On Using Wildly Invalid VI Scales

Objection to the Proposed Jupiter Wind Farm

Dr Michael Crawford

February 13th 2017

Cloustons VI Assessment methodology uses a number of dimensions (e.g. *Distance of View, Quantum of View, Viewer Sensitivity, and Visual Impact*) with *Visual Impact* estimates being derived from a combination of scores on measurement scales created for each of the other dimensions. ***Every one of Clouston's scales is fatally flawed in multiple ways***.

In addition Cloustons has created composite dimensions (*Magnitude* and *Visual Impact*) which depend on some combination of other dimensions. However, Cloustons has failed to specify the formulation through which scores on the various subscales are transformed into scores on the composite scales. Data in their report shows significant evidence of ***arbitrary application*** of whatever is their underlying formulation for determining scores on the *Magnitude* scale.

Nor has Cloustons provided any evidence of validating the scales they have presented or the predictive validity of the combination of *Magnitude* and *Viewer Sensitivity* to predict *Visual Impact*.

All scales have been structured in ways that cause scale truncation at the upper end, thus suppressing the magnitude of scores that can be reported, to the advantage of the developer and disadvantage of the local community.

Application of most of the scales depends on judgements by Clouston's staff but Cloustons have provided no data on the reliability of those scales in use, despite published research showing they are generally very low.

The methodology is the antithesis of rigour in scale construction and application. Since every assessment in the VIA depends on applying the methodology, it follows that every assessment and the conclusions in the VIA are consequently hopelessly invalid and must be rejected.

Measurement scales are a fundamental part of most science-based disciplines, including the social sciences. Professionals in those fields understand that scales must be constructed with care and in particular they must be validated, especially when used to predict scores on some other dimension. They also understand that there are limits on how scale values may be aggregated consequent on the numeric character of the scales. None of this awareness is apparent in the Jupiter Visual Impact Assessment.

Clouston's whole VI assessment (VIA) methodology rests on the application of a set of scales, and the use of those scales to predict *overall visual impact*. No evidence is presented in the EIS of:

- validation of scale construction
- awareness of the numerous problems in such scale construction
- awareness of the problems in subjective application of such scales
- validation of the use of the scales to accurately predict the claimed result (i.e. *overall visual impact*)

On the evidence (and lack of it) provided in the VIA, the whole basis of the methodology is scientifically bereft, and contrary to well established principles that apply to professionally constructed scales.

Consequently Cloustons has put forward a methodology that claims to offer rigour when in scientific terms it is the antithesis of rigour in scale construction and application. The conclusions in the VIA therefore are invalid and must be rejected.

A delicatessen or petrol station which dealt as casually with measurement would feel the full wrath of the National Measurement Institute and be shut down until it was corrected – despite the consequences being minute compared to invalid assessment in a wind farm VIA.

Hopefully the Department of Planning and Environment has someone with a technical understanding of the process of scientifically valid scale construction and validation.

The proponent must be told to produce a VIA using tools for which the validity has been scientifically and transparently established.

Clouston's VIA Use of Scales

Any matrix which purports to provide a result in its constituent cells based on the values of the vertical and horizontal dimensions inherently uses three scales. Each of the axes constitutes a scale. The results expressed in the cells also constitute a scale if they are ordered categories, as is the case in the VIA.

Cloustons VIA uses a matrix[1] with the dimensions *Magnitude* and *Sensitivity*. Each of those is presented as an *ordinal* scale (i.e. the magnitude on the scale increases as you move from one end to the other).

[1] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 59, Table 4.4.

The cells in the matrix contain scores on another scale *Overall Visual Impact* (or just *Visual Impact*), which also is presented as a set of ordinal values ranging from *Negligible* to *High*.

In addition, the scale *Magnitude* is actually a composite of four other notionally ordinal scales:

- *Quantum of View*
- *Distance of View*
- *Period of View*
- *Magnitude of Change*

Each of those introduces its own problems of scale construction, as does their consolidation into a *Magnitude* score, which will be discussed later.

