COGNITIVE DISSONANCE IN GEOTECHNICAL ENGINEERING

Burt G. Look
AGTRE PTY Ltd

https://doi.org/10.56295/AGJ6031

ABSTRACT

Cognitive dissonance is the incompatibility between any two elements of knowledge or belief. We often filter information that conflicts with what we already believe, to avoid contradictory statements occurring. The term can also be loosely used when contradictory statements occur. There are many contradictory statements in Geotechnical Engineering – a few of these will be discussed.

There are many common geotechnical practices that have historically served the engineering community well, but continue to be used in common practice, even when later knowledge has shown inconsistencies. Can risk management substitute for data? But even with data, why would a group of Engineers presented with the same data so often have different conclusions? When did data become a point of view? A sampling of such dissonance in common geotechnical practice is presented through case studies.

1 INTRODUCTION

Common practices in geotechnical engineering seem to have many silent inconsistencies that as an industry we do not question. There seems to be disconnect between what we know and what we do in practice. Why? If we have information, but have not applied that knowledge, have we learnt and progressed? This suggests getting updated or correct information is not the issue – changing out behaviour may be the problem. This paper does not offer any new geotechnical knowledge, but examines commonly known and "old" knowledge that still is not widely used in practice.

Leon Festinger (1957) proposed a theory of Cognitive Dissonance (CD) that arises from conflicting thoughts, beliefs, or behaviours:

- People strive for internal psychological consistency to function mentally. We strive for cognitive consistency, so when inconsistency arises, we seek to reduce dissonance
- The conflict between knowledge and behaviour creates dissonance and we are motivated to reduce the cognitive dissonance to protect a core belief. This occurs by:
  - Adding new parts to the cognition causing the psychological dissonance (rationalization) or
  - Avoiding circumstances and contradictory information likely to increase the cognitive dissonance (confirmation bias).

A famous example is found in the term "sour grapes" which originated with one of Aesop's fables, about a fox who cannot reach grapes that he wants. After numerous attempts to get the grapes he "decides" that he does not want the grapes. They are probably sour or not ripe. The fox rationalizes his failure (**Figure 1**).

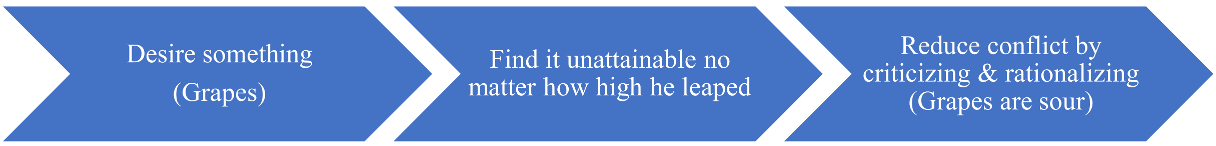


Figure 1: Sour grapes cognitive dissonance

Heuristics are patterns of thinking. There are many heuristic devices which helps using making quick decisions. These mental shortcuts of "in my experience" and rules of thumb, provide us with a simplified first pass on a way forward. However, we should not always settle for the simplified or convenient explanation. Heuristics may be right much of the time, but they are wrong often enough that they quite frequently lead us astray.

In civil engineering projects, an example of cognitive dissonance would be in airport concrete pavement construction. An engineer believes (is taught as part of their Civil Engineering training) that a lower concrete temperature during construction reduces the risk of concrete cracking. However, due to project constraints (e.g., tight schedules, weather conditions), a higher concrete temperature during placement is required at a site to complete that overnight pour over sig significant lengths. The conflict between their belief and the actual behaviour (pouring at higher temperature) creates cognitive dissonance. To reduce dissonance, the engineer might justify the decision by emphasizing other factors (e.g., strength gain occurs over time, avoiding construction delays) take precedence or rationalise control joints or curing methods are sufficient to mitigate potential cracking (**Figure 2**).

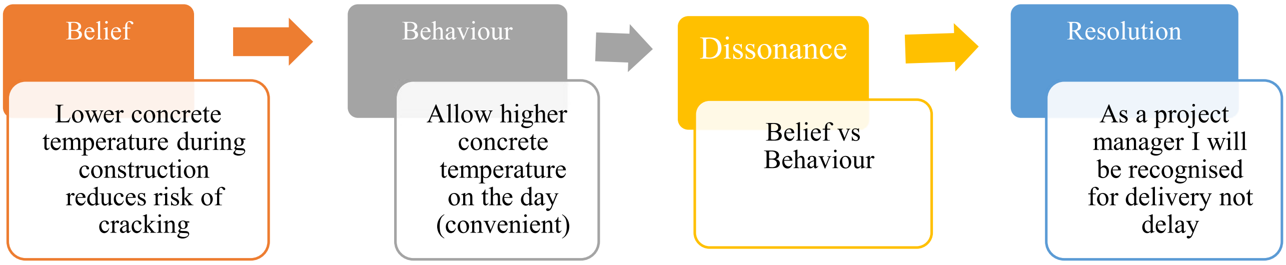


Figure 2: Concrete temperature during construction

The site engineer would have moved on to another project (or another job) when a decade later the concrete slab cracking occurs for those slabs poured at higher temperature during aircraft operation. An associated contractual litigation claim then results and many parties are affected - the contractor, sub-contractors, designer, sub consultant, geotechnical consultant, etc. The actual "cause" is not readily known when cracks are first observed and may take a few years of forensic engineering to reveal the cause of failure with experts deliberating over aircraft use, the pavement design, contractor's work, construction methodology, change of land use, ground conditions, services under apron climate change, etc. This is how legal cases occur and typically will not be ever reported.

This concrete temperature example also has a fallacy of the single cause, as a failure in a complex project is being attributed to only one factor while neglecting the other influences that contribute to the outcome. Lawyers involved in construction often pursue this oversimplification with a single cause. This ignores the interplay of various factors in construction projects. In this example, construction sequencing due to services may also play a part, or a prolonged dry period over a class "E" site. There is no new knowledge in this project example, only how knowledge is applied in practice, and how an early dissonance (**Figure 2**) may lead to a legal case.

Many common geotechnical practices also show cognitive dissonance. This paper covers some of these common practices used in industry which could be associated with:

1. Lack of awareness or knowledge
2. Accustomed practice i.e. "That's the way we do it" is mandated by their project elders. Some practices which historically served the community well continue to be used, even when current knowledge has shown inconsistencies.
3. Lag between state of practice and state of the art
4. Precedence of time or cost related factors
5. Not wanting to be different from what we know other practitioners are doing

If one uses a generation time as between 20 to 30 years, then if a known known has not been state of practice in 25 years, then this is being attributed to cognitive dissonance in geotechnical engineering. This paper examines knowledge prior to the year 2000, but which is still not widely used in current practice. This is an indicator of dissonance in practice, where knowledge is unable to supplant "that's the way we do it". These useful practices but with associated (**Figure 3**) misuse include:

1 Basic classification tests
   a. Applicable to transported soils but is also widely used as a direct indicator of behaviour in Australia where residual soils dominate
   b. Grading tests not representative of on-site conditions
   c. Plasticity Index (PI) as an initial indicator of likely expansive behaviour but is not representative of the whole sample

2 The standard compaction test is THE most basic test in earthworks control. Unfortunately, the optimum moisture content (OMC) is often used as a target and "best" value for compaction. The relevance to the EMC (Equilibrium moisture content) to construction practices is shown from case studies. The significant issues associated with the related modified compaction test and the associated laboratory California Bearing Ratio (CBR) as a design basis are briefly discussed.

3 The Standard Penetration Test (SPT) is arguably THE most common in-situ test and is used as "factual" in bore logs. The SPT is useful in obtaining samples and as a numerical guide to the soil strength. Counting "variations"

occurring at all sites are shown. Some (mis) interpretation occurs at low to high end values. An N value of 1 and 50 are typically automatically assessed as very soft in bore logs and being near pile driving refusal, respectively. Neither of these are correct as will be shown. Additionally, the Field N- value (N_{Field}) continues to be used as design value, yet a design value needs to be appropriately corrected for energy and overburden [$N_o(60)$]. Case studies with digital measurements instead of eye measurements, show the significant counting and energy corrections which vary with each drilling rig.

4 The Design Characteristic Value is used to provide a strength assessment of the ground conditions, which envelopes the site risk. A moderately conservative value is often based on the interpretative eyeball line on a graph from the data obtained or Tabled, and/or based on experience. Yet this interpretation of moderately conservative varies significantly in industry for engineers presented with the same site data. (Look and Campbell, 2013). Applying statistics is more defensible, yet the use of statistical models may lead to additional confusion by adopting the common normal distribution. Case studies are shown where the soil, properties are not normally distributed. A characteristic design value requires a best fit Probability density function (PDF) to have a relevant and defensible input to both probabilistic and deterministic design.

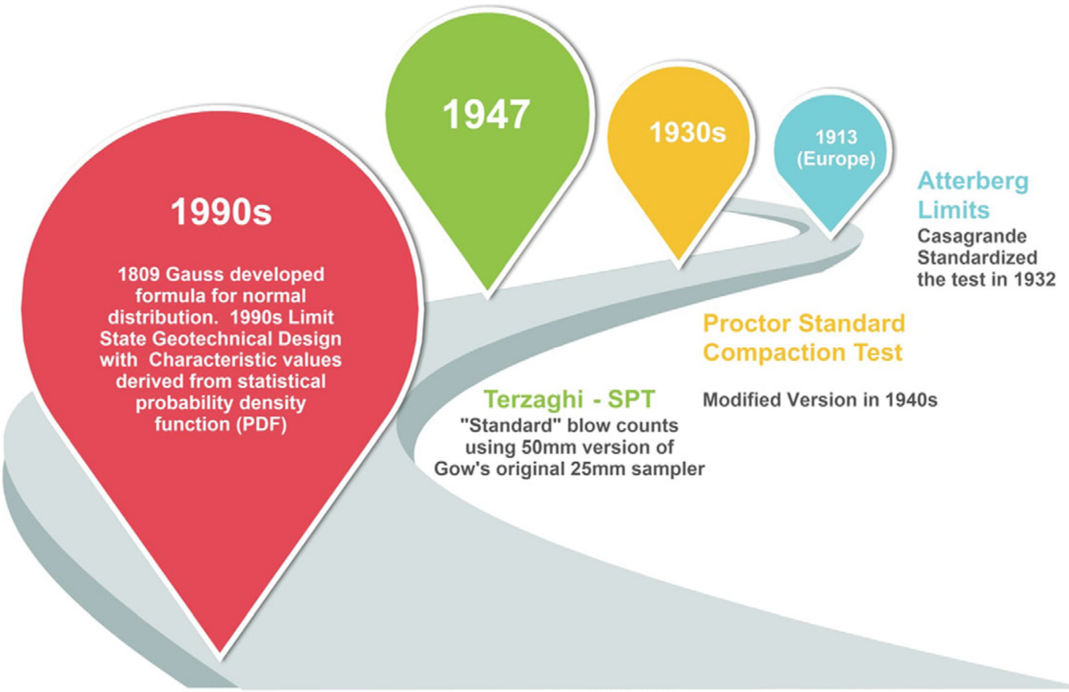


Figure 3: Useful common practices and its historical development

These geotechnical basics are highlighted herein as they are regularly sighted in reports. This paper is about application of existing knowledge rather than any new knowledge. The paper does not suggest all practicing engineers have these “errors”, but is used here to show that it has been observed to be regularly occurring.

The acknowledgments at the end of this paper provide background to the study of critical thinking and cognitive dissonance in management, decision making and social contexts.

2 HEURISTIC DEVICES

Ockham's razor is the problem-solving principle that recommends searching for explanations constructed with the smallest possible set of elements and can be paraphrased as “The simplest explanation is the best one”. This is a key principle of parsimony as a heuristic device. Heuristics are required so we can produce a speedy decision. These mental shortcuts allow us to make a judgement quickly with rules of thumb in complex situations. Heuristics rely on familiarity and availability of information and are derived from previous experience with similar problems. However, heuristics have both advantages and disadvantages as such strategies while time efficient may also lead to cognitive biases. Becoming aware of this might help you make better and more accurate decisions instead.

Parsimony is a guiding principle which suggests that we should prefer simpler explanations and solutions over more complex ones, all other things being equal. Parsimony helps guide our reasoning and decision-making in various scenarios. In a more general context, parsimony sometimes also refers to reluctance to spend resources (especially time and money).

Historically, our early geotechnical pioneers applied the principles of parsimony in many approaches we use today:

- By classifying a soil, we have an initial indicator of its likely properties
- By improving density, we have greater confidence in likely strength and modulus gains
- By applying a Standard Penetration Test (SPT) value we can assess the strength of soil
- By carrying out a CBR test, we have an indication of a subgrade strength

There are many useful correlations which can be applied to these indices. These heuristics have served the geotechnical industry well, and are essential in any preliminary geotechnical assessment. However, usefulness is different from truth. Indices in preliminary assessment sets the general direction, but should not be the only basis for a detailed geotechnical design. One must recognise that many common indices are simplified "knowledge". These guide early solutions, but project specific data and considerations trumps any index. The heuristics and associated cognitive dissonance are examined further with:

- Geotechnical investigation testing
- Assessing variability
- Analysis and design
- Construction quality control

Cognitive biases can be categorised in a number of ways and are widely discussed in the field of social engineering. Some are only subtly different to each other, while others may represent opposite ends of the same spectrum. A short list of common biases are shown at Table 1 with geotechnical examples that are explored in more detail in subsequent sections. There are many other biases not discussed in this paper.

