frontiers | Frontiers in Dental Medicine

TYPE Review
PUBLISHED 23 January 2023
DOI 10.3389/fdmed.2022.1084006

OPEN ACCESS

EDITED BY
Julian G. Leprince,
Institut de recherche sur les médicaments de Louvain, Université catholique de Louvain, Belgium

REVIEWED BY
Matthew German,
Newcastle University, United Kingdom
David Watts,
Victoria University of Manchester, United Kingdom
Ulrich Lohbauer,
University of Erlangen Nuremberg, Germany

*CORRESPONDENCE
Brian W. Darvell
b.w.darvell@bham.ac.uk

SPECIALTY SECTION
This article was submitted to Dental Materials, a section of the journal Frontiers in Dental Medicine

RECEIVED 29 October 2022
ACCEPTED 19 December 2022
PUBLISHED 23 January 2023

CITATION
Darvell BW (2023) Mechanical test relevance—A personal perspective on some methods and requirements.
Front. Dent. Med 3:1084006.
doi: 10.3389/fdmed.2022.1084006

Mechanical test relevance—A personal perspective on some methods and requirements

Brian W. Darvell*

Dental Materials Science, University of Birmingham, Birmingham, United Kingdom

Mechanical tests of various kinds are widely used in dental research to study the behaviour of its materials. Unfortunately, despite its often long history, the relevance of a test or its outcome is hard to discern. But even for tests that have a more apparent appropriateness, many details that ensure good accuracy and reproducibility appear not to be appreciated or understood in that there is little evidence in the published literature that the necessary steps and precautions have been taken. The present purpose is to examine a number of these aspects in the context of a selection of tests to illustrate the care and attention that are essential for sound results.

KEYWORDS

dental materials, strength, setting time, hardness, flow, film thickness

Introduction

Mechanical tests of several kinds are used in the dental literature for materials research for various reasons, most often for product comparisons, less commonly for fundamental studies. It has been observed over a long period by the present author that the protocols used are rather variable, and often idiosyncratic, in many details. Indeed, where such detail is in fact reported, this is too often with fatal flaws (1). Many are used out of habit, it would seem, for no better reason than that they have been used frequently before by others, and spurious claims of the need to generate results for comparison may be made. Superficial simplicity is attractive; detailed analysis for validity is an apparent anathema. It is unfortunately necessary to state that the repetition of a fault does not make it right. Tests that have been in use for long periods, repeatedly, whether through familiarity, habit or inertia are not validated, justified or endorsed simply by that repetition. It is necessary to examine whether there is value in their continuance and thus, ultimately, for the end goal of patient well-being through understanding.

What is striking is that very rarely indeed is there any analysis of the relevance of a test's circumstances to those of the material's service. Even more rarely is there seen much appreciation of the mechanics of such tests and need for careful set-up and execution. Many seem to be oblivious to the difficulties, pitfalls and subtleties of laboratory work of this kind. In fact, many seem to fall back on tests used in ISO product standards on the false assumption that they are the best possible (2, 3), and without realizing that many are quick and dirty (*pace,* ISO) for quality control (QC), safety and minimum efficacy. There is no pretence that they are the last word, and indeed may (and do) change from time to time as knowledge improves.

There are many types of tests in the present context, and a full review would require a dissection of very many papers. This is not a practical proposition: there are many shortcomings in such matters in the published corpus, and indeed misuse. The piecemeal approach, commenting on previous work one topic at a time, seems not to have much effect (despite citations), and plainly there is not much teaching of such things at undergraduate or postgraduate levels.

The present purpose then is to examine a few of the more common kinds of test, to identify principles and factors, dangers and nuances, and good practice where possible, without drawing up detailed protocols. Obviously, each application and material needs it own appraisal, preferably with an analysis of errors, and potential variations with justification would probably be rather numerous. The whole ground cannot be covered in short compass. Even so, where possible some pointers have been given, and alternatives suggested.

Reference will be made to many matters of materials science without elaboration, although some pointers are given [*via* (4)]. This is not meant to be a tutorial in that sense. It is left to the reader to pursue such items separately, although in several cases a cue is given for further specific reading. This is by no means an exhaustive list.

Preliminaries

At the outset it is necessary to establish what it is desired to determine. For this, it is valuable to employ the "First Three Questions" approach (1). Under this, it is essential first to identify what is the fundamental piece of information sought, the property or character, without prejudice with regard to potential measures or the methods that might be used. The second is to identify the natural measure of that property or character, and again without prejudice with regard to the methods that might be used or the equipment or expertise that might be available. The third step is to identify the proxy measure(s) that will be used in actuality because it is rare that a direct measure is even possible. Thus, for example, elastic modulus cannot be measured as such, it requires two other measures first and a calculation.

Only then, in the fourth step in the process, is it appropriate to identify the means by which the proxy measure(s) may be obtained. It is the purpose now to explore some of the various mechanical tests that might be invoked. Such proxies then lead to the definition of a Figure of Merit (FoM). This may be an actual material property, but more often it is an arbitrary measure that is assumed to represent some key characteristic.

The pragmatic aspects of time, costs, equipment availability and expertise do need consideration. But such decisions as are then involved depend on a cost-benefit analysis, on an individual experimental basis, and are beyond the present scope. Research involves an effort to do better, but faulty work cannot be justified by lack of facilities: care and attention to detail can overcome much. Ignorance of pertinent chemistry, physics or mechanics is no excuse.

Broad principles

The key word, it is suggested, is relevance. For a test to be relevant to the clinical context (which is assumed to be the overriding consideration here), some principles need to be kept in mind.

Theory

There are many fundamental properties that may underlie behaviour in service, yet it is frequently very difficult to obtain such values in a theoretically-sound fashion, or with good accuracy. Understanding of theory is a necessary prelude to designing and undertaking any test, to justify the choices made, or even to allow for assumptions and approximations: awareness of risks and interferences is always necessary. What may be used as a test for QC purposes may fall short in rigour quite legitimately, provided it yields information interpretable in context. That is to say that the purpose for dentistry is often functional or operational, not theoretical as such: e.g., what is the behaviour in service, what controls handling? What are the effects of defects and how can they be controlled?

Mode of challenge

In many cases this means that the mode of loading is such that the nature of the stress experienced by the test piece is ideally identical to that in service. Pragmatically, a close approximation if required because the exact conditions in what amounts to innumerable prototypes cannot easily be ascertained or mimicked. Consideration of the kinds of design (say, of a filling) that may be used may lead to more than one kind of stress field being relevant and thus imply more than one test is to be used, unless correlation has been unambiguously established (no such examples are known). This is to say that the mechanical (for example) circumstances of use need to be properly understood first.

Mode of failure

Despite seeming to mimic, say, the mode of loading, if the test piece does not fail in a manner seen in actual service the result is essentially meaningless and uninterpretable. For example, explosive fragmentation of a material in service in the mouth is never seen. The question then is what is seen in practice? The mode of challenge must be modified in order to achieve that kind of mimicry.

Test conditions

All materials have many properties that depend on the conditions under which they are tested. Thus, for example it is well understood that organic polymer systems are strain rate-sensitive, and most especially when working near the glass transition temperature (T_g) when dramatic changes may be found. However, such effects may also be seen in other than polymers when any component of a system is susceptible to flow, by whatever mechanism. Similarly, in any test where there is the formation of new surface, whether by deformation or fracture, the energetics of the process are dependant on the surface energy of the material, and thus the environment in which it finds itself (4, Chap. 10). The oral environment may in short be described as wet, warm and

salty, and generally slightly acidic. Such factors must be taken into account for condition mimicry.

It is important to note that ordinarily all materials in service in the mouth for any length of time will equilibrate with respect to the activity of the water of saliva or other body fluids with which they are in contact, to be essentially saturated (allowing for the presence of solutes). In other words, given that enamel and dentine are permeable, nothing stays "dry", whether prosthetic, restorative or endodontic, no matter its initial condition. Testing dry is therefore almost entirely irrelevant unless an initial condition is pertinent to handling, say.