Pervasive Defects in Clouston's VIA Scales

Appendix A provides an extensive dissection and evaluation of the seven scales used in Clouston's VIA. ***Every one of those scales is fundamentally defective in terms of scientifically valid scale construction***. Yes **EVERY ONE**.

Any graduate student at an Australian university who presented up even a couple of scales with these problems would be sent to an extensive remedial program.

In summary, the defects in the scales (explained in detail in Appendix A) are:

Quantum of View:

- **Conceptually flawed scale based on applying ideas of urban views to rural locations**
- **Category definitions reflect constructor bias**
- **Application of categories is subjective**
- **Inappropriate truncation created at upper end of scale**

Distance of View:

- **Conceptually objective scale which has been operationalised in a distorted manner**
- **Category cut points show extreme constructor bias which advantages the developer**
- **Category cut points inconsistent with VIA's own table of wind farm visibility**
- **Category cut points dramatically inconsistent with all independent research on wind farm visibility**
- **Inappropriate truncation created at upper end of scale**

Period of View:

- **Conceptually confused and consequently unusable scale**
- **Category cut points actually have important gaps**
- **Where categories are specified in terms of time duration, no indication of the application of those durations (e.g. is it per occurrence, per day, per month?)**
- **No justification provided for why different periods warrant one descriptive category rather than another**
- **Inappropriate truncation created at upper end of scale**

Magnitude of Change:

- **A wholly subjective set of categories based on abstract terms**
- **Claims to be "a quantitative assessment" but is not quantitative**
- **Has inconsistent definitions of what warrants a high score on the scale**
- **The dimension is confounded with the dimension *Quantum of View***
- **Inappropriate truncation created at upper end of scale**

Magnitude:

- **A composite scale for which the formulation that generates scores on the scale has not been revealed**
- **Reported results include obvious inconsistencies demonstrating arbitrariness in application**
- **Failure to provide any validation for the formulation of the scale**
- **Inappropriate truncation created at upper end of scale**

Viewer Sensitivity:

- **The operationalisation of the scale is inconsistent with the definition**
- **The definition refers to the sensitivity of individual people, while the category definitions are in terms of the characteristics of locations**
- **Operationalisation of the scale is confounded with magnitude of exposure**
- **Distinction between the *High* and *Moderate* categories is subjective**
- **Inappropriate truncation created at upper end of scale**

Visual Impact:

- **A number of the category descriptors are identical with those of other categories, just described with slightly different wording**
- **In application, the difference between categories, especially adjacent ones, is highly subjective**
- **There is no sense of magnitude associated with the "impact" referred to by each of the categories and therefore no sense of quantum of difference between them**
- **The scale is a rehash of the Magnitude of Change dimension (i.e. it is not a distinct scale)**
- **The scale is conceptually confused and fails to actually reflect Visual Impact as something perceived by viewers**
- **Scale attenuation is achieved through creative labelling which serves the interest of developers rather than parties actually subject to Visual Impact**

So, as the descriptions above reveal, every one of the scales used in the VIA is fatally flawed in multiple ways.

Reliability Testing and Predictive Validity

In social science there are occasions when researchers aggregate scores on multiple ordinal items to constitute an overall score. These are commonly known as Likert scales and there are important characteristics that differ from what has apparently been done in the VIA:

- the constituent items that are averaged are typically meant to be indicators of the same subjective thing, e.g. personal happiness, depression, or perceptions of the fitness of applicants for a role;

- there are tests done to determine the consistency of scores from individuals (e.g. test/retest) and the ability of scores to differentiate between subject instances;
- when the scales are of something external to raters, there are tests to measure the consistency of ratings given by different raters about the same test subject, i.e. the reliability of scores obtained with the scale (inter-rater reliability or generalisability coefficient); and
- critically the predictive validity of the scales are statistically evaluated, either directly in relation to the matter a scale is meant to express or in relation to other outcome variables which have an hypothesized dependence on the conceptual content of the scale.

