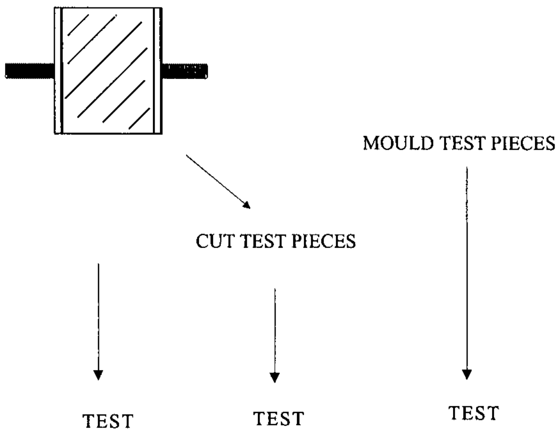

Figure 2-1. Choices for compression testing an engine mount

For both quality control and design or performance evaluation purposes, it is relatively clear when whole product testing is desirable. The question then becomes one of whether it is considered essential and, if so, how sophisticated the experiment should be. This can only be answered, albeit with great difficulty, by weighing the cost against the risks and values involved. It should not of course be forgotten that, although we may not know fully how to make predictions from material tests, for many products experience will have shown what level of material properties will be satisfactory. It would probably be fair to say that in the past the tendency has been to be somewhat frugal with product tests. There now seems to be a trend towards more product tests being specified in standards. Generally, more people want to see evidence of fitness for purpose and CEN, for example, have a policy of producing performance rather than construction/material standards.

On first reaction, this would seem to be wholly good in that logically performance tests on the product should give the greatest certainty that it will be satisfactory in service. However, it is extremely difficult and expensive in most cases to devise adequate simulation rigs. The pressures of standardisation are to demand that they are produced quickly and almost inevitably without any obvious source of funding. The most expedient route often has to be taken and rarely are there the resources to properly evaluate and refine the methods decided on.

The result can be methods which do not adequately fulfill their objectives in properly simulating service or are unnecessarily complex and unworkable within reasonable cost. Reproducability of rigs can be very bad. There is a world of difference between a rig for development purposes in one laboratory and multi laboratory product certification. If new methods are introduced which are ambiguously written or without full interlaboratory comparisons then problems and disputes are likely to follow. It can be concluded that it would be better to rely on material properties than on inadequate or ill-defined product tests but a well designed product test provides the best proof of fitness for purpose.

REFERENCES

1. BS 903 Part 2, 1997. Guide to application of statistics to rubber testing.
2. Hill A, Buchholz H-V and Wenzel K. ACS Rubber Division 147th Spring Meeting, Philadelphia, May 2-5, 1995, Paper 56.
3. ISO 5725, in 6 parts, 1994, 1998. Accuracy (trueness and precision) of measurement methods and results.
4. ISO TR 9272, 2005. Rubber and Rubber Products – Evaluating precision for test methods.
5. ASTM D4483, 2003. Standard practice for evaluating precision for test method standards in the rubber and carbon black industries.
6. Guide to the expression of uncertainty in measurement (GUM), BIPM/IEC/IFCC/ISO/IUPAC/IUPAP/OIML, corrected and reprinted 1995.
7. ISO/TS 21748, 2004. Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation.
8. BS PD 6461-4, 2004. Practical guide to measurement of uncertainty.
9. ISO 19004, 2004. Evaluation of the sensitivity of test methods.
10. ASTM D6600, 2000. Standard practice for evaluating test sensitivity for rubber test methods.
11. ISO 9001, 2000. Quality management systems – Requirements.
12. ISO/IEC 17025, 2000. General requirements for the competence of testing and calibration laboratories.
13. ISO 10012, 2003. Measurement management systems – Requirements for measurement processes and measuring equipment.
14. ISO 18899, 2004. Rubber – Guide to the calibration of test equipment.
15. Veith A G. Polym. Test., 7, 4, 1987.
16. White I. Eur. Rubb. J., 175, No. 2, 1993, p 10.
17. Spetz G. Polym. Test., 12, 4, 1993.
18. Spetz G. Polym. Test., 13, 3, 1994.
19. Veith A G. Polym. Test., 12, 2, 1993.
20. Brown R P, Soekarnein A. Polym. Test., 10, 2, 1991.
21. Leete J L. ACS Rubber Division 155th Spring Meeting, Chicago, Paper 9, 1999.
22. Koopmann R K. ACS Rubber Division 143 Spring Meeting, Denver, Paper 44, 1993.

http://www.springer.com/978-0-387-28286-2

Physical Testing of Rubber
Brown, R.
2006, VIII, 388 p. 30 illus., Hardcover
ISBN: 978-0-387-28286-2