There is a further, commonly unrecognized, factor for "strength"-related determinations: scale, *alias* test piece size or dimensions. For all materials there is a scale at which there is a change from plastic or ductile behaviour to brittle (ascending). All materials are plastic on a small enough scale, even alumina (4, pp. 752–753). Accordingly, it is necessary to consider the sizes of the objects in service to ensure that there is reasonable mimicry of scale (bearing in mind the exactness problem alluded to under *Mode of challenge*, above). However, should the critical transition scale be comparable with that of service, two more tests might be advisable – above and below that point as well as close to it. The conditions of use are frequently not tightly controllable in practical terms, even if specified (e.g., mixing ratio).

The Griffith Criterion, relating nominal stress at failure to flaw size and surface energy, applies primarily to brittle materials, but can be extended to include plastic deformation (4, p. 768). Thus, the structural scale of the material (e.g., grain size) has a bearing on behaviour, as does the presence, type and shape of flaws, whether internally or surficial. The processing (preparation) of test pieces must therefore take this into account such as to represent as closely as possible the state of the material as used in service, including surface finish where this is relevant to failure initiation.

Furthermore, many materials have chemical reactions (setting) that continue for some time after preparation, some may absorb water (affecting T_g in polymers and thus all properties, including size), some may further react with that water or components of the environment (e.g., CO_2, phosphate), some may lose some component (dissolve or leach). These effects are time-dependent.

We may therefore list some conditions of relevance as follows:

(a) temperature
(b) strain rate
(c) aqueous environment (water, solutes)
(d) surface condition (roughness, texture)
(e) internal structure (grain size, flaws)
(f) history (preparation method, time, temperature, environment)

There may be more. Each should be checked and controlled for relevance, required accuracy and practically-achievable precision. The demands of research are commonly much greater than for quality control, not less.

Accelerated tests

There are circumstances when long-term behaviour is needed, whether static exposure or fatigue loading in some sense. It is essential that comparability of outcome has been established to justify and validate the accelerated test – it is an assumption that the same processes will occur, just at a greater rate. Thus, for example, in static terms raised temperature accelerates reactions in general, but it also alters the properties of organic polymers with respect to T_g and thus free volume, and therefore water uptake, chain segment diffusivity, interfacial stress in composites, and so on. Likewise, fatigue testing needs attention being paid to strain rate where there is sensitivity to this, but also the potential for temperature rise due to work hysteresis – there are few ideal elastic solids in that respect in dentistry. Selection of duty cycle must also recognize what happens in the mouth – reversed stress might not occur, for example.

Similarly, lower pH may produce faster reactions, but care needs to be taken to ensure that they are similar in kind as pertains to actual service lest misleading interpretations result. There are other possibilities, such as irradiance, flow rate, tonicity.

Controls

Perhaps less feasible as a generality in many mechanical tests, some thought might be given to validating a test set-up by reference to the behaviour of known, stable, reproducible systems. Plainly, if a system cannot be relied on to give reliable outcomes, the problem becomes one of a reference system such that a relative value can be determined. The provision of a reference material and, by implication, a highly-reproducible test piece preparation process, has its own challenges. There is little evidence that this is done in practice except through comparison with earlier results for well-known products, where the same errors or defects may have similar effects on the outcome.

The properties

Compressive strength

The long-standing practice of axially loading a right circular cylinder to the point of collapse, when the calculation of force per unit cross-sectional area is used to yield a so-called "compressive strength", is superficially simple and easy to perform. It has several problems.

Firstly, there can be no such thing as "compressive" strength because it is not a feasible failure mode for solid bodies, i.e., in compression (5). Porous bodies may, under some circumstances, densify, and network or skeletal structures may collapse by local buckling of struts, but ordinarily rupture into several parts, sometimes explosively (most especially with ceramics and other brittle materials), is what is observed. Indeed, collapse is initiated in shear at the tip of "shear cones" formed at the ends. It might be useful in some circumstances outside dentistry to speak of the "load bearing capacity" of a column, but this is not relevant here.

Technically, this test is in fact quite challenging. It requires that the end faces are accurately plane parallel (and normal to the long axis) to avoid stress concentrations; this may be hard to achieve without special care in preparation. Grinding end faces by any means except a precision machine fails to reach the standard required and cannot ever be recommended. It also requires the load platens to be accurately plane parallel, and again normal to

the load axis, but the maintenance of this condition is onerous. Any test material that is at all abrasive with respect to tool steel or the like platen material (such as many filled resins, cements, and certainly ceramics) will slowly but steadily produce an indentation at the location site, immediately spoiling the required conditions. "Strengths" are then always lowered.

Recognizing the above difficulties, many resort to placing a slip of filter paper, sometimes wet, between each end and its platen. This fundamentally changes the behaviour of the system (5), so much so that the failure mode is changed, sometimes to vertical slabs, with very different loads at failure. The effect of "parasitic" stresses can be substantial. Indeed, it is argued that the friction at the contact areas in the unpadded test invalidates it in the first place because of the constraint it represents, yet providing zero restraint is essentially impossible, even with lubricants. The only viable approach in this respect is to use short cylinders of the same material as "padding" such that there is no lateral shear stress at the interfaces. This doubles the problems of achieving plane planarity, but adds the problem of accurate co-axiality.

What is sometimes overlooked is that the load at failure varies with the length of the cylinder. This is enough to betray the idea of a material property being measured. It does mean that for uniformity in outcome constant length between test pieces, within and between materials is critically important if comparisons are to be made or criteria set. Thus the tolerance on test piece length in ISO 24234:2021(E) *Dentistry – Dental amalgam* (6.5.4.3) of (8 ± 1) mm represents a substantial source of scatter beyond that of the material itself. Indeed "After ejection, the test-piece shall not be trimmed." It is also noteworthy that the method in that standard does not represent clinical methods.

The fabrication of cylindrical test pieces in many materials is also challenging. Avoiding the incorporation of large (and often invisible) defects by the usual "top loading" of a mould is very hard, and while it is often feasible to use a "bottom loading" method the risks are high. Any visible defect on the wall of the cylinder invalidates its use; this is not often done.

Constraint at the end faces results in the phenomenon of "barrelling" – a radial bulge at around the midpoint (irregularity would of course betray internal problems). This of course invalidates the assumptions of the "strength" calculation (this is more than Poisson strain), but it is especially noticeable with more plastic materials. The deformation results in a change of length at the same time – the effective axial ratio has changed in a way that depends on several properties, hence the values obtained depend on other than the "strength" of the material.

It is sometimes argued that because there is correlation between the values obtained for tensile and compression testing, deriving from a consideration of the Griffith Criterion, and since tensile testing is often very difficult for the material of interest, especially when brittle, then compression testing is a useful surrogate. The problem here is that correlation, depending on material and circumstances, results in a factor of anywhere between about 8 and 20, for "compressive" over tensile strength. This depends on the internal friction of the material (4, p. 751). There is no constancy.

Broadly, it is essentially impossible to interpret a "compressive strength", given that failure is in fact in shear and that the shear stress cannot ever be known from the test behaviour.

Nevertheless, neither this mode of loading nor the manner of collapse occurs in practice in dentistry in any context. It is therefore entirely irrelevant and offers no basis for a fundamental material characterization in any material.