It is quite clear that while the constituent subscales for *Magnitude* and *Visual Impact* might each hypothetically contribute to the magnitude of those two variables, they are not simply interchangeable or even correlated (e.g. for instance, there seems no reason to expect a correlation between *Distance of View* and *Period of View*; and *Magnitude* and *Viewer Sensitivity* are defined as two quite distinct concepts).

The scientifically appropriate way to determine the magnitude of contribution of each dimension to *Visual Impact*, and any relationships between them in so doing, is via analysis of variance or regression analysis, which allows the testing of more complex equations than simple linear ones. Unfortunately except in very simple cases regression analysis depends on having at least interval scales rather than ordinal ones and, as noted in this review, none of the scales used by VIA could be considered interval data. They are all, at best, ordinal scales.

There is no evidence offered in the VIA of any attempt to demonstrate predictive validity of the combination of dimensions used to ultimately derive *Visual Impact* scores. This may be because the flaws in each of the scales make it virtually impossible to do.

But saying it is very hard to validate a scale you have decided to create does not then offer any legitimacy for assuming the scale is valid. In fact, the reverse tends to be true. If you can't validate, because you can't otherwise measure the outcome of interest, it means any claim you make about knowing the magnitude of that outcome in particular situations is nothing more than a guess.

The same is true in relation to the predictive validity of a combination of dimensions being used to predict another dimension (in this instance *Visual Impact*).

Cloustons has done nothing to demonstrate the validity of its schema and examination of the multiple defects in that schema leads to the inevitable conclusion that the schema simply has no validity whatsoever and cannot be relied upon by the Department or anyone else.

Rater Unreliability

In addition to the multitude of inherent flaws in its scales, the Clouston's schema depends on individuals making subjective judgements using each of those scales. The expected consistency of different raters (what is termed inter-rater reliability) in using the scales is a critically important factor.

Cloustons provides no inter-rater reliability data for the scales, despite that being the norm in scientific and professional endeavours when using methodologies that depend on individuals providing ratings using various scales. However, generic research on the matter is available.

The *Evaluation of Methodologies for Visual Impact Assessments* review draws on peer-reviewed research which demonstrates that professional ratings of VI have low validity in predicting the VI actually experienced by people who live near a development. There are two reasons discussed in that review:

- Research studies show that the inter-rater reliability of professionals (i.e. the consistency between different individuals) when assessing the various factors commonly used to rate visual character is low, and the reliability of assessments about the difference between before and after a development are even lower[2]; and
- "The difference between what professionals value and what the public values is profound." [3]

Note the second part of that first point. The reliability of assessments about the difference between before and after a development is particularly low.

Feimer, *et al* investigated the reliability of ratings provided in tasks forming part of visual impact assessments. The research participants in their study "consisted of 60 BLM (Bureau of Land Management) staff members whose agency duties included VIA". [4]

They found[5] that "the rating of the degree of change to be imposed upon the landscape by land use activities is characterised by very little agreement when the number of raters is relatively small" and that "A perusal of the reliabilities for contrast and direct impact ratings (importance and severity) indicates an even more problematic state of affairs (see Table 2). The importance and severity ratings for both treatment conditions are in each case .05 or less, indicating virtually no agreement."

Note, perfect agreement by different raters would be indicated by an inter-rater reliability score of 1. So a score of .05, as reported by Feimer, *et al*, clearly means virtually no agreement between the various raters, despite the fact those raters were commonly involved with visual impact assessment duties for the US Bureau of Land Management.

So Clouston's methodology involves a process where research teams have found in multiple studies that reliability of judgements is low. It is incumbent on Cloustons to demonstrate that is not the case with their methodology. However, they have chosen not to do so, and therefore have no basis upon which to claim the multiple research findings do not apply to what they have produced.