Table 1: Selection of bias affecting our thinking

Common Bias	Overview	Issue	Geotechnical examples
Anchoring	Relying on the first piece of given information	Regardless of the accuracy of the initial information, it is used as a reference point, or anchor, to make subsequent judgments. Related to confirmation bias	Text book correlations overriding actual site test data or "averaging" between the two results. Correlation may not be applicable
Availability heuristic	Your judgement is influenced by readily available information and what springs easily to mind	Media coverage, sensational / recency of event or experience, and the unavailability of information influence decisions. Occurs when rushed decision required. Give yourself time and seek out other information. Statistics provide a clearer view of the big picture.	Using a one point CBR test value in design as tested at MDD and OMC. The CBR (max) is not directly related to OMC especially in expansive clay material. MC changes with time and so will the CBR
Circular argument AKA begging the question	Assumes the conclusion is the premise. A form of confirmation bias	No new evidence is provided to support the conclusion. Using a definition to define itself, and using the same statement as evidence to support a conclusion.	A driven pile will only refuse if SPT- N > 50. SPT > 50 means a pile will meet driving refusal. The latter is not correct but often applied by inexperienced engineers
Curse of Knowledge	Once you understand something you presume it to be obvious to everyone	This often causes a barrier to effective knowledge sharing as we are uncertain about what the other party already knows. Hindsight bias occurs after an event when as an "expert" we assign a higher probability of an outcome.	Geotechnical specialist neglects to adequately convey the significance of the many ground issues to the design engineer or in the contract documents (occurs with many specialists)

Common Bias	Overview	Issue	Geotechnical examples
Confirmation bias AKA cherry picking	Favouring information that confirms our existing beliefs	Dismissing information that does not support our belief (Ostrich effect). Also related to sunk cost fallacy. Avoid by being aware of your beliefs. We are typically more aware of our assumptions than of our biases. Test assumptions by searching for disconfirming evidence of your assumptions, to form factually supported arguments. Beware of repetition – a form of brainwashing - you begin to think that something is true simply because you have heard it so many times.	Text book correlations overriding actual site test data. Such correlations may have derived for a different soil type and Geology. Such correlations should not override site specific data and tests. Related to "Group think". If everybody else is doing something, then it must be right
Group think	The social dynamics of a group situation overrides the best outcome	The desire for harmony or conformity may results in an irrational decision making and to agree at all costs. A consensus decision occurs without critical evaluation with failure to listen to anyone with dissenting opinions.	SPT energy is required for design. If other companies are not measuring energy and it is not required in Australian Standards then is OK to use SPT (field) value as a design value
Information Bias.	Tendency to acquire more and more information. Believing that more information acquired to decide, the better, even if that extra information is irrelevant for the decision	The perception that more information leads to better decisions; a fear that mistakes will be blamed on the lack of information; collecting information just in case it is useful. Seeking information when it does not affect action. Collecting extra information has the potential to lead to data noise; cognitive overload unless there is an adequate supply of time or expertise to filter the information. Related to confirmation bias	There is more likely too little than too much ground data, therefore less likely for this bias to apply to geotechnics for data quantity. Or maybe this author is biased as a geotechnical engineer? Measurement bias may occur due to inappropriate tests, biased observations or instruments not calibrated
Survivorship bias AKA cherry picking	Assessing successful outcomes and disregard failures	Survivorship error arises because we look at the data we have but ignore the selection process that led us to have that data. Confusing correlation with causation. What is the data that is not present?	Boulder sizes on site are not sampled. Affects gradation. Atterberg, compaction, CBR tests removes material for lab testing.
Sunk cost fallacy	You cling to things that already have an invested cost and time.	Stick to the plan as you have invested heavily (time, money, effort) even when giving up is clearly a better idea. Unrecoverable past costs in present decision making. Stopping means admitting that whatever resources we invested up until then had been wasted. We avoid the embarrassment of admitting the initial decision was wrong	Construction quality control based on density testing – the most precise and best test available in the 1960s. More relevant and accurate modern day tests are unable to supplant the widely used existing quality specifications

The easier biases to recognise are first discussed. Many of the examples do have more than one bias associated. This paper does not present an easy solution as bias recognition is the focus. Additionally overthinking can be a hindrance and the heuristic used for quick decisions is still useful and should not be completely abandoned because it is not perfect.

The availability, group think and survivorship bias are the 3 most common and will be discussed with examples.

3 SURVIVORSHIP BIAS

Survivorship bias is the logical error of concentrating on the people or things that "survived" some process and inadvertently overlooking those that did not because of their lack of visibility (Dobelli, 2013). This can lead to false conclusions in several different ways. Survivorship bias leads to incorrect conclusions due to only studying a subset of the population. A common example of survivorship bias occurs when a smoker points to those in the room who have been smoking for 30 years and healthy. Not having the dead smokers in the room is a sampling bias on a health assessment. Survivorship bias in this case ignores what was not sampled. Examples of survivorship bias in geotechnical testing are:

- Gradings. Large cobbles / rock sizes > 300 mm are not picked up and 100% passing 300mm is often (incorrectly) seen in grading curves
- Residual soils. Significant amount of samples are not tested in Atterberg limit tests which require particles < 425 micron size for both the liquid and plastic limit. Inferring high plasticity if most of the sample is discarded is survivorship bias
- We place boreholes mainly in areas the drill rig can access, or we have permission to access, or where we do not have to create drilling platforms. Test holes may not be in the worst areas if soft / loose ground occurs adjacent to waterways.
- Strength on rock core samples are often biased towards the core lengths with intact samples, which may produce stronger rock strength than available over the whole profile. There is a need to test every 1m (or a pre-defined length) consistently.

A few of these are discussed in more detail.

3.1. CLASSIFICATION TESTS

Classification tests are based on gradations and Atterberg limits for coarse grained and fine-grained materials, respectively. Look (2019) discussed the application of these basic tests. Grading curves should always be used with caution in residual soils, and weathered rock. Curves with 150mm or 300 mm maximum particle size should always be used with visual evidence as testers typically will not include these larger sizes in their sample bag. A common occurrence of the survivorship bias.

Soil testers do not place large sizes into their sample bag as it exceeds the safe transporting lifting weight (**Figure 4**). Placing even 1 boulder size (> 200mm) in a sample bag would also lead to ridicule from peers, even when accompanied by another bag of the representative smaller size material, given 1 sample bag (20kg) is considered sufficient for a grading curve. The soil tester is asked to "obtain a sample to be tested". As large (over size) samples are not tested, the tester has satisfied their brief, although the grading is now incorrect. Contractual disputes may arise if a contractor later claims for removal of oversize and the "factual" grading says otherwise.

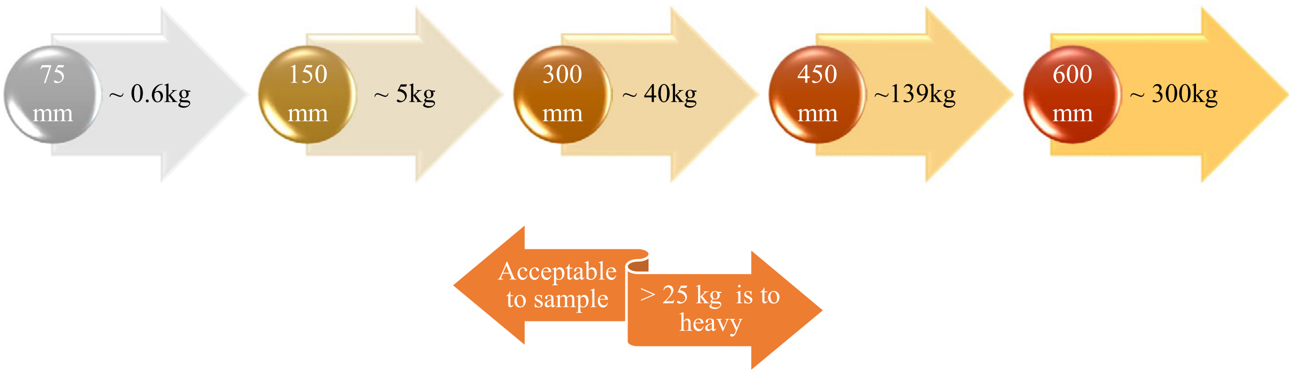


Figure 4: Sample sizes for testing and corresponding weights

3.1.1. Gradings

Since the 100% passing of the material tested does not represent the maximum size on site, the % sizes in the grading curve would be different with and without the site size (**Figure 5**). The factual numbers from the NATA test certificate are based on a survivorship bias.

For quality control testing the dry density ratio test (Field Density / Lab maximum dry density (MDD)) is used. The 20kg sample size obtained for the lab MDD should ideally extend to the layer compaction depth (**Figure 6**a from TfNSW, 2015) which was concerned with the moisture content gradient with depth.

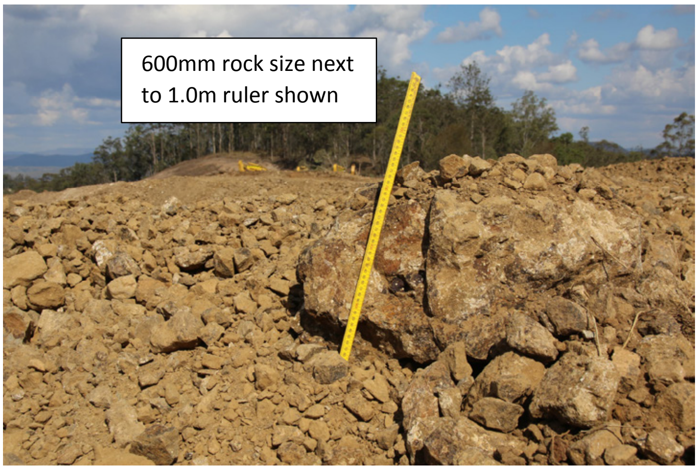
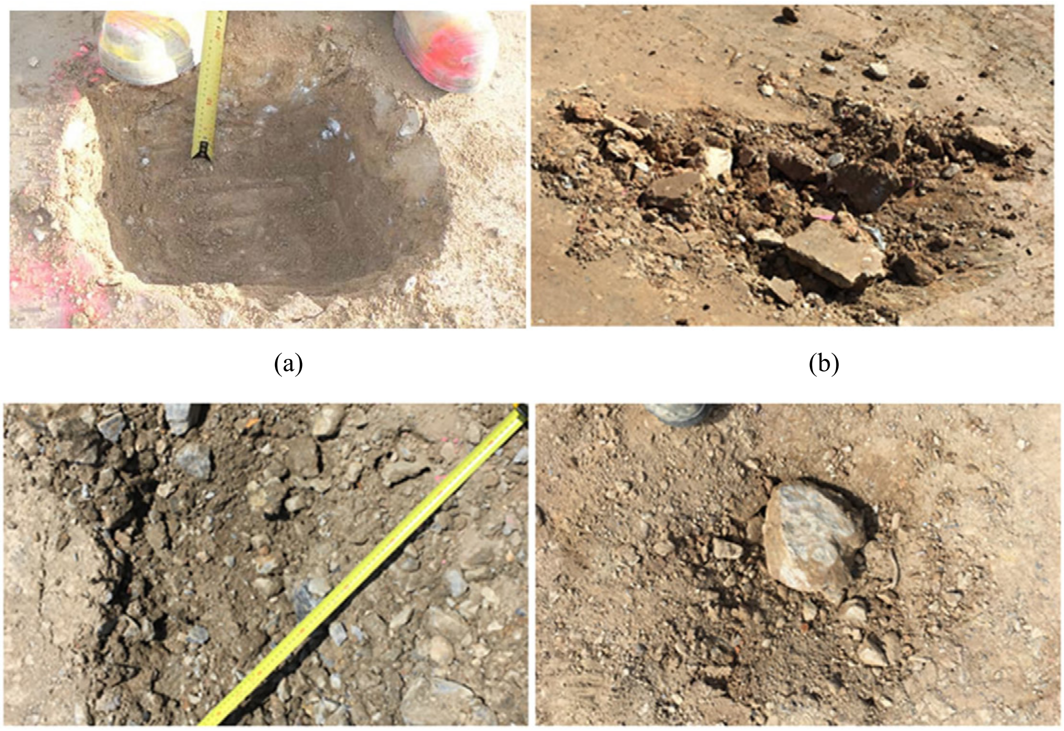


Figure 5: At this site 4 out of 5 pre compaction gradings showed 100% passing 150mm sieve size

The lazy approach of scraping the top surface (**Figure 6**b) to fill the sample bag also occurs with an associated error in particle sizing. For compacted weathered rock material, the crushed material from compaction is typically at the surface (**Figure 6**c), while the larger sizes is pushed down (**Figure 6**d). The sample arriving at the laboratory for grading or density testing may not have the larger particles if the full depth of the 200mm compacted layer had been sampled. This is the survivorship bias.



(a) (b)

(c) (d)

Figure 6: (a) Ideal 150mm deep sampling hole, (b) Shallow sampling, (c) Crushed material at the top, (d) Discarded material (> 100mm size) typically below surface sampled

3.1.2. Atterberg Limits

The plasticity index (PI) and liquid limit are used to classify the silt and clay behaviour. The Atterberg tests are carried out on the percentage passing the 425-micron sieve as sizes above impeded flow (for the liquid limit test) and cause the thread to crumble (for the plastic limit test). This was for testing of transported soils by Atterberg and Casagrande.

Look (2016) shows the Cumulative Distribution Function (CDF) for the % passing the 425-micron sieve used to carry out the PI tests for 285 samples examined for a large transportation project in Queensland (**Figure 7**). In carrying out the PI tests, typically only 60.4% of the sample was used in the test of that residual soil profile. Thus, the PI as an index of expansive clay behaviour can be misleading in residual soils due to the high percentage discarded in carrying out the test. This is another example of survivorship bias as sample is discarded in carrying out the test when residual soils are tested.