Diametral tensile strength

For various reasons (5), the loading of a right circular cylinder across a diameter has been assumed to yield a value for the tensile strength of a material. One of the main claims is that a Hertz contact stress analysis indicates a pure tensile stress across that diameter. In fact, such a test is in essence the same as the "compressive strength" test: failure occurs in shear along the apices of a pair of shear wedges. However, the coplanarity conditions for the platen contacts are made more severe by the length of the test piece; padding again does not help but confound the situation. Here now the platen contact is by definition a line – zero area, and thus represents a singularity, an infinite stress, which is impossible to sustain. This was ignored or overlooked by Hertz. Even in brittle materials this results in some flattening at the contacts (see the point on scale under *Test conditions*, above) – and this can be in the platens themselves, whether elastically or plastically (the latter wrecking the platens for any further use). This means that the shear wedge now depends on the yield point of the material (or of the platen), hence the load at failure also varies in a non-regular fashion. If there is flow, then it is time-dependent: cross-head speed matters. The equivalent of barrelling also occurs, of course, with sufficiently plastic materials. There are also changes in the behaviour due to the transition from the plane stress to plane strain conditions as the test piece is changed from a thin disc to a longer cylinder: axial deformation matters.

Thus, even if tensile strength is ever relevant in dentistry, this is not the way to measure it. Without doubt, the method does not yield a "tensile" strength. However, there is no such load system operating in any context. There is no mimicry in this crucial regard. This too is an unviable test. In fact, there are many variants of uniaxial loading tests, none is meaningful (5), except as providing a measure of "load bearing capacity" of an actual device or structure. Again, none are known to be relevant in dentistry.

Shear strength

There are only two fundamental strengths, tensile (Mode I) and shear, with shear being either in-plane (Mode II) or out-of-plane (twisting) (Mode III). It is commonly asserted that for many systems in dentistry it is appropriate for "bond strength" to be tested in shear (implicitly Mode II) because that is how failure is envisaged to occur. In fact, (pure) shear failure in dental contexts is extremely rare, common speech usage notwithstanding. The only kind of context where it could occur is illustrated by differential thermal expansion of a veneer on a planar substrate with a different expansion coefficient (and that substrate must be relatively massive to avoid appreciable bending – it can never be zero). As far as is known, this is not a clinical situation or occurrence. Close consideration of the actual failure of bonded

systems (without prejudice as to the cause of the "bond" – chemical or mechanical) shows that they are all in effect "peel" failures [e.g., (6)]: a crack is initiated at some critical point and grows through the interface, essentially in tension, even if some shear stress is present it does not dominate. Given the need for the mode of failure in a test to mimic that in a service failure (see *Mode of failure*) for the result to be meaningful, a shear test can only be contemplated when such failure has been documented. To date it has not. It is difficult therefore to understand the drive to test in an irrelevant fashion.

Even were such a context identified, there are severe difficulties in realizing the envisaged stress. It has been shown that in the frequently-used chisel edge, wire loop, and numerous other similar, but randomly-created arbitrary set-ups, stress concentration at the contact, displacement of the point of contact from the substrate face, thickness of the cement, and differential elastic modulus distortions (as well as several other matters of detail) all mean that pure shear at the (or either) interface is never achieved, failure is initiated at a stress concentration (and therefore unquantifiable), and separation is in effect by levering the components apart. The interfacial area is not directly relevant as a scale factor. Saying what is desired does not make it so.

As with bearing capacity (see *Compressive Strength*), if a circumstance occurs in service that has relevance to the ability of the system to withstand the challenge, then mimicking that challenge is a reasonable proposition. This is provided that it can be shown that the failure mode is indeed relevant, and a suitably abstracted and standardized approach can be designed, whether or not a theoretical or otherwise strict interpretation of the figure of merit thereby obtained is possible. The goal remains one of quality control.

In fact, there are many difficulties in creating a pure shear test in systems outside the dental context: it is a very challenging field, despite much interest from several points of view. Broadly, the differential deformation of substrates commonly spoils alignment. Even when that is compensated the in-line non-uniformity of the tensile stress in the substrate (Mode II) means that a uniform shear stress cannot be obtained: elastic modulus is always finite. Again, if a service condition is mimicked, the bearing capacity can be studied, but theoretically fundamental values may remain inaccessible.

Thus, if "retention" of one material on another is of interest (and clearly it is in several dental contexts), then the relevant load and failure conditions need to be identified and mimicked, whether through an (attempted) exact or abstracted model.

As it stands, there is no shear test in dentistry in any context that withstands scrutiny for what is claimed, yet it may be that there is a "retention" measure that is of use. The problem is that wishful thinking prejudices observation and interpretation.

Tensile strength

The other fundamental strength, as mentioned, is tensile. Testing in this mode would appear superficially to be very straightforward, but there are limitations, often serious. For ductile and low elastic modulus materials such a test is forgiving. Small malalignments due to test piece imperfection, or in the grips, are accommodated. For brittle materials, such as ceramics (assuming such a test piece can be fabricated), it is very difficult indeed. The implied flaw sensitivity means that surface condition is critical: this must represent the surface state of the material in service as closely as possible. In addition, the parasitic stresses due to malalignment become serious to the extent that premature failure may be frequent, but hard to detect as such. Any failure that occurs in or around the grips is invalidating. Increased section ("dumbbell" or "dogbone") is commonly essential, relying on the St Venant Principle, but so too is padding if the gripping action itself is not to be damaging. This is true especially for wires where the section is predetermined and uniform. Universal joints may assist in some respects, but here they may create problems if the resulting load axis is off centre: any bending must be avoided. What is often hard to achieve is a good enough grip without prestressing the test piece in the process, especially when the grip faces are independently adjustable and thus not symmetric and simultaneous in action.

For miniature test pieces, these concerns are especially challenging, and failures during handling, especially for bond tests, may be common. Often discarded, such data are in fact important and should always be reported. (Miniature test pieces cut from one piece of prepared material are of course correlated, not independent. Statistical analysis of such data must take this into account: "within" variance cannot be the same as "between" variance, in general.)

Sometimes such test pieces are pre-mounted on a carrier that preserves their integrity in a handling stress-free manner, at least in principle, until mounted on the test machine, when some kind of "release" is used to enable the test. In such cases it becomes more difficult to control temperature and hydration and thus avoid stresses arising from either change, and indeed provide the necessary test conditions.

It is necessary, however, to return again to the question of relevance. There are no circumstances in dentistry where direct tension actually pertains, even if it is deduced that the primary cause of a failure is a tensile stress.

For wires, there may appear to be no other choice to get a "strength" but again, while not ever observed to fail in this manner in service, tension might provide a useful figure of merit. Even so, relevance ought to be established. What complicates matters is the necking of even slightly ductile metals that means that the cross-section at failure is not the original. There then follows a debate as to which is to be used for the strength calculation, and why. No value for dentistry is discernible, even if there is theoretical interest for development purposes.

Yield point

For metals, including wires, and given that in dentistry none are truly brittle, the controlling property is yield. This is the stress at which permanent (as opposed to elastic) deformation occurs, because that is when the dimensions of the device are irretrievably altered and function may be lost. This applies whether we are dealing with a partial denture clasp, an implant framework, a cast crown, or an orthodontic appliance, and indeed can apply to

polymeric materials, although then other things may be happening as well.

One problem is the definition. Ordinarily defined simply as departure from pure elastic behaviour, this cannot be measured directly. Even where the load-displacement (stress-strain) plot is ostensibly linear, resolution is the problem. However, true linearity is often hard to find and prove, even if assumed at small displacements: the definition of elastic limit is not the same as proportional limit, and that is the problem. The common, pragmatic, engineering approach is to use "proof stress": identify the stress at which a predetermined (but entirely arbitrary) strain offset from apparent linearity has occurred (4, p. 13). It is assumed that permanent deformation has been proven for the system of concern by proof-of-principle measurement of a gauge length after unloading sample cases. The determination of this essentially requires the use of a ruler, or a software equivalent, but plainly has errors different in kind from those of other test methods, and demands an accurate calculation of strain in the first place since that is the index value.

The method used for this is commonly direct tension, using an extensometer on a convenient gauge length. In bending, where cast metals are concerned, the calculation of strain is more elaborate but may be more useful, especially since this mimics the service condition of interest.