---

[2] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, Transportation Research Board of the National Academies, Washington DC, 2013, pp. 34-37 and 39-40.

[3] *Op cit*, p. 139.

[4] Nickolaus R. Feimer, *et al*, "Appraising the Reliability of Visual Impact Assessment Methods", Presented at the *National Conference on Applied Techniques for Analysis and Management of the Visual Resource*, Incline Village, Nevada, April 23-25, 1979, p. 290.

[5] Nickolaus R. Feimer, *et al*, *op. cit.*, p. 292.

Summary

Cloustons VI Assessment methodology uses a number of dimensions (e.g. *Distance of View, Quantum of View, Viewer Sensitivity, and Visual Impact*) with *Visual Impact* estimates being derived from a combination of scores on measurement scales created for each of the other dimensions. ***Every one of Clouston's scales is fatally flawed in multiple ways***.

In addition Cloustons has created composite dimensions (*Magnitude* and *Visual Impact*) which depend on some combination of other dimensions. However, Cloustons has failed to specify the formulation through which scores on the various subscales are transformed into scores on the composite scales. Data in their report shows significant evidence of ***arbitrary application*** of whatever is their underlying formulation for determining scores on the *Magnitude* scale.

Nor has Cloustons provided any evidence of validating the scales they have presented or the predictive validity of the combination of *Magnitude* and *Viewer Sensitivity* to predict *Visual Impact*.

All scales have been structured in ways that cause scale truncation at the upper end, thus suppressing the magnitude of scores that can be reported, to the advantage of the developer and disadvantage of the local community.

Application of most of the scales depends on judgements by Clouston's staff but Cloustons have provided no data on the reliability of those scales in use, despite published research showing they are generally very low.

The methodology is the antithesis of rigour in scale construction and application. Since every assessment in the VIA depends on applying the methodology, it follows that every assessment and the conclusions in the VIA are consequently hopelessly invalid and must be rejected.

Thus the content of the VI Assessment is false or at best misleading, since the instruments used to determine effect are not just flawed but biased in multiple ways which favour the developer.

Whether that was done knowingly by the proponent and its agents is for the Department to investigate. The consultants ought to have known that what they were presenting was false or at least misleading – unless it is assumed that VI consultants are not expected to have any knowledge about the scientifically valid way to develop and validate the scales they create and use to perform their assessment.

Appendix A: Detailed Analysis of Clouston's VI Scales

The Cloustons VI schema uses scales for the following dimensions:

- *Quantum of View*
- *Distance of View*
- *Period of View*
- *Magnitude of Change*
- *Magnitude*
- *Viewer Sensitivity*
- *Visual Impact*

Each of those scales is examined in detail below. Every one of them has multiple defects as scales for what they are purported to measure.

In addition, they all share one particular critical defect. That defect is scale truncation, whose effect is to constrain the size of scores that can be reported by the scale.

Scientifically valid scales need sufficient granularity to adequately express significant diversity within the population of interest. It is therefore important to avoid scale attenuation, i.e. compressing the scale into too few categories or truncating it at one end.

The nature of scale attenuation with the *Visual Impact* scale is described in detail in the relevant section below. All the other scales use the same form of scale truncation, which is addressed here.

They all use four categories, labelled as *Negligible*, *Low*, *Moderate*, and *High*. From the wording, it is clear that the *Moderate* category is meant to indicate some form of mid point in the scales. However, the scales then have two categories below that "mid point", yet only one category above it. The effect is to truncate scores at the upper level of the scale.

This is like constructing a thermometer so that it can never read above 30°C. Then when people are sweltering in what would otherwise be recognised as 35° or 40° heat, you assure them that "No, it's only 30°" – because that's all your special thermometer allows to be shown.

With a well constructed scale, there will be few occurrences of the extreme scores, unless the extremes express an absolutely fixed limit (e.g. zero where negative numbers are impossible).