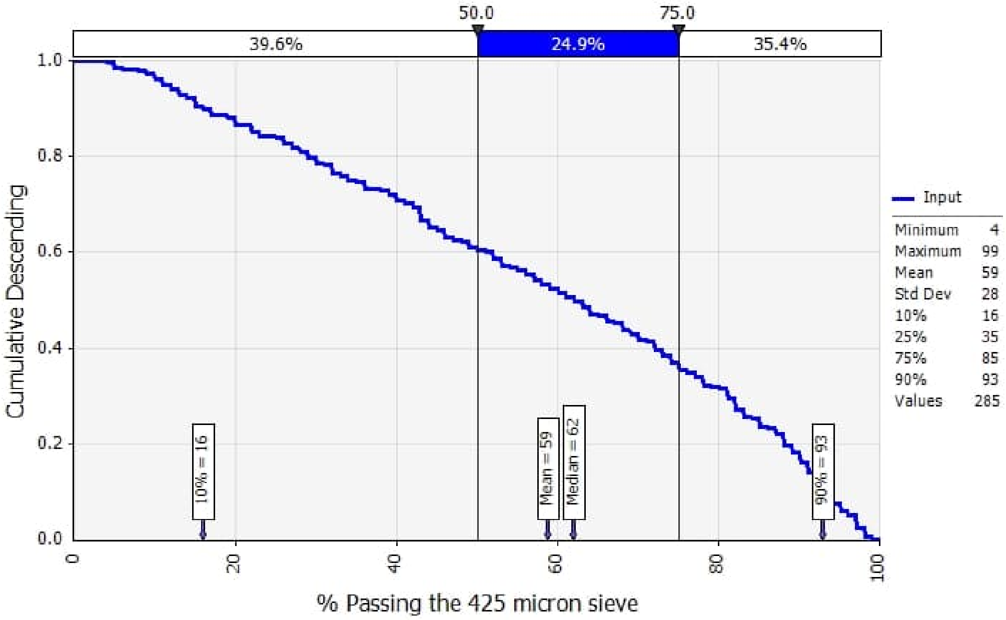


Figure 7: Cumulative distribution of fraction used in PI tests

The misclassifications associated with using the PI only in residual soils was overcome by using the Weighted Plasticity index (WPI). The WPI is the product of the plasticity index and the number percentage passing the 425-micron sieve (Look, 1995 and 2016). Using the PI as a criteria of expansive (**Table 2**) and non-expansive clay results in 20.2% of samples in residual materials having a PI less than 12% (non-expansive) across sites in Queensland (Look, 2016). If a WPI of 1200 is adopted to account for material actually used in the testing, then 51.4% of residual soils are now non expansive.

Table 2: Plasticity Index and survivorship bias – throwing away material as part of test

Plasticity (PI) Criteria		Weighted Plasticity Index (WPI)		Potential for Volume Change
Typical PI range	Likelihood	WPI	Likelihood	
< 12%	20.2%	< 1200	51.4%	Low
12% ≤ PI < 22%	31.8%	1200 ≤ WPI < 2200	17.9%	Medium
22% ≤ PI < 32%	21.6%	2200 ≤ WPI < 3200	13.9%	High
PI ≥ 32%	26.4%	≥ 3200	16.8%	Very High

This fundamental test of using PI only as an indicator of soil volume change potential has an inbuilt survivorship bias as part of its procedure, and a “correction” is required.

3.2. COMPACTION TESTS

For this residual soil data set in Look (2016), the percentage passing the 19mm sieve was also analysed. Oversize is any figure retained on the 19mm sieve according to the Standard compaction test procedure. Above 20% “oversize” the test is not applicable associated with the compaction tests. For this data 18.5% of the compaction test should be considered invalid and approximately 50% requires a correction factor to be applied to the maximum density reported (**Figure 8**). A correction factor was not sighted in any of the MDD results.

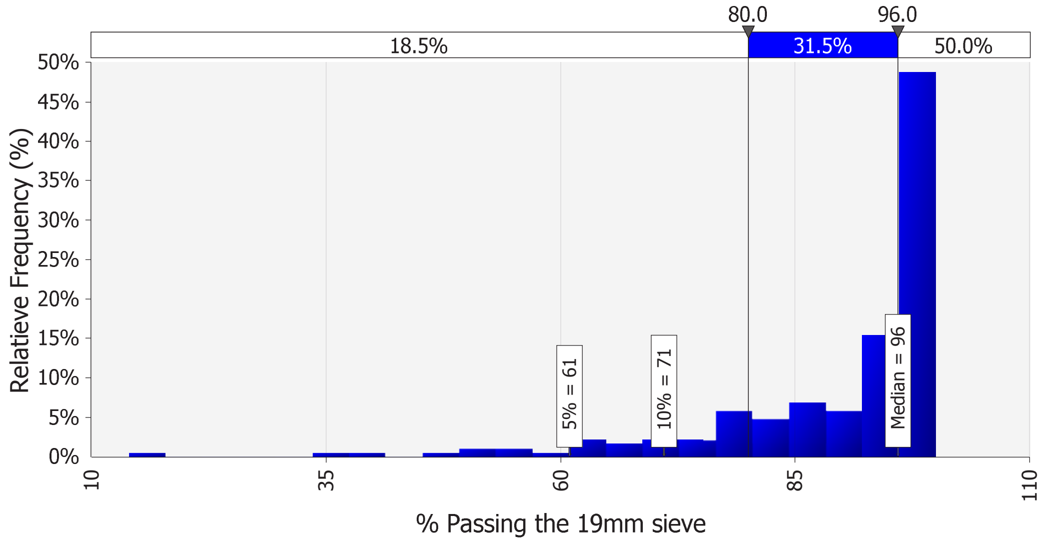


Figure 8: Distribution of Percentage passing the 19mm sieve when residual soils were used

The gradings were caried out independently of the compaction test results. The site supervisors and contract administrators have ticked all the required boxes for quality control testing with no invalid tests reported at this site. This is the standard of practice for a successful project.

3.3. CBR TESTS

Similar to compaction tests, a maximum size limit occurs for CBR testing. In residual soils and weathered rock, a high percentage of material is removed from the laboratory sample tested for a CBR value.

Engineers would typically suggest the CBR of a highly weathered rock should be higher than the CBR for a residual soil or extremely weathered rock. However, the laboratory CBR test would not necessarily show such results. Extremely and weathered rock was shown to provide a “peak” soaked CBR value of 13% (Lacey et al., 2016). The field testing using a Light falling weight deflectometer (LFWD) equipment to characterise the *insitu* modulus parameter showed a clear increase in value for this project (**Figure 9**). An unsoaked test would have provided a more direct comparison, but that data was not available.

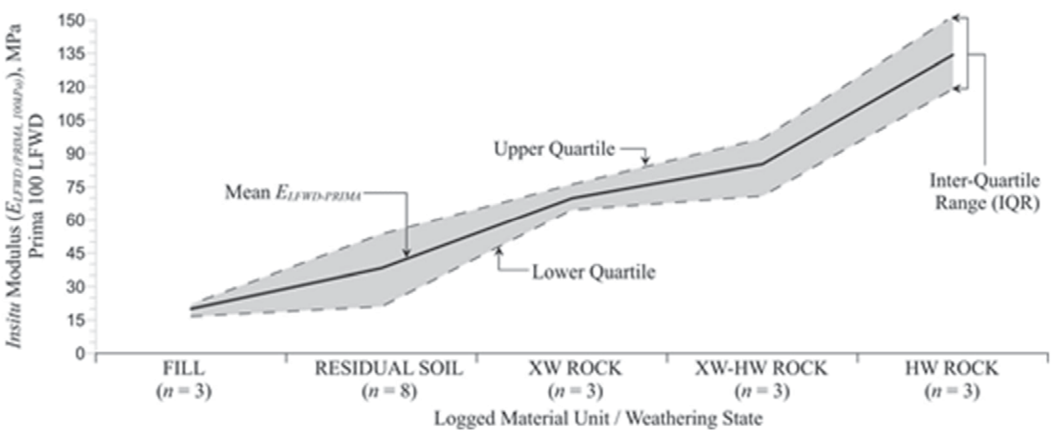


Figure 9: *Insitu* modulus measured by Prima 100kPa test stress ($E_{LFWD-PRIMA}$) categorised by weathering state

The laboratory determined, a soaked CBR result plateauing across the XW to HW rock of 13%. This concept is illustrated in **Figure 10** to show the effect of the relict rock structure and oversize particles present test. It should be noted that the CBR test is not applicable for samples with greater than 20% over-size materials (as per AS1289), yet continues to be used for such materials, due to the absence of a more reliable or convenient test. In such materials the survivorship bias skews the test results.

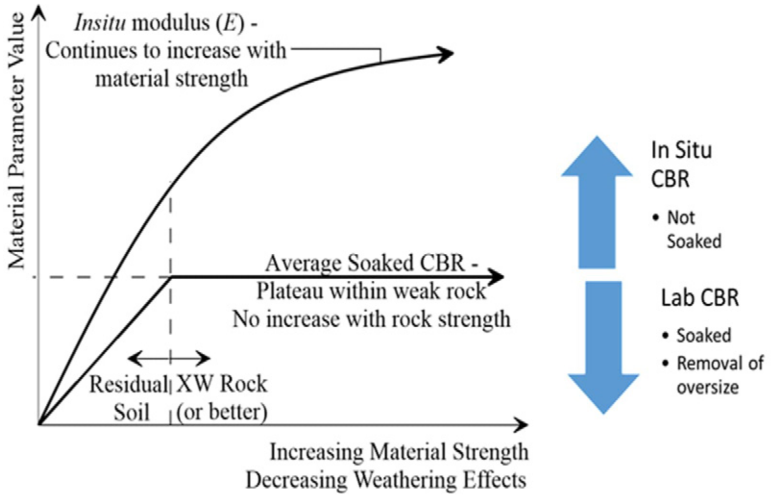


Figure 10: Conceptual difference between *insitu* and soaked CBR values over residual soil / weak rock profile

4 AVAILABILITY BIAS

The availability bias relies on immediate examples that come to one's mind when deciding and is inherently biased toward recently acquired information. Making decisions quickly make sense, and we may be biased towards what is available. Dobelli (2013) states we prefer wrong information to no information and relying on information that is readily available or easily recalled may not be best for decision making. Our judgment in frequency and probability of events is affected by the availability bias. Examples of availability in geotechnical testing and analysis are:

- We place boreholes mainly in areas the drill rig can access, or we have permission to access. Availability of access means test holes may not be in the worst areas of soft / loose ground adjacent to waterways /wet areas.
- Limitations in direct shear tests, results in the omission of particle sizes larger than 1/6th of the height of the test apparatus. Where such oversize exists in situ even after compaction, then inadequate size laboratory shear box testing results in over conservative design values. This is availability of test equipment.
- CBR testing at the OMC can be misleading as a design value.
- Correlations used in analysis
- Factor of safety versus probability of failure
- CBR and compaction tests. The % above 38mm size are discarded (strong sizes) resulting in natural soils being under-predicted → In the early days (1930s) of the confident 150mm compacted layer with the equipment of the day. Hence a 75mm max size (½ compacted thickness) is the maximum grading size seen in pavement granular crushed rock specifications, but 38mm maximum for grading curves with high specification materials given the CBR plunger and compaction hammer was 50mm diameter, and maximum sample size needed to be less
- Equipment availability

A contractor often has large excavators on site. Excavation of extremely (XW) and highly (HW) weathered rock often occurs with such equipment when top of rock is specified. The availability of this equipment is different from the smaller backhoe which may have been used in the site investigation. The removed "soil" is then replaced with imported compacted material which has less the 20% of the strength of the replaced material (**Figure 11**). At the site shown in the photo the SW Andesite rock is evident. An additional excavation cost of $250,000 applied at this site in 2013, despite such remove and replace not being required.

This is an example of both availability and confirmation bias. The latter occurs because the weathered rock material is soil like and the observation of "soil" being removed by the excavator. By definition XW and HW can have greater than 50% soil material within the rock structure.

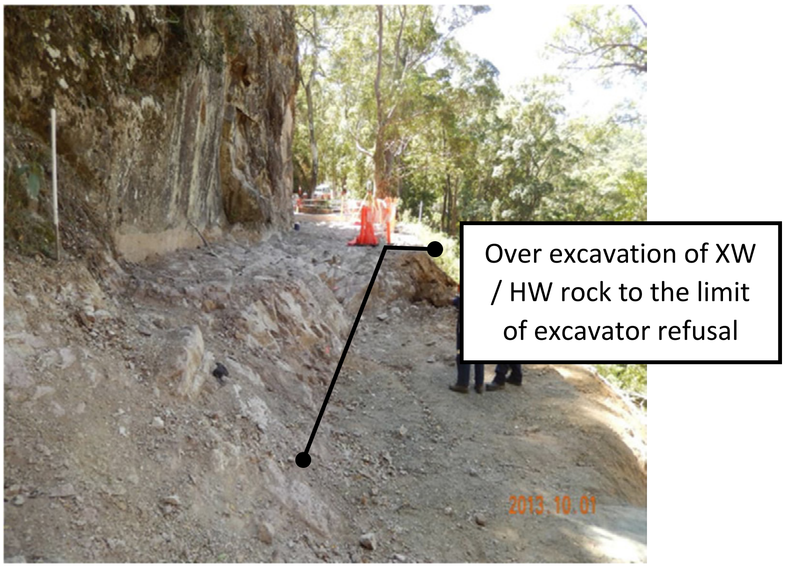


Figure 11: Rock level using the available 20 t excavator

3.4. CORRELATONS USED IN ANALYSIS

There are many correlations used in practice which are useful in the planning phase or preliminary design, but should not be used in the detailed design phase. As an example, the coefficient of consolidation (C_v) with the liquid limit (LL) is often sighted in many text books. Look (2023c) questioned this graph reliability as the 3 data points from another historical text book coincided with this graph as shown in **Figure 12**. A general purpose graph that may have been derived from 3 data points in the 1970s should not represent present practice.

Data from 10 Queensland sites with soft clays (126 data points total) are plotted for various load increments in **Figure 13**. The neat and simplified relationships from **Figure 12** are no longer apparent and with the best relationship with the high load increment at 400 – 800 kPa (R^2 = 0.42), which is unlikely to occur in practice. This is using Casagrande's t_{50} approach. This data highlights the NAVFAC charts to derive c_v from LL and widely seen in many text books, should not be used in detailed design due to variability of data, and poor correlations. Due to anchoring and availability biases, reports have been sighted which override project specific test data in favour of this correlation chart.