One matter of concern is that in general service, once yield has occurred, the device is considered failed. But, in strength tests, any process that allows the load to increase beyond that point (i.e., the material shows any flow or ductility) means that the measured "strength" is then higher, which may then make it entirely irrelevant to the concept of failure in practice. For brittle materials this does not arise. The question then is which is required for assessing the serviceability of a material in its context? It would appear that yield is the only sensible criterion. Indeed, it is for such tasks as burnishing the margins of gold crowns and inlays, for example, when "failure" is not the concern but the goal. Likewise, for the forming of orthodontic wire appliances: the yield has to be attainable with hand tools.

Bend strength

Bending of a bar or sheet of material is perhaps the most common deformation encountered for devices in dental service, whether these are denture bases, bridge or implant connectors, isthmus extensions in occlusal fillings, partial denture clasps, rubber dam clamps, or crown posts; in such cases fracture is a concern. A convenient means of addressing such a property is, appropriately enough, a bend test. This usually involves a rectangular beam and may be addressed using 3- or 4-point bending. With the test piece resting on two narrow cylindrical supports often called "rollers", the load is applied through either one central or two symmetrically-placed "inner" rollers.

The geometrical considerations of the test rig are simple: the rollers must be exactly parallel if twisting of the test piece is to be avoided, the parasitic stresses of which torsion complicate the system appreciably. The same applies to the test piece, of course, and thus the mould or other fabrication technique becomes critical: all lengthways surfaces must be flat and pairwise parallel as well as adjacent faces properly square to each other. There are obvious challenges with materials such as filled resins which may undergo appreciable shrinkage and thus not remain true to the mould when set. Less obvious is distortion in ceramics due to differential cooling rates: viscosity changes rapidly on cooling, and a part that is cooler tends to hold the local size while the remainder relaxes, shrinks further, and generates an internal stress that may distort the final product. This might be alleviated by grinding, but at the risk of generating significant surface flaws. A similar effect is seen in cast metals that must machined to size: the distortion resulting from simple reduction can be countered dimensionally by successively finer cuts. In both cases internal stress may remain (a prestress) that complicates behaviour in the test.

Failure initiates (or is meant to initiate) on the convex (lower) surface, in tension. Accordingly, the surface finish matters very much in the Griffith Criterion sense, and following the precepts above that surface ought to be in the ordinary service state in terms of roughness, texture, and orientation of that texture when it is not isotropic. Corner defects – chips and bubbles – must be avoided. While it is wise to examine fractographically the fracture surfaces to determine the site and nature of failure initiation, this is time consuming and not always practical. But, it can be said that, in a 4-point test, failure at or outside the inner rollers means that the test is invalid, as would also be implied in a 3-point test if the crack is appreciably distant from the centre (in both cases a cause should be sought). The as-moulded state, assuming that the mould surface is smooth and texture-free on a scale less than that of the natural roughness of the material, is commonly the reference condition against which the effects of finishing techniques may be assessed – when these are relevant to service. It is pointless to consider such in cases where the finishing is applied only to the free, compressed surface in practice where cracking is never seen.

The ends of such test pieces are relatively unimportant in terms of surface shape and finish, but while the elementary concept of the nature of such tests implies that the end overhang is of no consequence, in fact the stress field due to the outer supports extends outwards into that overhang and thus it modifies the behaviour inside. It is not feasible to test with the test-piece length exactly the same as the outer roller spacing – the curvature in the test means that the arc length is greater anyway: the intended support would be lost. Whether that overhang is big enough to matter in the face of the usual scatter in results needs to be ascertained, unless the usual approach of the "semi-infinite" approximation is adopted. Here, the overhang length is made great enough that no further effect is detectable on further lengthening. Adding an amount equal to the depth of the beam may be enough; twice that amount would be safe, as a generality.

An obscure problem arises from the effects of Poisson strain in a bent beam: it does not remain rectangular in cross-section. This so-called anticlastic curvature (4, p. 573) means that there is a stress concentration on the corners at the outer rollers, and likewise a stress concentration at the centre line contact under upper rollers. The tensile stress on the convex surface in the effective test zone is also concentrated on the edges, making their state all the more critical, and especially so the larger the deflection. This problem is not known to have been investigated.

The key distinction between 3- and 4-point bending is the volume of the material effectively being tested. The tensile stress in 3-point bend is theoretically at a maximum only on a single line, although the critical zone is considered to extend inwards and sideways somewhat. The calculated failure stress thus depends on the chance of finding flaws along that line and nearby, i.e., given their bulk frequency and size distribution, remembering the nature of the Griffith Criterion: stress × size$^{1/2}$. In a 4-point test, the effective volume tested depends on the separation of the upper rollers and thus is much larger than that under test in 3-point bend. Such strengths are therefore not surprisingly lower on average than is found in the 3-point test for the exact same material. This is not a defect of the 4-point test – it reflects the statistical nature of failure in a realistic way. Comparison of the two kinds of result is therefore rather pointless. The only question is which is better for assessing service behaviour? Does a 3-point condition occur in practice?

It can be noted further that the 4-point test may itself have variation in outcome according to the separation, c, of the upper from the outer rollers in relation to the full span, L: $c = L/3$ is common, but $c = L/4$ is also used: the ratio is arbitrary and odd values may be encountered. None is definitive, but all give different results, on average, because of the volume effect.

Similar remarks apply to biaxial flexure, where a central load, commonly through a ball, is applied to an edge-supported disc. Again, the results differ because the tested volume is proportionately larger, although edge defects are entirely avoided. However, the size of the edge overhang still affects the stress system. There may be a correlation between uniaxial and biaxial results, but no strict "conversion" is possible.

Very often beam test pieces moulded from more or less fluid materials end up with a "flash" due to mould overflow, or sometimes due to a poorly-closing mould. No attempt should be made to remove such flash completely if it is on the edges of the tension surface: the risk of creating edge flaws that will prove critical is too great. The extra material involved is usually so small as to have no discernible effect on the final results, especially if this extends laterally. In fact, it would be better on the upper, compression surface, assuming the other face is flat enough to be tested sensibly.

The moulding process itself for such materials may cause problems. It is known that filled resins contain bubbles (presumably oxygen-free) as supplied, while powder-liquid systems such as cements frequently incorporate air in the mixing process, and especially if close to the "minimum mixing liquid" (4, pp. 44,46,443). If the mould is "closed" (only open at the top), is overfilled and covered by a plate that is then compressed, there are two complications. The flash tends to be greater because flow through a narrow gap of a viscous material gets to be very slow, but also the bubbles are themselves compressed, providing an internal and unfavourable prestress that may compromise the measured failure stress. The material cannot be relaxed. Indeed it may be impossible in such circumstances to obtain sharp lower edges because the trapped air cannot escape. Such difficulties can be overcome by using a knife-edge mould that minimises the resistance to sideways flow and allows a good cut-off under the closing plate, but a mould that is in fact open at both ends, allowing lengthwise flow and thus full relaxation (7) (for which a little time may be required for a viscoelastic material). A further benefit is that tabs are thereby created at each end that allow handling, especially with tweezers for example, without risk of damaging the test length.

"Rollers" are in principle meant to avoid lengthwise frictional constraint creating parasitic stresses. In practice they can never turn, even if nominally meant to, because the "bearing" resistance is too great under load. The diameter of the roller can compensate at the outer supports as the test piece rotates around them, but this requires a close match of the calculated movement, allowing for the imposed curvature and the tensile expansion of the lower surface of the beam, the depth of the beam and the test material's modulus of elasticity (i.e., related to the deflection), to the roller circumference traversed. Accordingly, results may not be comparable if any of these things vary (i.e., between reports). Nevertheless, friction at the rollers is inevitable and thus so are sources of parasitic stresses. Again, we end up with a figure of merit rather than a true material property.

Friction and movement at the rollers also means that wear is inevitable, and the more so the harder the test material. That friction will then change with time. As with the platens in compression testing, inspection of the rollers should be a routine, and replacement made as necessary. If the profile changes there will also be widthwise parasitic stresses. Consideration might be given to lubricating the outer contacts (and also the inner for 4-pt bending) with a high-pressure grease.