For instance, the RFS has a Fire Danger Rating scale. At the upper end the categories are *High, Very High, Severe, Extreme, Catastrophic*. Each of those categories denotes a high level of fire danger but there are significant differences in the degree of danger as one progresses up the scale. Importantly, actual instances of *Catastrophic* days are quite infrequent. Consequently, it reasonably marks the end point of the scale.

This is also why the top temperature on Australian household thermometers is typically 50°C. We don't expect to see that temperature, but temperatures in the forties are certainly possible, though rare.

With the Clouston's four point scales, the *High* category is ***not*** infrequent[6], indicating that the scales have been truncated, as we would logically expect given two categories below the "mid point" and only one above, for dimensions which have no natural fixed limits at the upper end.

This is emphasised by the fact that in that data, no instances occur of the lowest category, *Negligible*. So one category of the four turns out to be absent from the data and all the occurrences fall into the three higher categories.

The aberrant distribution of scores does not stop there. Across the five dimensions using four point scales, *High* scores occur ***three*** times as often as *Low* scores. That provides very strong evidence that the scales are top end truncated.

Suppose that for each of these dimensions Cloustons had included a fifth category – *Very High*. The worst that could have happened is that there would have been no cases scoring in that category, i.e. the same as for their *Negligible* category. Or there would have been a number of cases to which *Very High* scores turned out to be applicable.

Of course doing so would also mean having to reconsider category definitions and reformulating them in such a way that there was a meaningful distinction between *High* and *Very High* for each of those scales. Not only would that have likely moved some cases from *High* to *Very High* but also some from *Moderate* to *High*.

Clouston's choice of scale categories for the four point scales logically implies scale truncation. The data reported by Cloustons confirms it***. All of those dimensions have scales with inappropriate truncation at the upper end, skewing the scores reported in a way beneficial to the developer.***

Quantum of View[7]

Assessment:

- **Conceptually flawed scale based on applying ideas of urban views to rural locations**
- **Category definitions reflect constructor bias**
- **Application of categories is subjective**
- **Inappropriate truncation created at upper end of scale**

Clouston's description of this dimension is:

> "The quantum of view relates to the openness of the view and the receptor's angle of view to the scene. A development located in the direct line of sight has a higher impact than if it were located obliquely at the edge of the view. Whether the view of the Project is filtered by vegetation or built form also affects the impact, as does the nature of the view (panoramic, restricted etc.). A small element within a panoramic view has less impact than the same element within a restricted or narrow view."

[6] See pattern of scores in *Visual Impact Summary – Representative Viewpoints* table, *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 102, Table 4.8.

[7] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60, Table 4.6.

The four categories are:

A direct view of the Project or its presence (sometimes in a very narrow or highly framed view), where the Project occupies the greater proportion of the view frame.	High
A direct view of the Project or its presence in a broader view where the Project occupies a moderate proportion of the view frame.	Moderate
An oblique, highly filtered or largely obscured view of the Project or a view where the Project occupies a very small section of the view frame.	Low
Only an insignificant part of the Project is discernible.	Negligible

It can be seen the four categories are all subjectively defined and all require subjectivity to apply the categories. There is no clear distinction between each category and the adjacent one(s).

The definition and categorisation uses a distinction between a "direct view" and an "oblique view" without defining either but seeming to imply something about the orientation of the turbines to the residence. If so, that implies some primacy to one aspect of the residence and its curtilage. In reality, the main areas of use and activity, and thus exposure to the wind farm, often have no relationship to what in a formal sense is the "front" of the residence.

If one were talking about the view of Sydney Harbour from an apartment window, the "direct" / "oblique" comparison may have some relevance. However, that is not what is being evaluated for Jupiter and betrays Clouston's urban blinkers.