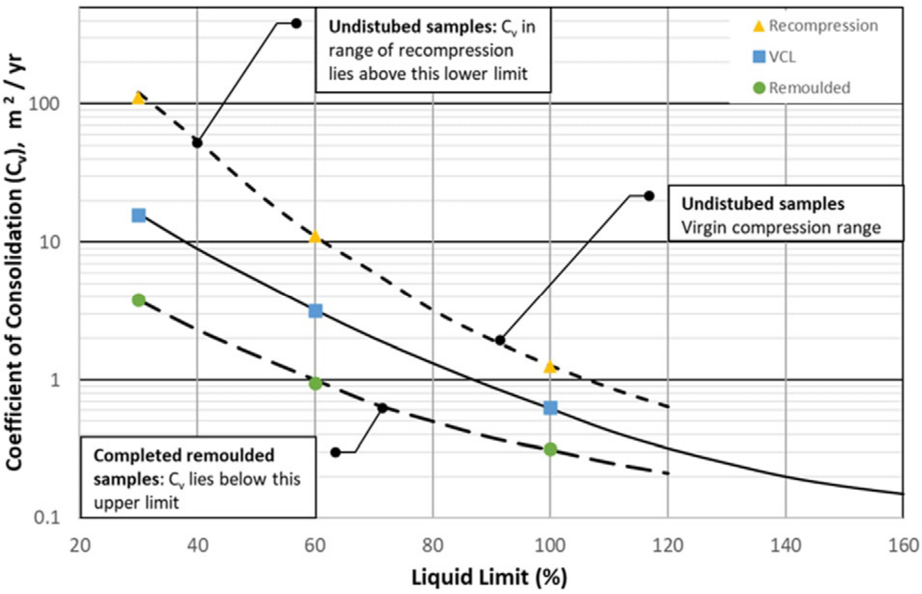


Figure 12: Relationship between cv and LL replotted with different units (NAVFAC, 1986) showing the 3 data points Tabled in Lambe and Whitman (1979)

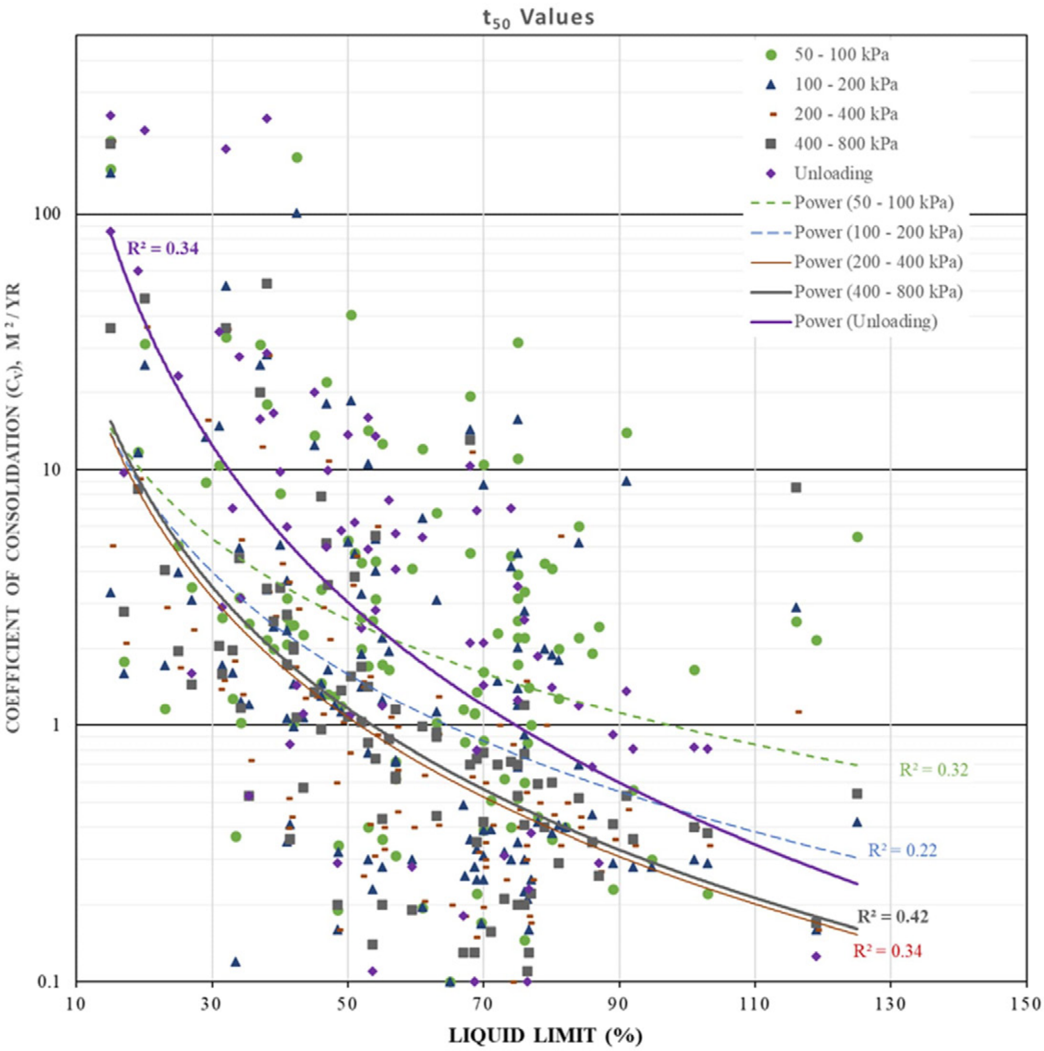


Figure 13: Queensland consolidation data (126 consolidation tests) using Casagrande's t_{50} approach

4.1. SHEAR BOX TESTS

It is widely accepted that oversized particles have a significant effect on the strength of a soil mixture and their effect cannot be ignored to obtain accurate shear strength values. Available shear box sizes are typically between 20 mm - 40 mm. AS 1289.6.2.2 (1998) for direct shear test using a shear box requires a particle sizes larger than 1/6th of the height of the test box. Therefore, the maximum particle size permitted in conventional shear testing is approximately 6.5 mm, However, large scale shear boxes with depths varying above 150 mm and yielding an allowable maximum particle sizes upward of 25 mm are available and more representative where residual soils and compacted weathered rock (which breaks down to soil but with "oversizes") are used as fill material.

The requirement for proper testing procedures and the effect of sample sizes are discussed in Nakao and Fityus (2008). They found the larger particle sizes (9.6 – 19mm) used in the larger 300mm X 300mm X 190mm shear box increased the friction angle by 4˚as compared to a 4.75mm maximum size in a smaller shear box of 60mm X 60mm X 50mm, even though the coarser gravel made up only 9% of the sample tested. All test results below assume suitable testing rates and preparation to the required testing standards have been carried out.

4.1.1. Sunshine Coast site

Lime treated material was used in a site in the Sunshine coast to minimise importing of fill material. This material was derived from weathered mafic volcanics (described in Chew et al, 2012; Boyer et al., 2013) with significant large sizes on site when compacted. The test results show using the larger shear box increases the angle of friction by 4° for this material (**Table 3**).

Table 3: Effect of shear box size on test results

Shear box tests (mm)	Treatment	Cohesion (kPa)	Friction angle (°)
60 X 25	4% hydrated lime cured for 7 days	65 105	34 31
140 X 30		60	38

4.1.2. Toowoomba site

Typically, 60mm size shear box is often used commercially. Testing at that smaller shear box size during design showed results of friction angle of 30˚ to 36˚ The larger shear box therefore seems warranted where large sizes would be used *insitu* despite the higher cost and time associated for testing. However, obtaining 50 kg size samples for testing is required, which is 100 times greater than for the small size shear test. This example shows that the availability or cost attraction of the smaller size shear box affects the test results as different particle sizes then applies.

At the Toowoomba Range Second Crossing Project (TRSC), the observed field testing results returned higher strength properties than the laboratory tests (refer also to section 6.5). Laboratory investigation uses different sized direct shear boxes and with three different maximum particle sizes of; 75 mm, 53 mm and 19 mm, were carried out to observe the effect which these particles typically excluded from tests have on the friction angle. The results of the particle size using the 300 x 300 x 160 mm is shown in **Figure 14**. The friction angle changes dramatically as larger particle sizes are allowed in the test. Note the 53mm and 75mm sizes exceed the maximum particle size allowed for the size shear box according to the testing standards, and this was for research purposes only to assess the potential effect of sample size. Additionally, this is the peak friction angle and the ϕ_{cv} friction angle may apply.

At this site, the maximum materials size pre and post compaction during construction trials were as follows:

- 75mm size: 15% and 1% greater pre and post compaction, respectively (noting survivorship bias).
- 53mm size: 33% and 7% of sizes greater pre and post compaction respectively
- 19mm size: 67% and 32% of sizes greater pre and post compaction, respectively

The 75mm size test would be unwarranted due to compaction breakdown.

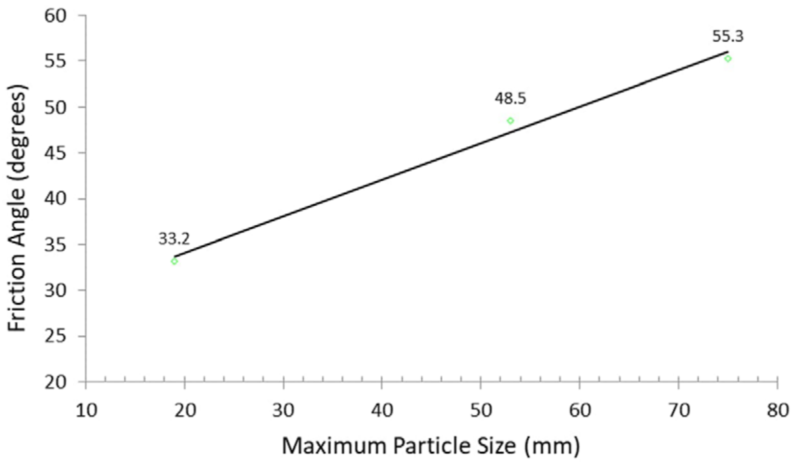


Figure 14: Friction Angle effect for maximum particle size using a 300 x 300 x 160 mm shear box

4.2. OPTIMUM MOISTURE CONTENT

In using density control as a quality control parameter, many engineers associate the "Optimum" Moisture Content (OMC) as a target which is associated with the Maximum dry density (MDD). This is the laboratory compaction model and useful for construction purposes.

Look (1995, 2006) shows the long-term equilibrium condition (EMC) is more important than a short-term OMC construction compaction target. This is similar to the equilibrium soil suction based on the Thornwaithe Moisture Index

(TMI) climatic zone in AS 2870 (2011). The EMC in terms of the moisture ratio varies with rainfall (**Figure 15**). Any CBR value tested at OMC is not specifically considering the long term moisture condition which affects the design CBR value. The OMC is also at a specific compactive effort in the lab. The compactive effort in the lab varies considerably based on the type and size of field equipment. Thus, the moisture content required to achieve a target MDD will vary from the lab value. Look (2016) highlighted this in a case study with 5,640 construction quality control test results with 45.7% of test values within 90% to 110% OMC, but acceptable compaction densities were being achieved outside of that moisture content range.

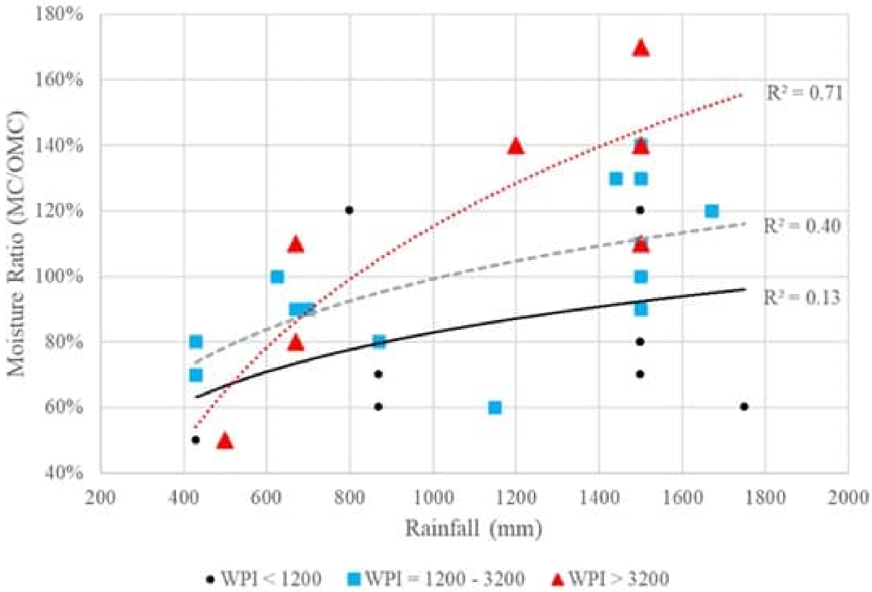


Figure 15: Equilibrium Moisture Content for varying WPI

For highly expansive clays with high weighted plasticity index (WPI), the movement rather than the strength governs the design. At low WPI, the compaction moisture content is poorly correlated to rainfall and therefore targeting the OMC seems reasonable. This would also apply to imported materials such as crushed rock.

For high WPI materials, the EMC is more important than the OMC, and therefore, while constructing at or close to the OMC should be targeted but biased towards the EMC. The design moisture content is the EMC and this influences the CBR value. Yet in practice geotechnical reports show the CBR value at OMC. This is an example of the availability bias. With no design data at the EMC the CBR at the OMC is then applied. Note the OMC would be acceptable for crushed rock materials and low WPI natural materials.

For the one- third of natural materials where the EMC ≠ OMC a CBR not at the OMC then applies. This is also the anchoring bias where the "correctness" of the CBR at the OMC is being incorrectly applied for high WPI materials.