A further source of difficulty is the span-depth ratio. The simple beam theory equation for strength, as with all such, assumes unchanged geometry. Obviously this cannot be true, but the deviation becomes greater with the deflection (which affects the parasitic stresses due to roller friction, and so on). To control this, the span and depth of the beam may be adjusted, but the range of sufficient validity is limited. Because bend tests involve a shear stress across the depth of the beam as well as the longitudinal tension and compression, deformations and thus the stress of interest are affected. The ratio of span to depth affects the deflection and thus the calculated strength (and modulus of elasticity). There are recommendations for sufficient validity, and approximate corrections can be applied (7), but again there will be a lack of comparability of results if these factors vary between reports.

Fracture resistance: Hertzian ball on disc

One reason for the use of the "compressive strength" test is the observation that masticatory loads on occlusal surfaces have a large compressive component, which has also led to the widespread use of so-called "crunch the crown" tests where a full device is loaded to failure, but which have been deprecated as uninterpretable at best (8). However, noting that, in broad terms, occlusal fillings and crown materials are loaded *via* a rounded cusp, are supported on a compliant substrate, tend to fail in tension from the lower surface, and do not explosively fragment, a "Hertzian" contact ball-on-disc test has been designed and applied successfully to ceramics (9), silver amalgam (10), and glass ionomer cement and filled resin (11), and been validated in several respects (12).

In outline, the test involves loading a disc of the test material by a hard spherical indenter, the disc resting freely on a dentine-mimic (in terms of elastic modulus). Validity is established by bottom surface-initiated radial cracking, and the absence of top surface-initiated cone cracking (which occurs if the disc is relatively too thick). The appropriate scale of thickness needs to be determined for each class of material.

This test satisfies the general requirement for mimicry in respect of mode of loading and, importantly mode of failure in that two or only a few pieces result, and failure is "quiet" (in fact, sometimes full separation does not occur). It is not in general possible to determine a strength as such because the stress at failure needs to be determined numerically, which is not a practical proposition for routine use. However, the load at failure is a figure of merit that can be used to distinguish between products and details of production. Significantly, the lower surface is "as prepared", against a glass surface say, and needs no treatment, as would be the case in practice. Further, the condition of the upper surface is of little or no concern because it is not involved in the fracture process. The diameter is also irrelevant beyond being big enough – the semi-infinite approximation is used. Indeed, the condition of the cylindrical surface is then of little concern, defects and edge chips have no effect. This means that filling the mould and ejecting the test piece, a common source of difficulty for many "packed" materials, now become less onerous; a tapered mould may be used. The most important region is the central zone of the lower surface. The substrate is cheap. In addition, the geometrical requirements for the test rig are minimal: no alignment problems arise beyond the minimal demand for the lower surface interface with the substrate to be normal to the load axis and that the ball contact is central.

There is some geometrical similarity with the biaxial flexure test, as will be apparent. However, the distinctions are that the test piece diameter is now not a factor (overhang-free results), and that the manner or design of support is not relevant: there are no edge contact effects.

Care again has to be taken in respect of wear. The substrate (conveniently, a glass-filled nylon) needs to be machined flat, as may the test piece lower surface if cut from a stock piece, but it may deteriorate with repeated use. It is compliant enough that a small amount of wear has no effect (as the test piece bends to conform before failure), but it should be replaced if this becomes apparent. Likewise the indenter (a hard steel bearing ball is convenient) will wear with abrasive materials and develop a facet with very hard ceramics. Rotation can deal with that simply by moving to a new unaffected area for contact, but in any case such balls are also relatively cheap and can be discarded as necessary.

Elastic modulus

The stiffness of a material is of wide interest and concern in dentistry. Broadly, all devices are generally assumed to function in their designed shape: simply put, they fit. Deformation from that shape is taken to be detrimental, whether this is the problematic redistribution of stresses, leakage, or just the ability to fulfil their primary purpose. Plainly, no material is infinitely stiff, so the question becomes one of how much deformation is tolerable in the circumstances. Alternatively, for orthodontic appliances, the elastic spring nature of materials is commonly fundamental to their operation: design is based on elastic displacement forces doing work.

Modulus of elasticity (E) is commonly envisaged and measured as the slope of the stress-strain plot in direct tension, and indeed this is the most direct way, again using an extensometer over a convenient gauge length. For wires, some polymeric systems (such as impression materials, denture base acrylic), and cast metals, this might be straightforward in terms of test piece fabrication. Certainly surface condition is less of a problem than in strength testing, as is gripping because the load applied need not be very high, but for cements and filling materials this is not at all easy.

We may then rely on symmetry: at low displacements the value in compression is the same as in tension – it is a smooth curve through zero. This means that it is possible, if indeed the load is kept small, to use a moulded right circular cylinder in axial compression. The problem then is determining strain: the load string compliance must be known so that it can be deducted from the cross-head displacement. Typically, the machine used is stiff enough that the frame and connectors are of negligible concern at low loads (although it should still be checked). This is not true of the load cell, which is designed to be a Hookean spring device with a substantial displacement in order to have the requisite sensitivity and resolution. This is commonly overlooked in many such contexts, but may easily be larger than the test piece compliance. Designs vary substantially; each must be calibrated, not assumed.

An alternative that may be better suited to filling materials and the like is to use a bend test (for what is called "flexural modulus", but it is the same property), where relatively large displacements may be achieved at low loads, giving good sensitivity with small-scale test pieces. With wires, a cantilever arrangement with non-contact displacement measurement can yield exquisite sensitivity. Indeed, such an approach demonstrated that many reports have given suspect results, even allowing for genuine scatter in true values. There are many sources of error that must be considered (13), a principle that ought to be applied to all aspects of any test system, no matter the mode or purpose if consistency and accuracy are to be attained, and preferably with good precision.

Further problems are backlash and "toe". No test machine can be made without some slack in the system if it is to move at all. This means that complete, microscopic reversibility is not attainable: some drive and crosshead movement is inevitable before a reverse load will appear. Whether this matters depends on the set-up, but ignoring it can be a problem. Likewise, no matter how well made and assembled is a test rig, to say nothing of the precision of the test piece and its contacts with the machine (especially under the rollers of a bend test system), it takes some finite load for a system to bed-in and show a reasonable approximation to linear load-displacement behaviour: all machine parts are deformable in a non-linear fashion when the contacts are not planar. This results in a "toe" in the plot. It is real, but an artefact and unhelpful. It must be ignored in calculating the strain. It is also possible for a load-cell to show some "deadband": a small load must be applied before any output is observed. This is meant to be negligible in modern devices, but it cannot be if a chart recorder is used

where slack in the pen drive system is not avoidable. Such effects might also be discernible in any associated electronics (14). Careful checks are essential.

In a bend test, and even if a displacement gauge is used instead of relying on cross-head movement, local crushing under the impossible to sustain line stress under the rollers means that the displacement is greater than expected because an indentation has been created. Effectively, backlash has been created. It can be dealt with by loading first to the maximum intended (with a dwell time, as in an indentation hardness test, if appropriate), then repeating for the actual measurement, given that no further crushing will (or should) occur.

Generally, the Young Modulus of Elasticity is what is calculated, which does not take into account Poisson deformation, using only original dimensions Commonly, this pragmatism has little impact on the use or interpretation of E, especially when the strain involved is small (4, p. 8).

Indentation hardness

Hardness is, in common speech, not well defined: the quality or condition of being physically firm, unyielding, rigid, tough … the degree of resistance to abrasion and scratching (15). Clearly, firmness and rigidity (elastic modulus), toughness (work to failure or fracture toughness), and being unyielding (strength?), conflict with other concepts in the present context, and thus with each other. Resistance to abrasion and scratching is distinct and more relevant to dentistry, but rarely addressed. On the other hand, measurement of resistance to penetration is both relevant and commonplace through what are termed indentation hardness tests, a self-explanatory and operationally-defined property.