People in rural residential localities spend little time looking out their windows or "front" door, but a lot of time outside working and enjoying their properties. Even if one accepts a Departmental imposition of restricting VI to the curtilage, Cloustons has no idea in which direction residents will most often be looking. In a great many instances it will be from what Cloustons would imagine is the back of the house, or the sides, or from a work shed.

For rural properties the whole concept of "direct" or "oblique" is a fiction whose sole purpose is to allow the VI consultant to deem where they would like to imagine the resident will commonly be looking, so the consultant can minimise the assessed magnitude of impact.

If *Quantum of View* is restricted to the extent of a view that is occupied by wind turbines then there is some limited justifiable intellectual basis for the scale. However, we are still left with the fact that there is no clear cut distinction between adjacent categories and thus in each instance the choice is subjective.

Cloustons apparently understands little about what constitutes disruptive elements in a rural landscape for normal viewers of those landscapes.

It does not take a large stain to spoil the appearance of a dress, or a suit, or a shirt. Because a single stain is such a breach of the visual integrity of the object, it draws the eye and conveys a sense of visual disruption out of proportion to the size of the stain relative to the object as a whole.

Industrial structures in a rural landscape, where people expect and want to see only natural and rural views, are the equivalent of a stain on a dress or suit. Only a small amount of the visual expanse needs to be touched in order to ruin the whole view.

Thus what constitutes a *High Quantum of View* for those who love a particular landscape is almost certainly very different to what constitutes a *High* value for those who live in urban and industrial environments.

Distance of View[8]

Assessment:

- **Conceptually objective scale which has been operationalised in a distorted manner**
- **Category cut points show extreme constructor bias which advantages the developer**
- **Category cut points inconsistent with VIA's own table of wind farm visibility**
- **Category cut points dramatically inconsistent with all independent research on wind farm visibility**
- **Inappropriate truncation created at upper end of scale**

Clouston's description of this dimension is:

> "The effect the Project has on the view relating to the distance between the Project and the visual receptor. The distances are from the approximate location of the nearest WTG."

With one caveat noted below, ***application*** of this dimension is objective, since it is simply a matter of distance to the nearest turbine. However, the actual cut points for the categories appear to be highly selective on the part of Cloustons and impossible to rationally justify.

The four categories are

0 – 2 kms	High
2 – 10 kms	Moderate
10 – 15 kms	Low
> 15 kms	Negligible

The bands specified do not correspond in any way with those presented earlier in the VIA for wind farm visibility[9]

0 – 3 kms	Highly prominent
3 – 6 kms	Prominent
6 – 10 kms	Prominent in clear visibility
10 – 20 kms	A small element in the landscape
20 – 30 kms	A minor element in the landscape

As reported elsewhere, the thresholds for *Highly prominent* and *Prominent* are themselves drastically shrunken compared to what has been determined in independent studies by research teams with international reputations[10]. So the VIA thresholds in VIA Table 4.1 were already highly suspect. The cut points for *Distance of View* are even more so.

In kilometres, the progression of distance bands in VIA Table 4.1 are 3, 3, 4, 10 and 10. For the *Distance of View* categories, they are 2, 8, 5, unlimited. That sequence of 2, then 8, then 5 is wholly unlike anything found in independent studies. It appears to have no rationale other

[8] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60, Table 4.6.

[9] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 47, Table 4.1.

[10] Less than half the distance indicated by the BLM Study, Offshore Study, and Stevenson & Griffiths, as described in *On Ignoring Inconvenient Published Research: Objection to the Proposed Jupiter Wind Farm*, Dr Michael Crawford, February 2017.

than to deliberately shrink the first band within which viewpoints may be rated *High* on the *Distance of View* dimension.

In addition, the caveat mentioned earlier is that Table 4.6 has a footnote for *Distance of View* which seems to flag the idea that in practice the VIA authors may ignore the cut points where they choose, in order to allocate lower ratings on this dimension than the cut points would otherwise specify.