4.3. CALIFORNIA BEARING RATIO (CBR) TEST

High strength often occurs at dry of OMC compaction. Soil suction has a significant effect on shear strength. As saturation increases, the material's strength decreases significantly, but with large strains to failure as shown by Seed and Chan (1959) and discussed in Leroueil and Hight (2013). The peak CBR does not necessarily occur at the MDD (**Figure 16**). This shows the peak strength is NOT associated with the MDD and OMC, and this has historically been known for over 60 years. The anchoring to the OMC has led to the CBR also being (incorrectly) tested at that moisture content and is discussed in Look (2023, 2024b) on constraints to accurate testing in earthworks practices.

The peak CBR is often not coincident with the OMC / MDD as shown in **Figure 17** for a CH Cooroy clay when 5-point CBR tests are caried out Typically, a high plasticity clay would have its high strength wet of optimum and below MDD. A higher Density ratio does not necessarily produce a higher strength or modulus – although this is an implicit assumption (Look, 2023a). In this material a design using a CBR of 2% would require a subgrade improvement as compared to a CBR of 3%.

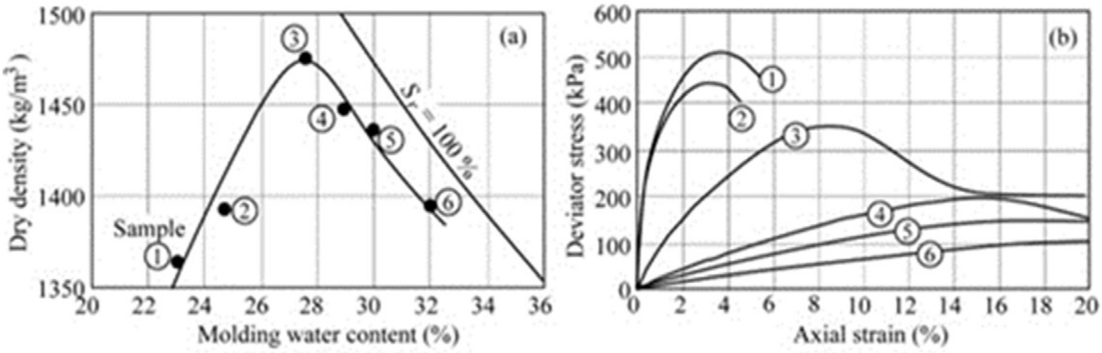
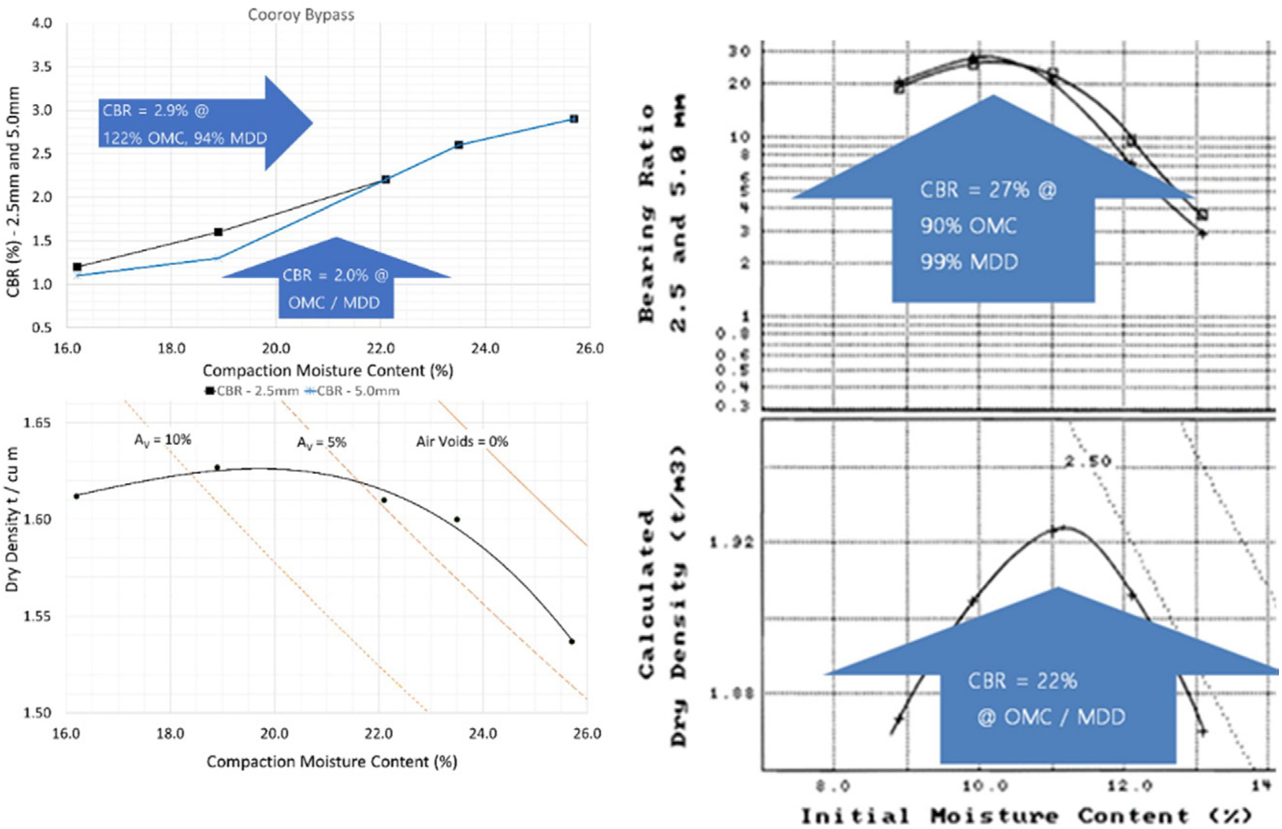


Figure 16: (a) Influence of moulding water content on the dry density and (b) stress-strain relationships for compacted samples (Seed and Chan, 1959; here from Leroueil and Hight, 2013).

A granular material would have its peak strength dry of optimum (below MDD) as shown in **Figure 17** for the Clayey Sand (Look, 2023a). There is a 5% CBR difference between the peak CBR test value and a CBR at MDD / OMC. However, for all practical purpose, both values in this example are representative of a suitable subgrade and would have little effect to the design and construction process, with an upper bound CBR subgrade value of CBR 15 often applied.

Moisture content governs the CBR value as shown in dendrogram analysis Look (2023a) with density poorly correlated to the CBR. Yet this assumption governs design and quality of construction.

This is another example of both availability and anchoring bias where moisture content on site is changing each day and season, so becomes difficult to measure. We therefore anchor to the density measurement during construction as it is a very precise measurement and is less variable. This false correlation of long term strength and modulus being defined by density is discussed further in section 5.2.



(a) CH clay **(b) SC Sand**

Figure 17: Maximum CBR is not coincident with MDD/OMC for a Clay and Clayey Sand

4.4. SAMPLE SIZE

Sample size often affects the test results as was illustrated with shear box test results (Section 4.1). A few other common examples are used to illustrate this availability bias in common practice.

4.4.1. Pocket penetrometer

The pocket penetrometer (PP) is a useful and simple test commonly used as a measure of the undrained shear strength. Its simplicity, low cost and availability often tempt many field engineers into its overuse. Many engineers prefer the less subjective PP numerical value in classification rather the tactile assessment of finger pressure on a larger sample for the strength classification of firm, stiff, etc. Due to the small test area the PP will likely overestimate the strength in a fissured clay.

Some tests also use the PP on an SPT sample. The split spoon sample is highly disturbed and the PP should not be carried out on undisturbed samples. The argument heard is that any test result is better than no test result. (Some of this may be forced by a civil or structural engineer insisting on this test occurring). During an SPT the soft to firm soils are compressed to a stiff state and the very stiff to hard are disturbed to a stiff clay (Look, 2004). Thus soft, firm, stiff, very stiff, hard all seemed to measure "stiff" with the PP on an SPT sample. The PP should therefore not be used on SPT samples. The availability bias to easily obtain such a test result does not mean it is useful, with often meaningless results and leads to a wrong data for analysis. The PP should therefore be used only on undisturbed samples.

4.4.2. Undisturbed 50mm samples

The Rankine lecture by Rowe (1972) states the required the minimum size of samples for various tests. In Queensland 60% of undisturbed samples are obtained with 50mm tubes (U50) and then pass on to the lab for testing. (Note most other States use larger size samples). The laboratory then tests what sample size is provided, not what the testing standard requires. For example, Look (2023c) shows that for Queensland data, the coefficient of consolidation (c_v) using a 75mm sample size is likely to be 3.5 times the value using a 50mm sample. The analysis that follows is based on that "non-standard" data and cannot then step outside of that wrong result – even when knowing the lab test is wrong. Requesting additional testing at what is likely an advanced phase of design, will most likely be rejected by project managers. Using what is available takes precedence over what is accurate.

The cheaper U50 sample may be appropriate in many cases where no consolidation tests or macrofabric does not govern (such as moisture content, index and density tests or strength tests without macrofabric). There is no singular rule for obtaining sample size. But U50 size samples are inappropriate for consolidation testing or fissured clays.

Using these size samples is also a group think bias as the contract drillers in Queensland use U50 samples as a de facto standard unless told to bring another size sampling tube. Hence suggesting many geotechnical consultants in Queensland use a U50 (as a standard). It is OK to be wrong as long as others are using the same (incorrect) sample size.

4.5. FACTOR OF SAFETY

Factor of safety provides a useful decision tool in slope stability analysis. However, a higher factor of safety does not necessarily indicate a higher probability of failure (**Figure 18**, Lacasse, 2013). Additional data on material and layer variability has to be available to obtain the probability of failure (P_f).

And while the factor of safety does indicate a possible probability of failure, this would vary with each project. Factor of safety should therefore be classified as a probability of failure index only. A probability analysis does require additional data to assess the material variability.

5 GROUP THINK

The availability bias also applies to the examples in this section although for discussions purposes, the group think bias is being emphasised. Group think or herd behaviour bias is related to the false consensus effect and social proof. We gravitate to the popularity of decision or what is commonly done. This suggests that people will adopt certain opinions if the majority is doing it as we strive for consensus. This has served us well as hunter gatherers if one person in the group bolted, it would be unwise to rationalise the correctness of that action (is there a lion or something that looks like a lion or some harmless creature?). Although there is no survival advantage in current times, there is a social benefit. Examples provided are:

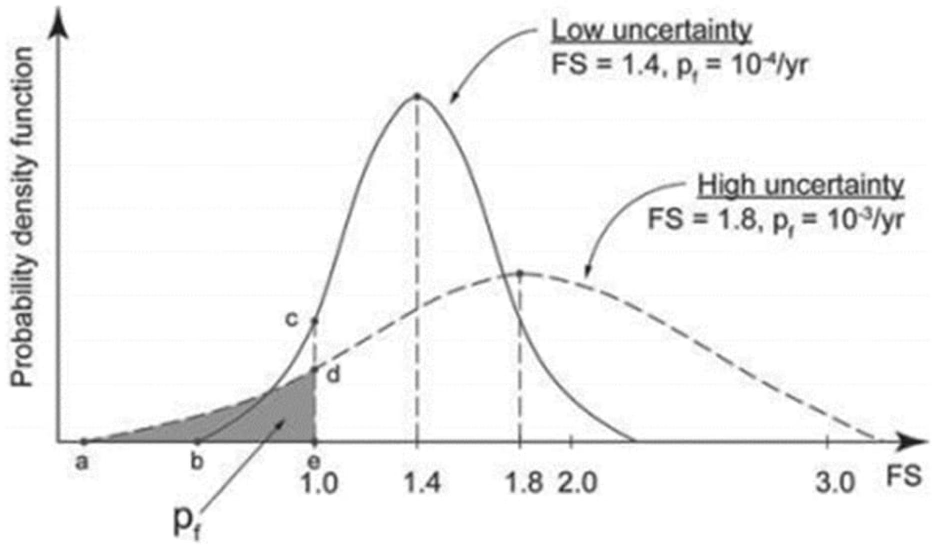


Figure 18: The probability of failure is not always coincident with the factor of safety (Lacasse, 2013)

- o The standard penetration test provides an N - value which is often used directly in design.
- o Applicable correlations
- o Sample size in testing

One should recognise there is a commercial advantage in not pursuing a higher quality test result and this bias is arguably simply a commercial decision to win the job and not a bias.

There is often a correlation bias associated with this group think. Correlation bias refers to different types of fallacies related to the interpretation of correlation. There may be an illusory correlation, which is the perception of a relationship between two unrelated events or traits due to existing biases or expectations. There may also be confusion of correlation with causation, when one assumes correlation implies a causal link.

5.1. THE STANDARD PENETRATION TEST

The Standard Penetration test (SPT) is one of the common in situ tests used in geotechnical engineering to determine properties of subsurface soils. Various corrections should then be made to the in- situ value to allow for type of hammer, energy corrections, etc. (Skempton, 1986). The test evolved from being a sampler only to counting the blows during sampling when Terzaghi (1947) introduced the term SPT. Terzaghi's original strength classification for granular materials was simply three parts:

- o Loose < 35%
- o Medium 35 – 65%
- o Dense > 65%

The current applications includes very loose and very dense. The implications of extrapolating to high and low values are discussed further in section 6.4.3

Some common fallacies with measuring and using the SPT N- values (Look, 2019) are:

1. The N - value is factual
2. $N_{\text{Field}} = N_{\text{Design}}$
3. Similar correlations for all Geology
4. N ≤ 1 implies a clay is very soft
5. N > 50 a pile will refuse.

5.1.1. Test values counting to 150mm increments

The SPT requires an accurate count of the SPT Blows (N-values). Look (2016) showed that 150mm is a counting target. It is not actually measured (**Figure 19**). An eye measurement is 150mm ± 41mm. This represent an 11% counting error in the test. The reported N value counted in 150mm increment therefore represents the closet integer value to that increment. Experienced drilling supervisors logged each of these as 150mm increments. The SPT is not factual but an interpreted number by the field supervisor.