In fact, such hardness (H) is generally related in a simple way to yield stress Y (in tension): $H = cY$, where the value of $c = 1 \sim 3$ depends on whether the structure of the material allows collapse (densification) (4, p. 26). Thus we see a direct means of assessing the effective yield point and therefore a value that is perhaps more representative of what matters in practice.

The principle of the test is to apply a probe, the indenter, to a surface, allowing the load to rise to a predetermined value, then on removing the indenter measure the size of the indentation, thus its area and the stress that caused it. Note that this is the inverse of a normal strength test where a (nominally) fixed area is subjected to a monotonically rising load when collapse is then commonly prompt and complete. At first contact of the indenter there is a stress singularity, but the stress then falls quasi-exponentially as flow (material displacement) occurs until equilibrium is reached. This implies that all such tests require time to settle, and a dwell or "hold" time is usually specified for reproducibility. This may be longer for such as polymer-based materials where appreciable creep on a timescale of some seconds occurs.

There are many styles of indentation hardness test. Some (Brinell, Mayer, Pfund) employ a spherical indenter such that the geometry of the indentation varies with depth, complicating interpretation and calculation of a relevant stress. The measured hardness is therefore not a smooth function of load applied but a series of partially-overlapping curves. This lack of scalability means that they are deprecated for general work but can find convenience in simple QC tasks. Other tests use an indenter of regular shape such that the geometry of the indentation is constant with depth. Such tests include the use of pyramidal (Vickers, Wallace, Knoop) or conical indenters (Rockwell) and in principle allow indent size to be a continuous smooth function of load.

The stress has commonly been calculated based on the indentation surface area, but projected area is now recognized as more proper, although for regular shape indenters this only amounts to a small scale factor difference and has no great significance in practice. But, it is to be noted that indentation hardness, however defined, is a stress, and the units need always to be specified, with the type, load and duration.

The indentation size is normally measured optically, taking the mean of two diagonals, for example in a Vickers test, to calculate the area. Regular geometry also means that depth can equivalently be measured automatically by the instrument, avoiding uncertainty as to the location of the end of the diagonal, which might not always be clear. Larger indentations are of course better in terms of the precision of either measurement (and an error band calculation made on that basis), but any corner cracking invalidates the test as the assumed conditions no longer apply. Materials with appreciable recovery can obscure the diagonals' ends, especially polymer-based materials; measurement of indentation depth can be preferred in such a case. Nevertheless, considerable care is always required in measurement, with good lighting and magnification, to get the most accurate data (assuming that an eyepiece graticule, if used, has been properly calibrated against standard scale).

The approach seems to be relatively straightforward, but there are a number of problems that must be dealt with for reliable results. Firstly, vibration must be avoided. The vertical motion from a vibrating floor or bench means that the indenter is hammering the test material, making the indentation larger and the material appear softer. This can occur from local footfall, vehicular traffic, a nearby lift, or other laboratory equipment, and is very common. A vibration-proof (balance) bench is recommended. Equally, impact is to be avoided so the lowering rate must be controlled. For materials that have a surface texture (especially roughness), the optical measurement of indentation size can be difficult unless this is of the order, say, of some 50 times that of the texture scale; indeed the hardness value will be suspect otherwise. Likewise, if the material has a composite structure, the indentation must be large enough to average effectively over many grains (but large subsurface pores can lead to faulty results anyway). This applies to porous materials as well. Thin sheets cannot be tested because of the anvil effect – support from the substrate platen: the thickness must be at least 10 times the depth of the indentation. Similarly because of the strain field around an indentation, they must be sufficiently spaced as to not interfere with each other – 4 times the diameter or diagonal is usual. As indicated, for materials that show appreciable flow, the dwell time must be chosen to be sufficiently long to allow full settling. Needless to say, the indenter must be clean: some materials will stick and leave a film or fragments behind that compromise the next test. Of course, since flow past the indenter is involved, frictional retardation must be present. This implies that lubrication, whether deliberate or unintentional, will have an effect and must be considered. But, as with all tests

that create new surface, if the conditions of interest include water, this must be present during the test and materials not allowed to dry.

Despite some similarities in behaviour, it is not in general possible to interconvert hardness from one system to another in a consistent or reliable fashion.

One advantage of indentation hardness is that a number of tests can be made on a single test piece (provided the spacing stricture is observed) allowing both spatial variation to be assessed and better statistical information. But this then implies that variation with time can also be monitored on the assumption of a material undergoing change such as setting or water sorption. This would provide a sound alternative to tests such as the Gillmore Needle (see below) and permit a more rational, operational definition of setting time – say, 95% of the ultimate value. On the other hand, if a rational minimum strength were to be identified, the time taken to achieve this could be measured in a simple and material- and time-economic fashion. This might need a large-scale indenter (in tool steel, for example), such as Vickers pattern, but for full mapping the load can be changed as required to avoid indentations being too large or too small – continuity is assured if the structure scale and thickness constraints are honoured.

Flow

The ability of a paste-like material to flow to fill spaces and conform to topography is a fundamental concept for the function of many materials in dentistry, such as filling materials, cements, sealers and the like. Technically, characterizing flow implies both a qualitative and a quantitative description of the rheology of the material. In the absence of bubbles, the elastic component of the shear deformation involved in flow is entirely negligible because the bulk modulus of liquids is so very high in comparison with the stresses employed. Further, the retarded compliance is also generally negligible before network formation, and while some systems might end up stress-frozen when setting proceeds far enough while the flow is being driven, recovery is generally not a concern because the material is then constrained anyway. In other words, what is sometimes termed the Newtonian (viscous) component of the behaviour only is of interest. Measuring this properly requires elaborate equipment and technical expertise while the dental need is for a pragmatic QC method. Accordingly, the focus is changed to the net effect by measuring the final spread of a quantity of material compressed between parallel plates under a static load, as in what is called parallel-plate plastometry (16).

This is not an unreasonable approach given that many fluid systems involving large orientable molecules may be appreciably pseudoplastic (lowered viscosity at high strain rates), but also powder-liquid systems (such as cements) may have an appreciable yield point: a minimum stress is required for flow to occur at all. Further complicating this is the fact that such mixtures, at proportions considered desirable on grounds of other demands such as strength, i.e., a large volume fraction of solid, are typically dilatant. Thus the viscosity increases with shear strain rate due to particle interactions – collisions and jamming. Given that both the shear stress and shear strain rate vary continuously in such a set-up, any proper analysis of behaviour as a function of time is essentially impossible, but equally sufficient time must be allowed for the system to come to rest for reproducibility of outcome, given the sensitivity that may be inferred for many factors.

There are a number of factors that require tight control. Clearly, temperature must be well defined, uniform and stable. The sensitivity of polymeric systems to temperature is very great, being close to a negative exponential (4, p. 78), but even water has a negative power law dependency (17), implying that aqueous systems are also sensitive. The parallelism of the plates is clearly critical. The flow rate in a Newtonian system depends on the separation to the fifth power (18). In a dead-weight loaded system with a "loose" upper plate, as is commonly used for such work in dentistry, this must be accurately parallel at the outset. Maintenance of parallelism then requires a perfectly radially-symmetric initial blob of material, very accurately centred with respect to the upper plate, and the dead-weight similarly accurately placed. This is assuming that the blob is completely homogeneous so that its rheological properties are too. These conditions are not sensibly approachable by hand. It is therefore essential that the upper plate be constrained to remain parallel to the lower plate throughout the test, that is, not have the freedom to tilt (nor rotate). If the parallelism is constrained, the blob placement becomes of no great importance, but its shape remains important. Naturally, the weight of the upper plate (and its guiding system) must be taken into account in the defined load to be applied, and friction in that system negligible.