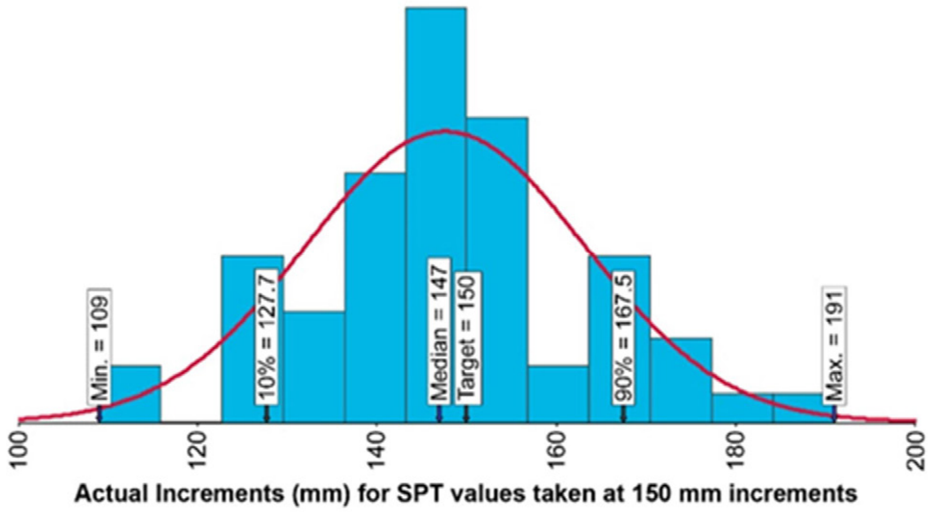


Figure 19: Digital measurements of SPT N- value.

5.1.2. Using SPT value in design

Having obtained the field value (N_{Field}), an appropriate energy conversion is required for the SPT to be applied to design (N_{Design}). Australian Standards do not currently specify energy requirements in the SPTs procedure. Ergo, energy is not measured. Any correction (if applied) is based on the international literature. Most developed countries have a Standard to measure the SPT energy – Australia does not. Depending on one's favourite reference the energy correction can vary from 0.8 to 1.7 for the trip or automatic hammer, which are generally used in Australia (Look 2019). Look (2016) showed a case study, where 3 drilling rigs were in use, that 18% of 1120 SPT measurements in granular materials would change classification. The energy correction varied from 1.5 to 1.0 with the latter value assumed for the drilling rig that was not on site at the time of the energy measurement.

Most design correlations that use the SPT N - value are based on the N_{60} (energy corrected value). Yet engineers in Australia continue using the in-situ N value (N_{SPT}) even though this practice has been outdated for some time.

$$N_{60}\, E_{60} = N_{SPT}\, E_{SPT}$$

The rationale for not applying an energy correction varies from not being aware that a correction is required, or lack of confidence in the wide range of values suggested in the literature (understandable), and not enough budget to carry out an energy correction evaluation. This rationale for not applying an energy correction can be further justified by "every-one else is doing the same". This is an example of group think and social proof. If everyone else is not measuring an SPT energy value then it is not required. Yet an energy correction and using the N_{60} value for design has been widely known since the 1980s (Skempton, 1986), and applies to most the correlations related to the N_{60} values. Australian "standard" practice (although not stated) implies using $N_{60} = N_{Field.}$

5.1.3. High and low SPT values

Another common bias is observed with SPT at high and low values. This may be also due to no energy correction.

Look (1997) highlighted the limitation of Australian Standards for SPTs. This is explicitly titled "The Standard penetration test procedure in **soils**". The "soil" test has limitations when applied to residual soils and weathered rock. Yet that standard is used in practice in Australia in the absence of a Standard **for rocks**.

Using data at the coring interface where an SPT was obtained and a point load index test was carried out within 1.0m, Look (2004) showed that an inferred SPT extrapolation between 60 and 120 had little relevance to strength (**Figure 20**). The inference within that range could be a meaningless range of very low to high strength. The inferred values above 120 show some correlation with strength. The UK and USA standards allow additional consideration for testing in rock which would improve such extrapolations. Additionally, there was a poor correlation if all values were used. When the data was separated to defects less than or greater than 60mm, in the adjacent core run, there was a poor relationship with defects greater than 60mm. The practice of using the cheaper augered holes and an SPT to refusal, instead of obtaining core data results in an unreliable N- value data. This is the availability bias occurring when the inferred SPT N- value is then used for analysis and design.

A driven pile will only refuse if SPT- N > 50. The reverse logic (circular argument) is often applied i.e. if SPT > 50 means a pile will refuse. This is a flawed logic of a singular high SPT value indicating "pile refusal", but is still used by inexperienced (and sometimes experienced) practicing engineers. Adams et al. (2010) present data for driven prestressed concrete piles in Queensland and the refusal criteria based on the SPT N-value factored for geology. This pile refusal (a near constant set mm / blow) varied from:

- $N_{Field} \geq 160$ for shales at a set of 4 mm / blow
- $N_{Field} \geq 90$ for sandstones at a set of 2 mm / blow

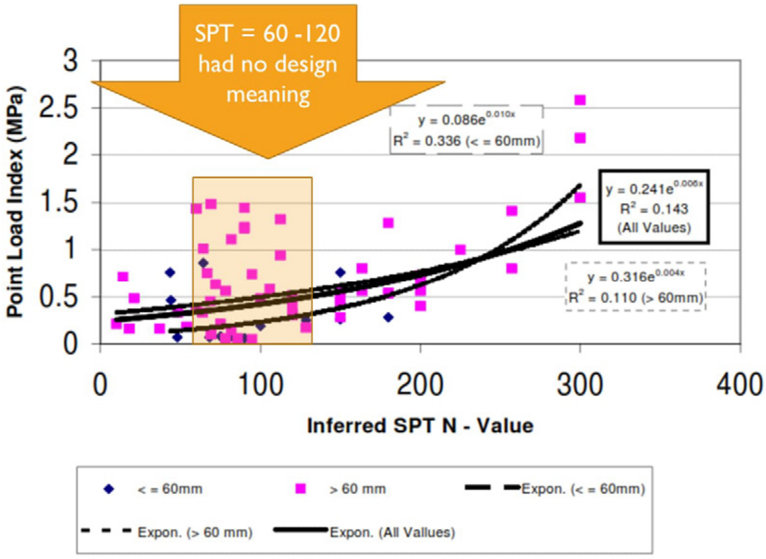


Figure 20: Effect of Defects on N- Value (No. of points = 73 with outliers removed)

The rock or soil type, and the type of pile, influence the depth of driving.

At the other end of the spectrum, many borehole logs show a very soft clay for an SPT N value ≤ 1. A flawed circular argument is being used as follows:

- A clay is very soft, therefore N value ≤ 1 is correct. However, the reverse logic does not necessarily apply.
- If N value ≤ 1 does not mean a clay is very soft.

A simple bearing capacity check using the weight of hammer, rods and anvil will show the SPT split barrel will "fail" into firm clays ($C_u < 50$ kPa) with an SPT N value ≤ 1. Adjacent CPTs or vane shear tests typically provide much higher C_u values. The clay may be very soft, but without an energy correction or other test, the correlation to very soft may be an illusory correlation. The misuse of SPTs in fine grained soils is discussed further in Reid and Taylor (2010).

This is an example of both a circular and group think argument often applied in strength classification.

5.2. THE DENSITY ILLUSION

Dry Density ratios (DDR) are applied widely in quality control for earthworks testing. Yet because of its widespread usage, this now acts as an impediment to the application of alternative and modern day methods of testing (Look, 2024b). Density testing is a lag indicator and unable to provide a reliable indication on the ground strength or modulus when geotechnical and pavement designs are based on these parameters. DDR has low accuracy, but high precision and simplicity. The measurement technology of the time was limited for other key parameters and DR was chosen as the key QC measurement. Look (2018) highlight correlations errors or assumptions in using density ratio. These include assuming:

1. There is a direct and dependable relation with density and CBR, strength or modulus;
2. The reported density measurement is accurate (refer survivorship and availability bias);
3. As the density ratio increases the CBR, strength or modulus increases;
4. The higher energy imparted to the soil via higher number of passes, increases the density ratio and hence also the CBR, strength or modulus;

5. The additional passes improve the DDR because the field density increases with the maximum dry density unchanged for a given soil.

These assumptions are each discussed in Look (2018) as they are all incorrect for some materials. Excessive compaction may result in strength degradation although density ratio may be increasing.

Several devices have been available for the past decades and while used in design should also be used in quality control testing. Density has historically satisfied our basic needs in quality assessment. Performance-based subgrade design aggregates influences in

- Density
- Moisture Content
- Material quality
- Underlying material

Correlating to density only, is blind to the significant influences of the other factors influencing strength and modulus. Most alternative measurements combine all the above in one measurement, while the current approach is to measure each of the above independently. Depending on the equipment there can be different proportions of these factors being measured. A 95% density ratio does not have a singular strength or modulus value as measured with one-to-one testing. The decision is between the past comforts and success of using density measurements (confined to our basic needs) vs the higher order needs and benefits of alternative modern testing equipment and discussed in (Look, 2018, 2023a, 2024b). A 95% DDR does not have a singular strength or modulus value – although this is an implicit assumption (**Table 4**). This is not rock fill testing as all materials conform to "soil" after compaction.

Table 4: Modulus and strength at 95% DDR (Look, 2021)

Fill Material Origin	PLT Ev2 (MPa)	In situ angle of friction φ (°)	
		Smooth	Padfoot
Sandstone	70	45	45
Interbedded Siltstone/Sandstone	40	41	39
Basalt	65	39	43

Comparing DDR to other equipment measurements often leads to a poor correlation, since a multivariate relationship is required. Dendrogram analysis is used to illustrate this inter-relationship (Look 2023a, 2023b). Yet field supervisors often ask for a correlation to provide a linkage with DR as the de facto standard. A higher DR does not necessarily produce a higher strength or modulus – although this is an implicit assumption. Typically, a 95% DR is specified at subgrade level. However, this DR does not correlate directly with modulus values as shown in **Figure 21** for 3 materials in trials tested with 3 different source materials of residual soils at the Toowoomba site (Look 2023a).

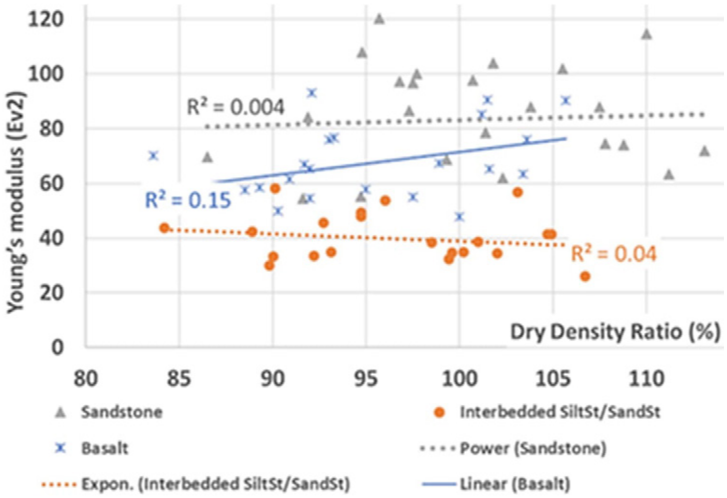


Figure 21: Poor correlations between modulus with DDR.

In residual soils and weathered rock materials being compacted to a "soil" embankment, the material structure is destroyed at higher number of passes. The DDR is increasing but quality test using different test equipment all show the strength and modulus is not increasing.

The emphasis on density has led to the (incorrect) belief that it is the key parameter, yet it is an index only, i.e., we assume an increased density ratio means an increased strength or modulus or reduced permeability. Density has historically satisfied a 1 parameter need in QC assessment. Other quality parameters are assessed independently.

The availability and group bias occurs when we assume the density test relates back to the design strength or modulus. Even with "modern" tests available (pre 2000), the quality test requires such direct modulus tests to correlate back to the density (which is an index only). An incorrect anchoring bias.

5.3. CORRELATIONS AND THE DESIGN CHARACTERISTIC VALUE

Look and Griffiths (2001, 20024) investigated the correlation between Point Load Strength Index ($I_{s(50)}$) and Unconfined Compressive Strength (UCS) test results for rocks within the Brisbane area. Two key findings of that investigation were:

- The UCS to $I_{s(50)}$ multiplier of 24 commonly used at the time (2001) should not be applied for Brisbane rocks.
- UCS ~11.2 $I_{s(50)}$ applied for most Brisbane rocks (Figure 22**:** a). Varying from 5 for the Phyllites to 18 for the Brisbane Tuff unit. The effect of test direction on the correlation is shown in **Figure 22**. Trends do change when sub groups ae considered. This is Simpson's Paradox with misleading aggregation of data.

This shows how established correlations derived from elsewhere (at that time) did not apply – a group think bias. The availability bias occurs when an established correlation is then applied to the Brisbane rocks.

The availability of trendline options in an Excel spreadsheet, also does not mean one can correlate all values. A reasonable high R^2 value is a first indicator of correlation, but does not always make the correlation valid. For a correlation to be valid and to avoid inference errors one must evaluate outliers and non-normality.

The trend lines are shown for axial and diametral values and with all data points at the Gateway Bridge site (**Figure 23**) (Look and Wijeyakulasuriya, 2009). The box and whisker plots show outliers which represents less than 1% probability (**Figure 24**) and should be removed. Other statistical analysis techniques such as examining residual values or Q-Q plots (**Figure 25**) would show such inconsistencies if all values are used without removing the outliers. The Q-Q Normal Plot creates a quantile-quantile (Q-Q) plot for a single variable. If the data are essentially normal, then the points on the Q-Q plot should be close to a 45-degree line. The 3 outlier diametral residual points (**Figure 24**) are seen as high z values. The z values measure how many standard deviations away a data point is from the mean of a distribution.

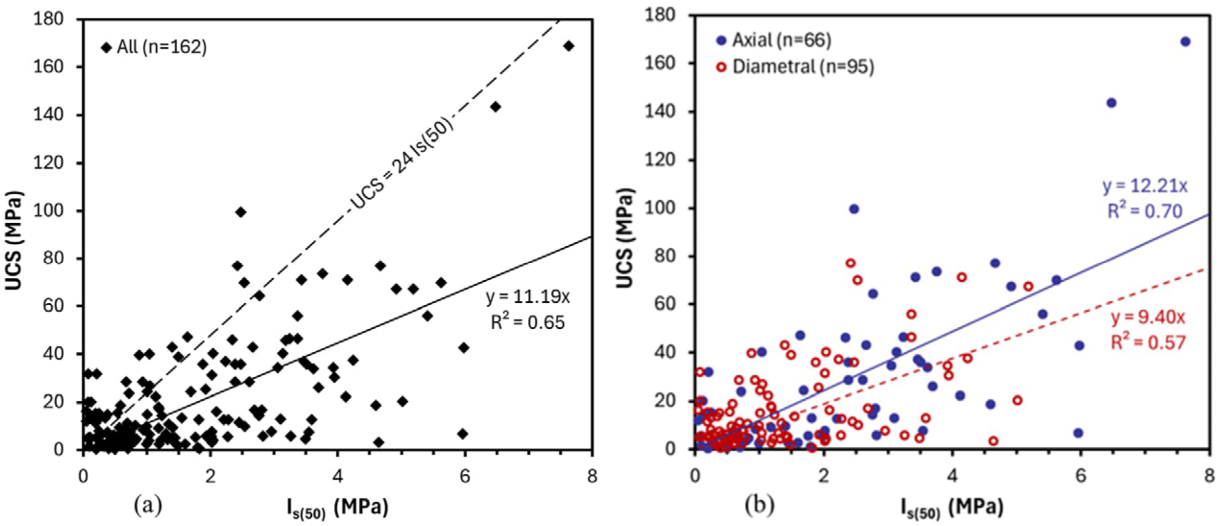


Figure 22: Brisbane rock strength (a) UCS to $I_{s(50)}$ correlation (b) showing effect of anisotropy (Look and Griffiths 2001, 2004)

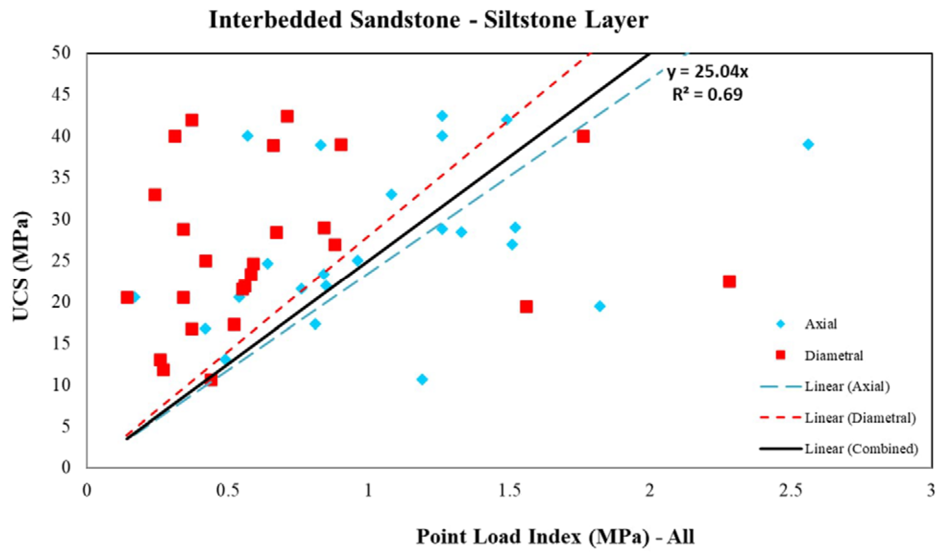


Figure 23: Gateway Bridge Pier 6 - UCS to $I_{s(50)}$ correlation

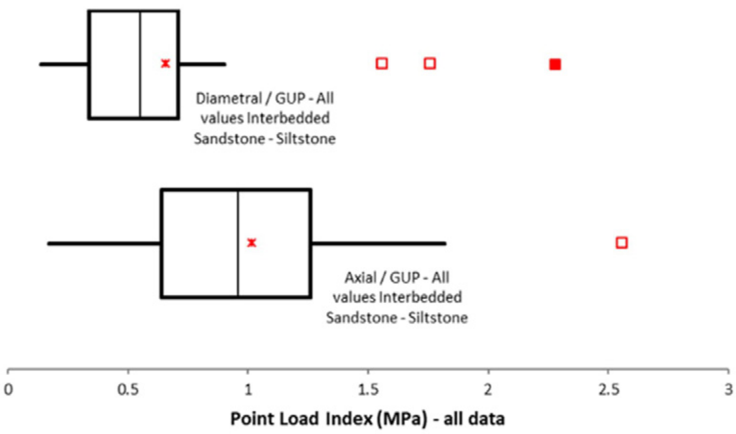


Figure 24: Gateway Bridge Pier 6 – Box and whisker plots showing outliers

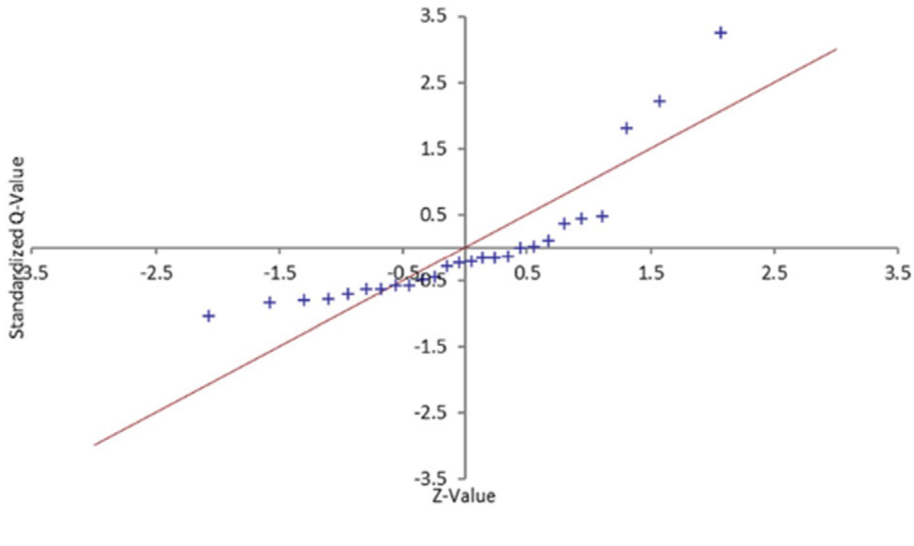


Figure 25: Gateway Bridge Pier 6 – Q-Q normal plots of diametral values $I_{s(50)}$

Determining the characteristic design value is a key task in any geotechnical analysis. Applying a default normal probability distribution function (PDF) has limitations in processing geotechnical data. Look (2015) show case studies when the normal PDF should not be used to avoid statistical irregularities. One case study is the Gateway bridge where the rock socket design is based on the rock strength, derived from point load index values (**Figure 26**). Using goodness of fit tests various PDFs are compared and ranked. A 10-percentile strength value with a best fit PDF is 8 times higher than the normal PDF, which predicts an unrealistic negative value with 9% likelihood. Where large Coefficient of Variations (COVs) occur, the best fit PDF is consistently not the Normal PDF, and should be avoided in the statistical assessment of parameters (Look, 2025).

The use of automatically generated trend lines and normal PDF are examples of both availability and group think biases.

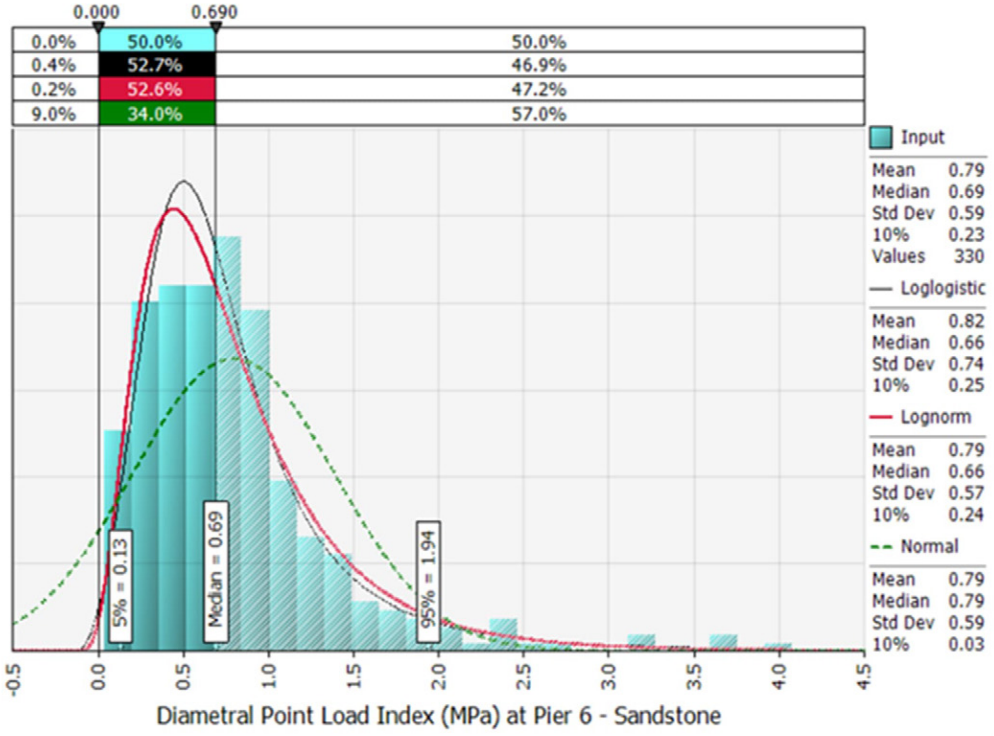


Figure 26: PDFs for the diametral I_s (50) at the Gateway Bridge Pier 6

6. SUNK COST EFFECT

The sunk cost effect is manifested in a greater tendency to continue an endeavour once an investment in money, effort, time or other resources has been made. Even if a new course of action would result in better outcomes, there is a psychological desire to continue with the existing course of action so that the effort expended so far has not been a waste.

A typical example occurs within acquisition projects. Suppose a project spends several million on a capability option. If a cheap alternative is consequently offered that would save money overall, the sunk cost effect suggests that the current, more expensive option will still continue in order to prevent those millions already spent from being perceived as wasted. But this is not rational if an overall saving could be achieved.

In the geotechnical engineering (GE) world we tend to "specialise". The team involved in the site investigation (SI) may not be involved in analysis and those involved in analysis are often not involved in design or construction. During analysis or design, the modeller often would wish improved SI field or lab data not covered, but as that money is from a different bucket, The analysis then proceeds with the missing data as they are often unable to request that completed SI phase of the work. Analysis and design often proceed in parallel with the SI as time for delivery often governs a project. Conservative analysis parameters are often not changed even when SI data shows otherwise. With project budget limitations rework may not then occur. A preliminary vertical and horizontal road alignment may not be the final design alignment. The SI may then not be to the required depths in some instances and tests are off the final alignment. Without a staged SI many geotechnical reports have missing data or irrelevant test data.

Look (2024b) showed a case study where geotechnical test data (over 2,000 pages) was provided to the contractor as part of the tender information for 4 road widening areas. There were 100s of test certificates with no summary data on location provided. Yet due to initial to final design changes, most of the test data provided no longer applied to the tender documents. The test data did not summarise where the testing was located, but the sunk cost test data was still provided as "background" to tenderers. The tenderer's quantity surveyor incorrectly used all data provided for each site to assess material quality for cut to fill and with a resulting contractual dispute when site data did not reflect site conditions. **Figure 27** shows CBR tests at site 4, where only 21% of pre tender site data was relevant. Other alignment test data such as Atterberg and Emerson tests suffered a similar issue. The post tender testing (during construction) did not reflect that "suitable" cut to fill material which was estimated to be available at this site 4.

The sunk cost test data provided to tenderers provides data noise and information bias.

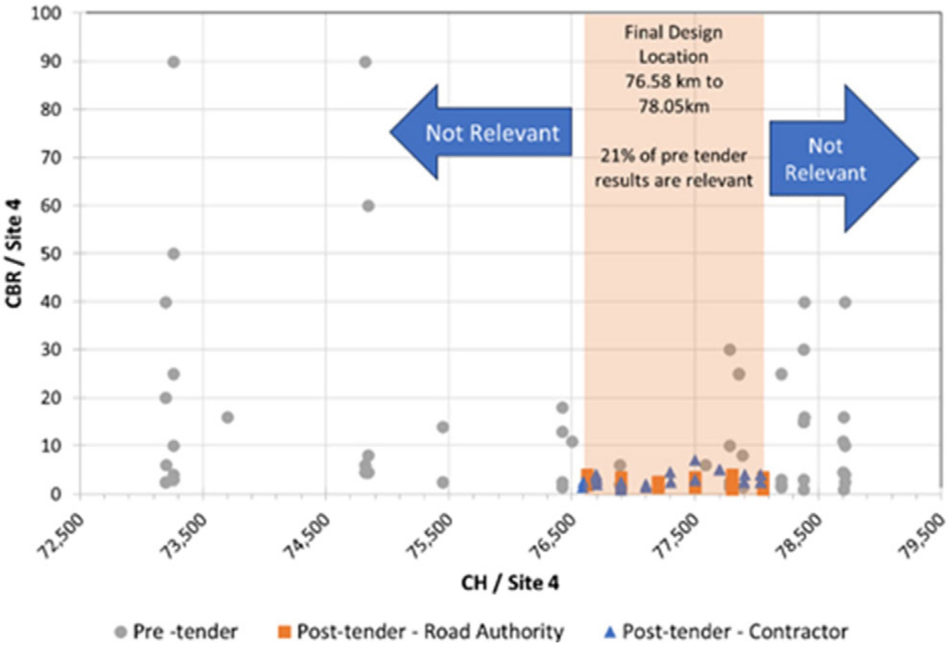


Figure 27 CBR data at site 4 - pre and post tender.