The problem with hand-mixed materials is ensuring that mixing is complete, i.e., that the mixture is indeed homogeneous. If it is not, then the volume fraction of solid varies from place to place, with severe effects on the local behaviour. There are simple guides to mixing properly based on the scale of the mixture and the scale of the particles (4, pp. 445–446), the principles being very general but commonly ignored if understood. Note that the limits of visual acuity prevents clear observation of the condition, even with strong colour contrast – which aid does not apply to materials relevant here.

However, the assumption is that the proportioning has been done accurately. Often, the liquid component is dispensed by drops, and the powder by a scoop. While this may be a convenience in practice, for test purposes it is inadequate except to demonstrate the scale of variability due to such crude techniques. It is therefore essential for material characterization purposes to weigh both components (while avoiding appreciable evaporation of liquid). Similarly, the use of a volumetric approach for the blob to be tested is fraught with difficulty: packing a syringe can incorporate bubbles that subsequently are included in the blob, the resolution of the volume delivered from such a device is very poor, and clearly the calibration of the volume markings needs to be validated beforehand. For small volumes, the difficulties become greater very rapidly. A fixed mass is plainly a better method for precision and accuracy, but of course the end result depends on the volume and this implies in effect a density measurement, to calibrate the mass to be used. It might be that dispensing by volume provides the simplest (if crude) method, but then scaling the final result by the mass (determined after the test) would provide a good standardization scaling of the measure; this requires a calibration curve to determine the form of the relationship, assuming that final thickness is not constant.

Density measurements offer their own difficulties because of the need to determine volume in the absence of bubbles; a careful calibration is essential.

The shape of the blob also has a bearing on the radial flow pattern observed: boundary circularity can hardly be expected. It would also be complicated by air bubbles, of course – negligible viscosity. Large voids should be visible at the end and their size perhaps be used as a criterion to invalidate that test. Clearly, an elongated blob cannot be self-correcting, and too great a divergence from circularity can be considered disqualifying, albeit perhaps unnecessarily.

The manner of loading also needs to be defined and controlled: too fast and dilatant materials may lock up, but pseudoplastic ones overspread; too slow and the reverse may occur. As with indentation hardness testing, vibration must be avoided as the additional forces may be enough to modify the outcome.

The lack of circularity means that identifying diameters for averaging is required, as the test is ordinarily envisaged. However, image analysis techniques are now commonplace and cheap such that an area measurement could be made quite easily. The area can then be scaled by mass, as above, to yield a more precise measure of the outcome. However, if this is acceptable in principle, there is an easier way: measure the thickness of the material at the end of the test. This can be done to a high resolution by a displacement gauge (non-contact, for preference, as dial gauges and LVTDs apply a variable load that is hard to control for) on the upper plate system. Indeed, measured image area, thickness and mass provides all needed information for calibration.

Finally, the plates themselves need to be scrupulously clean and of course unworn. There is a risk that some materials will exhibit slip at the interface, affecting the bulk flow (16), hence the avoidance of residues matters, in both directions. Many cement powders and fillers (zirconia, as prime example, but most others anyway) are abrasive with respect to glass, and minute scratches are to be expected. Any visible sign of wear should require that the glass plates be replaced, for the avoidance of doubt.

The measure itself can only be regarded as a proxy for the outcome of a set of possibly complex behaviours under artificial circumstances. It can have no direct correlate with any actual service condition or behaviour. The value chosen as a criterion of acceptability is therefore entirely arbitrary. It can only be judged in relation to the behaviour of actual products that on general grounds are considered to be acceptable in clinical handling, and thus by extension the value of this "property" is also acceptable. There do not appear to be any direct operational criteria. Likewise, the usual use of a fixed time for the observation of the outcome, as opposed to a final rest value, introduces a further arbitrary aspect – the time chosen. If this is appreciably longer than the time taken in the clinical procedure, which at best would not be very long – seconds rather than minutes – the distance from clinical interpretability increases. Since in at least some cases the setting process occurs on a similar timescale, it is difficult to judge the importance of the measure. Clearly, the loading that occurs in the clinical use of materials such as sealers may be on a very short timescale – possibly applied rapidly rather than gently. Further consideration needs to be given to the clinical relevance and interpretability of the results of such a test, and other conditions applied, such as temperature, which is commonly just uncontrolled ambient, unspecified. Reproducibility cannot be assumed.

Film thickness

Considered to be a critical factor for luting and sealing materials, this measure can be seen to depend primarily on solid particle size maximum and how dilatant the material is: its tendency to lock up creating load-bearing structures on a scale greater than the typical particle size. It is thought to be important because thin lutes are stronger in the structures they create, that in the case at least of crowns, veneers and the like better final occlusal dimensional accuracy is obtained, and that resistance to erosion and washout is greater.

Film thickness is measured in a very similar way to flow in that a portion is placed between parallel glass plates and loaded for a fixed time. The key difference in the usual embodiment is that the glass plates are small in relation to the volume of material such that it is extruded all around the periphery (and much smaller than are used for flow). Once the gap is filled to the edges, the rate of change of separation depends on the cube of the gap size and the fourth power of the radius of the plates, when these are circular (18). This means that this value must be tightly controlled for consistency. It also means that the edges of the working faces of the plates must be sharp ("square") for that value to be meaningful.

The blob size and shape do not appear to be critical, provided the gap is actually filled by the end of the test. The material thickness is readily measured or calculated. Most of the other requirements remain to be observed and controlled: parallelism, mixing ratio (if relevant), load accuracy, load rate, cleanliness, wear and temperature. Bubbles might not seem to be a critical concern, but as they affect flow overall they must be avoided. The plates themselves must be accurately sized, and accurately aligned so that the area between them is accurately defined.

Square plates are supposed in some quarters to be equivalent to circular if the areas are the same on the grounds that the (final) stress is then identical. However, the flow pattern is different enough as to affect the closing velocity (19), and thus when the material is strain-rate sensitive, as must be the case in general in the context (as for flow, because it depends in part on flow), the results may vary. It can be shown numerically that the closing velocity difference is about 3% lower in the case of the square plates (at any thickness). The assumption of equivalence is therefore flawed. The sensitivity of the outcome to this needs to be ascertained.

While superficially film thickness would appear to be an absolutely bounded character (unless solid particles were crushed, which remains a possibility for some materials), in practice the rheology in relation to the volume fraction of solids (i.e., mixing ratio for relevant materials) and their particle size distribution and shape, specifically the dilatancy of the mixture, has an overriding control. Considerations of loading rate, temperature, setting process effects (the rheology changes and reacting particles get smaller), and time of observation remain open.

Setting time: Gillmore needle

The eponymous Gillmore needle was introduced in the 19th century by a US Army engineer concerned with the setting of Portland cement. It followed similar approaches used for similar purposes over about the previous 200 years: many variations of tip size and load (the term "needle" is misleading: it is a right circular cylinder, i.e., with a plane end, albeit carrying a weight). Broadly, it was taken as a means of ascertaining the development of strength through what can be seen to be a load bearing capacity test (as discussed above, *Compressive strength*), but in a punch on plane surface mode: it is a punch strength criterion. The practical criterion is the appearance or not of a noticeable residual mark after the test. It is widely used in dentistry for the determination of "setting time". There are two sizes of device, supposedly to identify "initial" and "final" set, the latter stress 16× that of the former. It is still in use in ASTM C266 for cement and the like.

The primary problem is the definition of setting time. If the goal is to know in an operational sense when sufficient strength has been attained to withstand some challenge, such as when the next task is to be performed, where an expected load can be identified, there is a practical interpretation. In that sense its use could be justified. However, if it is to ascertain the progress of setting in a chemical sense it is being misunderstood and misused. It is clear that, even with Portland cement, both the final strength and the time taken to get there vary: the rate of development of strength varies. It is easy to see that, in dental contexts at least, a material that never reaches a sufficient punch strength to bear the stress involved cannot be said to set (4, pp. 63–65). Accordingly, if the test or its like is to be used at all in the present context it is essential that the criterion of required punch strength be identified with respect to the service conditions: a pragmatic approach to a practical context. We might then refer to the material being "sufficiently" set for the work to continue without compromising the integrity of that material. As far as is known, such a definition of required strength has never been made for any dental material to which this test has been applied. The assumption has always been simplistically that the result is meaningful with regard to actual setting, which plainly it is not. Bearing in mind that many materials may be operationally-functional long before reactions slow to imperceptibility (in most if not all dental systems there is no absolute end point), the pragmatic interpretation of "sufficiently set" would be better – a "setting time" it never is.