7. CONCLUSIONS

This paper highlights some of the many biases occurring in practice. There are many useful heuristics that have served industry well but such anchoring indices points may not be relevant to all projects. Usefulness is not the same as truth. **Figure 28** provides a likely "truth" in some of the commonly useful geotechnical indices. Some of these were discussed and categorised with survivorship, availability, group think and sunk cost biases illustrated

- A soil grading "fact" may have a sampling bias;
- The plasticity index may have a survivorship bias when used in residual soils
- A factor of safety is a useful index but may not necessarily provide the probability of failure;
- The lab model of OMC and MDD are useful construction indices but does not consider the related CBR or EMC more relevant to long term conditions
- Use of a trend line automatically generated or normal PDFs has associated bias with outliers included

The industry seems to be polarizing towards those who carry out site investigations and those who carry out analysis. This discussion on site characterisation and assessment sits between these two "specialisations". Models do not occur only at the analytical phase. Field and laboratory tests are also test models. There are also statistical models before the analytical model. We need to be aware of the cognitive biases that exists in "Standard" test practices. Even a NATA registered reported test result is not 100% fact on site conditions. Such biases transfer to the analytical and design model.

The heuristics used for quick decisions is still useful and should not be completely abandoned because it is not perfect. Being aware of the associated biases was highlighted.

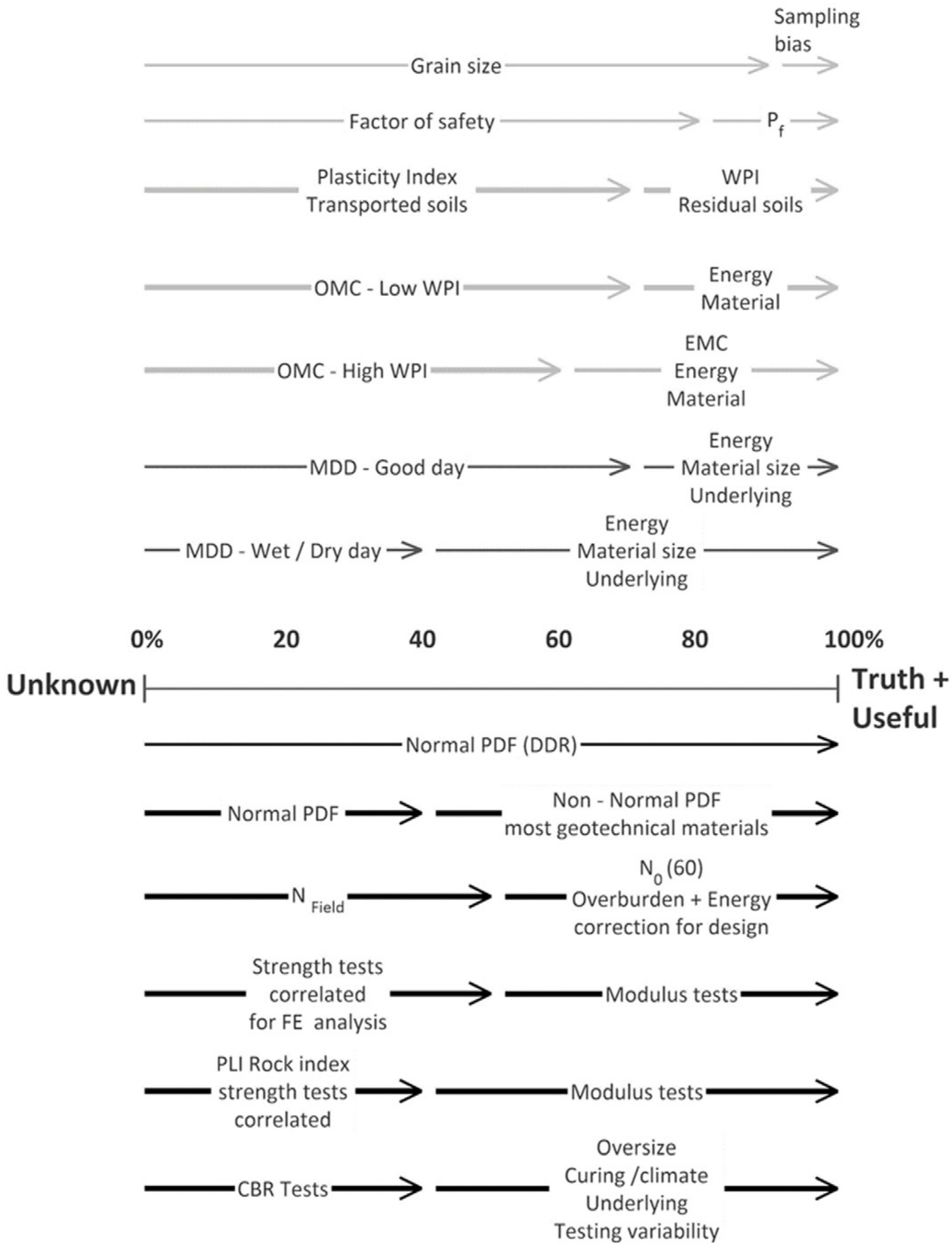


Figure 28: Degrees of truth and usefulness in various geotechnical indices

ACKNOWLEDGEMENTS

The concepts and ideas used herein were first formulated in post graduate courses in philosophy (critical thinking) some 25 years ago at The University of Queensland. This was followed by many books and favourite DVDs such as Steven Novella (2012) from The Great Courses who states, “Questioning our own motives, and our own process, is critical to a sceptical and scientific outlook. We must realize that the default mode of human psychology is to grab onto comforting beliefs for purely emotional reasons, and then justify those beliefs to ourselves with post-hoc rationalizations”. Micheal Roberto (2012) states “people around you filter information, often with good intentions” and in his course, examines why leaders and organizations make poor choices, digging deep into cognitive psychology to help us understand why well-intentioned, capable people blunder. Rolf Dobelli (2013) provide examples and cognitive biases in “The Art of Thinking Clearly”. These favourite books and lectures have each been listened to or read dozens of times and set the scene for exploring “Cognitive dissonance in geotechnical engineering”.

As part of the Australian Geomechanics Society (AGS) practitioner award (2018), this presentation was planned to be delivered at local Chapters in mid-2020. However, even when COVID-19 restrictions were first lifted and plans were again made some States were still not covered due to the on again – off again border restrictions. This paper documents a few of the concepts presented at local chapters.

CRediT authorship contribution statement

Burt Look: Writing - original draft.

8. REFERENCES

Adams, B., Look B., Gallage, C. (2010). Assessment of pile driving refusal using the standard penetration test. *Proceedings of the 11th International Association of Engineering Geologists Congress,* Auckland, New Zealand, September, CD-ROM, pp 2641 – 2648.

Chew K.C, Look B., Boyer M (2012). Environmental Evaluation of Landslide Remediation Options. *11th Australia New Zealand Conference in Geomechanics*, Melbourne, pp 314 – 319.

Boyer M, Chew K and Look B (2013). A case study of a hillside slope instability in south east Queensland, *Australian Geomechanics Journal*, Vol 48, No. 1, pp 77 – 84.

Dobelli Rolf (2013). *The Art of Thinking Clearly*. Sceptre publishers.

Lacasse, S. (2013). 8th Terzaghi Oration: Protecting society from landslides – the role of the geotechnical engineer. *Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering*, Paris, France, September 2-6, 2013.

Lambe, T. W. and Whitman, R. V. (1979), "Soil Mechanics", *John Wiley and Sons*, New York.

Lacey D.W, Look B.G. and Marks D.F. (2016). Use of the Light Falling Weight Deflectometer (LFWD) as a site investigation tool for residual soils and weak rock. *5th International conference on geotechnical and geophysical site characterisation,* Gold Coast, Australia.

Look, B.G. (1995), "The effects of volumetrically active clay embankments on roadway performance", PhD thesis, *The University of Queensland.*

Look, B G (1997). The Standard Penetration Test Procedure in ~~Soil~~ Rock. *Australian Geomechanics Journal*, No. 32, December, pp 66-68.

Look B. G. and Griffiths S. G. (2001). An Engineering Assessment of the Strength and Deformation Properties of Brisbane Rocks. *Australian Geomechanics Journal*, Vol 36, No. 3, pp 17 – 30.

Look B. G. and Griffiths S. G. (2004). Rock Strength Properties in Southeast Queensland. *9th Australia New Zealand Conference in Geomechanics, Auckland, N.Z.*, Vol. 1, pp 196 - 203.

Look B G (2004). Rock Strength at the coring interface. *Australian Geomechanics Journal*, Vol 39, No. 2, pp 105 – 110.

Look B G (2004). Effect of Variability and Disturbance in the measurement of Undrained Shear Strength. 9th *Australia New Zealand Conference in Geomechanics*, Auckland, N.Z., Vol. 1, pp 302 - 308.

Look, B G (2005). Equilibrium Moisture Content of volumetrically active clay earthworks in Queensland", *Australian Geomechanics Journal*, Vol 40, No. 3, pp 55 – 66.

Look, B. G. and Wijeyakulasuriya, V. (2009). The statistical modelling of rock strength for reliability assessment. *17th Int. Conf. on Soil Mechanics and Foundation Eng., Alexandria*, Vol. 1, pp.60–63.

Look B and Campbell N (2013). Variability in selecting design values in Queensland - A calibration of professional opinions. *Australian Geomechanics Journal*, Vol 48, No. 1, pp 51 – 66.

Look, B.G. (2015). Appropriate distribution functions in determining characteristic values. *12th Australia New Zealand Conference in Geomechanics*, Wellington, New Zealand, P014

Look B (2016). The Weighted Plasticity Index in Road Design and Construction. *Australian Geomechanics Journal*, Vol 51, No.3, pp 55 – 66.

Look, B.G. (2018). Compaction QA Limitations: Benefits of alternative testing methods. *AGS 2018 Victoria Symposium on Geotechics and Transport Infrastructure*.

Look B.G. (2019). Geotechnical basics – Common practices to avoid. *13th Australia New Zealand Conference in Geomechanics*, Perth

Look, B.G. (2021a). An earthworks quality assurance methodology which avoids unreliable correlations. 4th *International Conference on Transportation Geotechnics*, Chicago, USA. In: Tutumluer, E., Nazarian, S., Al-Qadi, I., Qamhia, I.I. (eds) Advances in Transportation Geotechnics IV. Lecture Notes in Civil Engineering, vol 166. pp 179 – 192, Springer.

Look, Burt G (2021b). Strength and modulus changes associated with compaction of residual soils and weathered rock. *Australian Geomechanics Journal*, Vol 56, No.1, pp 79 – 104.
Look, Burt G (2023a). Earthworks testing and the density illusion. *Proceedings of the 14th Australia and New Zealand Conference on Geomechanics*, Cairns 2023 (ANZ2023).
Look, Burt G (2023b). Cairns, Smithfield Bypass Project: Site for alternative quality testing equipment. *Proceedings of the 14th Australia and New Zealand Conference on Geomechanics*, Cairns 2023 (ANZ2023).
Look, Burt G (2023c). Assessment of the coefficient of consolidation with Queensland data. *Australian Geomechanics Journal*. Vol 58, No. 2, pp 105 – 120.
Look, Burt G (2024a). Precision measurements constrain the implementation of more accurate testing in earthworks quality assurance. *Proceedings of the XVIII European Conference for Soil Mechanics and Geotechnical Engineering*, Lisbon, Portugal ECSMGE 2024
Look, Burt (2024b). Material subgrade variability in site characterisation. *Proceedings of the 7th International Conference on Geotechnical and Geophysical Site Characterization*, Barcelona, June 2024
Look B. G. (2025). The effect of coefficient of variation and distribution functions in determining characteristic values. *Proc. of the 9th International Symposium for Geotechnical Safety and Risk (ISGSR),* Oslo, Norway, August 2025 (unpublished to date)
Nakao T and Fityus S (2008). Direct Shear Testing of a Marginal Material Using a Large Shear Box. *Geotech. Testing Journal.* Sep 2008, Volume 31, No. 5, pp 393-403
NAVFAC (1986) Soil mechanics – design manual 7.1, *Department of the Navy, Naval Facilities Engineering Command,* U.S. Government Printing Office, Washington, DC
Novella Steven (2012). Your Deceptive Mind: A Scientific Guide to Critical Thinking Skills. *The Great Courses, audio/video lecture series,* The Teaching Company DVD.
Reid Alan and Taylor John (2010). The misuse of SPTs in fine soils and the implications of Eurocode 7. *Ground Engineering,* July, pp 28 – 31.
Roberto Micheal (2012). The Art of Critical Decision Making. *The Great Courses audio/video lecture series*, The Teaching Company DVD.
Rowe P.W. (1972), The Relevance of soil fabric to site investigation practice” *Geotechnique*, Vol. 22, No. 2, pp 195 – 300.
Skempton A.W. (1986). Standard Penetration Test Procedures and the Effects in Sands of Overburden Pressure, Relative Density, Particle Size, Ageing and Over-consolidation. *Geotechnique,* 36 (Sept), pp 425 – 47.
Standards Australia *AS 1289 6.31. – 1993.* Determination of the penetration resistance of a soil - Standard Penetration Test. *Standards Association of Australia, NSW.*
Standards Australia. AS 1289.6.2.2-1998. Methods of testing soils for engineering purposes: Soil strength and consolidation tests – Determination of the shear strength of a soil – Direct shear test using a shear box. *Standards Association of Australia,* NSW.
Standards Australia A.S. 2870 - 2011, “Residential Slab and Footings - Construction” *Standards Association of Australia,* NSW.
Transport for NSW (TfNSW- 2015), Field density testing by using a nuclear density gauge. *Technical Guide L-G-002. Roads and Maritime.*