Assuming, then, that a sensible criterion can be set, which might involve adjusting the stress applied up or down from an ASTM-defined value, there are several factors that must be carefully controlled to obtain a useful result. Firstly, the stress applied: the tip diameters are defined as 1/12th and 1/24th inch respectively. For 1% accuracy in the tip area, the precision of the machining must be ∼±10 μm and ±5 μm respectively. "Equivalent" 2 and 1 mm tips clearly fall outside those limits, and if used the mass of the assembly must be adjusted accordingly – assuming that the original specification is to be followed. For comparison, the tolerance on tip diameters in ISO 6876 *Dentistry — Endodontic sealing materials* is ±0.1 mm. The stress tolerances are therefore ∼±5% and ±10% respectively, for precise masses. The tip face must be accurately planar with a perfectly square edge, no rounding. In addition, the test surface needs to be accurately planar, and the tip lowered precisely perpendicular [as illustrated (20)] to avoid stress concentrations. Hand-held devices are unacceptable in the context, although commonly encountered. Duration of load must also be specified and controlled because this is in fact a form of indentation hardness test and a material with appreciable flow or creep will mark according to the time allowed. Even an operational definition of "sufficiently set" does not automatically mean no flow. Furthermore, the test bench must be vibration-free, as even slight "hammering" can produce a mark. The lowering must also be impact free, which by hand is very difficult indeed, even with a guide.

Even under such controlled circumstances, many materials will show a mark at any time after "setting" because they have or develop a surface texture that means that there will be sufficient local stress concentration as to flatten that texture. Accordingly, the criterion must be adjusted from "no mark" to "no further change". Systematic location and timing of the succession of trials allow such a judgement to be made in retrospect. Even in the absence of texture, a sharp-edged indenter means that there is a stress singularity implied, which stress cannot be sustained of course, meaning that a mark will be made in principle on any material at any load, if of low enough relative hardness. Whether it is visible depends on the compliance of the test material and the means of observation – microscopy is never used. This is an unsatisfactory ambiguity. Such an edge will, of course, then wear, even if handling is sufficiently careful as to not let it be damaged by inappropriate contacts. Replacement at intervals is implied as a necessity.

In summary, a punch bearing capacity test could be defined in operational terms, but this requires a determination in advance of a "satisfactory" strength to be identified for materials in a specific context of use. This has never been done. In addition, the conditions for performing the test satisfactorily while tight are not all onerous, but must be observed for reproducibility. Even then, there are practical difficulties.

Conclusion

It can be seen from even this brief excursion that there are very many aspects to mechanical testing that may have escaped attention in the course of routine use of commonplace tests. Most have considerable effect on the validity, reproducibility and interpretability of the outcome, especially when this is a crude Figure of Merit masquerading as a material property or an acknowledged proxy for a behaviour seemingly of interest in a clinical context. It can be seen, however, that there are ways of avoiding many dangers, with appropriate care and attention to detail. Of course, there are other tests, and several aspects that have not been included above (21), but it is hoped that the kinds of problems, concepts and treatments mentioned in this personal overview will enable more careful thinking about how to get useful results. There is no easy answer to any such problem, but assuredly if the science of dental materials is to advance, whether for development or clinical purposes, the effort must be made. The situation in this regard has barely changed since first remarked (22). It may well be that issue will be taken with some of the

points made here, whether of kind or in detail, whether through errors of commission or omission. I make no claim to being comprehensive or rigorous. The point is to make the effort. I am still learning.

Author contributions

Conception, writing, editing, approval: BWD.

Acknowledgments

I am grateful to the reviewers for several useful suggestions and the identification of some infelicities which have allowed me to improve the text.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Darvell BW. The first three questions [guest editorial]. *Aust Dent J*. (1995) 40 (6):397–8. doi: 10.1111/j.1834-7819.1995.tb04840.x

2. Darvell BW. Misuse of ISO standards in dental materials research [guest editorial]. *Dent Mater*. (2020) 36(12):1493–4. doi: 10.1016/j.dental.2020.10.014

3. Schmalz G, Watts DC, Darvell BW. Dental materials science: research, testing and standards [guest editorial]. *Dent Mater*. (2021) 37:379–81. doi: 10.1016/j.dental.2021.01.027

4. Darvell BW. *Materials science for dentistry*. Woodhead: Cambridge (2018).

5. Darvell BW. Uniaxial compression tests and the validity of indirect tensile strength. *J Mater Sci*. (1990) 25:757–80. doi: 10.1007/BF03372161

6. Gale MS, Darvell BW. Controlling dentine penetration in computer microleakage tracer mapping. *J Dent*. (1997) 25(2):129–36. doi: 10.1016/s0300-5712(96)00020-6

7. Musanje L, Shu M, Darvell BW. Water sorption and mechanical behaviour of cosmetic direct restorative materials in artificial saliva. *Dent Mater*. (2001) 17:394–401. doi: 10.1016/s0109-5641(00)00097-x

8. Hooi P, Addison O, Fleming GJP. Strength determination of brittle materials as curved monolithic structures. *J Dent Res*. (2014) 93(4):412–6. doi: 10.1177/0022034514523621

9. Dong XD, Darvell BW. Stress distribution and failure mode of dental ceramic structures under Hertzian indentation. *Dent Mater*. (2003) 19(6):542–51. doi: 10.1016/s0109-5641(02)00103-3

10. Darvell BW. Effect of corrosion on the strength of dental silver amalgam. *Dent Mater*. (2012) 28:e160–7. doi: 10.1016/j.dental.2012.06.001

11. Wang Y, Darvell BW. Failure behaviour of glass ionomer cement under Hertzian indentation. *Dent Mater*. (2008) 24(9):1223–9. doi: 10.1016/j.dental.2008.02.002

12. Baig MS, Dowling AH, Fleming GJP. Hertzian indentation testing of glass-ionomer restoratives: a reliable and clinically relevant testing approach. *J Dent*. (2013) 41:968–73. doi: 10.1016/j.jdent.2013.04.004

13. Tian K, Darvell BW. Determination of the flexural modulus of elasticity of orthodontic archwires. *Dent Mater*. (2010) 26:821–9. doi: 10.1016/j.dental.2010.04.007

14. Darvell BW, Lee PKD. Gould frequency converter: zero error correction. *Lab Practice*. (1987) 36(3):82.

15. Oxford English Dictionary. Available at: https://www.oed.com (October 21, 2022).

16. Peek RL. Parallel plate plastometry. *J Rheol*. (1932) 3:345–72. doi: 10.1122/1.2116499

17. https://water.lsbu.ac.uk/water/water_unexpected.html (October 21, 2022).

18. Murray MD, Darvell BW. A reappraisal of the physics of denture retention. *Int J Pros*. (1989) 2:234–42.

19. Cheneler D. Analysis of a coupled-mass microrheometer. Chapter 3, p. 55–74. In: RT Kelly, editors. *Advances in microfluidics*. London: IntechOpen (2012). p. 55–74.

20. https://www.globalgilson.com/gillmore-needle-apparatus (October 21, 2022).

21. Darvell BW. Adhesion strength testing – time to fail or a waste of time? *J Adhesion Sci Tech*. (2009) 23:935–44. doi: 10.1163/156856109X440966

22. Darvell BW. A polemic on behalf of a poor cousin. *J Dent Res*. (1989) 68(5):843. doi: 10.1177/00220345890680051901