Period of View[11]

Assessment:

- **Conceptually confused and consequently unusable scale**
- **Category cut points actually have important gaps**
- **Where categories are specified in terms of time duration, no indication of the application of those durations (e.g. is it per occurrence, per day, per month?)**
- **No justification provided for why different periods warrant one descriptive category rather than another**
- **Inappropriate truncation created at upper end of scale**

Clouston's description of this dimension is:

> "The length of time the visual receptor is exposed to the view. The duration of view affects the impact of the Project on the viewer - the longer the exposure the more detailed the impression of the proposed change in terms of visual impact."

As can be seen in the table below, this scale is a strange combination of supposedly objective (specific duration) and subjective (e.g. "significant part of day") category definitions. The four categories are

Significant part of day: usually residential property	High
1 to 5 minutes	Moderate
1 to 10 seconds	Low
Less than 1 second	Negligible

First it is obvious that the categories are not continuous in their coverage of time. Where does 30 seconds fit in? How about 10 minutes? No scale should have that uncertainty.

Second, it is unclear whether the time references are meant to be per trip, per day, or per year. No scale should have that uncertainty.

The application to residences appears to be subjective. For 11 of 13 private viewpoints evaluated in the VIA, *Period of View* is rated *High*. For the other two (9 & 10, covering 16 residences) *Period of View* is rated *Moderate*. How on earth would Cloustons be able to decide that all the people involved would actually see turbines from their properties for only 1 to 5 minutes (presumably in this case per day)? These are simply arbitrary ratings.

Most of the categories in this dimension are intended for public spaces, in particular for people travelling the roads. Cloustons is conceptually confused about those viewers and consequently has operationalised a confused and inappropriately applied dimension.

[11] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60, Table 4.6.

Road travellers can be separated into two broad categories: those who are essentially passers-by (e.g. travelling from Goulburn to the South Coast for an excursion) and those who are regular travellers in the area, many of whom drive the same roads every day, e.g. to and from school, to and from work, or other local activities.

The visual impact on passers-by is basically irrelevant. During a couple of hours driving from Goulburn to the Coast, they may see turbines for 10 or 15 minutes once or twice a year. Big deal!

It is totally different for regular travellers in the area. In their case, a large part of their travel will be wind farm exposed. They are likely to experience it at least twice a day (going from and returning home), and do so 5 – 7 days a week, every week of the year. The exposure is massively different from that of passers-by. In many instances they will also have exposure from their own residence, so there will be a strong cumulative effect.

The VIA and this scale make absolutely no attempt to come to grips with this. For this purpose the appropriate unit of analysis is not some viewpoint on the road but individual regular travellers in the area.

To do that, requires a travel analysis for residents and workers in the area and the cumulative experience they will have. That is certainly more demanding than simply taking a few snapshots from a couple of selected points on the road. ***It is the only way to determine the actual visual impact on the travellers who matter***, i.e. those who spend a lot of time travelling in the vicinity of the wind farm.

Magnitude of Change[12]

Assessment:

- **A wholly subjective set of categories based on abstract terms**
- **Claims to be "a quantitative assessment" but is not quantitative**
- **Has inconsistent definitions of what warrants a high score on the scale**
- **The dimension is confounded with the dimension *Quantum of View***
- **Inappropriate truncation created at upper end of scale**

Clouston's description of this dimension is:

> "Scale of change is a quantitative assessment of the change in compositional elements of the view. If the proposed development is largely similar in nature and scale to that of existing elements in the vicinity, the scale of change is low. If the development radically changes the nature or composition of the elements in the view, the scale of change is high. Distance from the development would accentuate or moderate the scale and variety of visible elements in the overall view and hence influence this rating."

This is a wholly subjective set of categories based on abstract categories. It therefore can be expected to have inherently low reliability. In addition the categorisation is confounded with "Quantum of View" and "Viewer Sensitivity" (as defined).

[12] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60, Table 4.6.

The four categories are

Elements within the view would greatly dominate existing features in the landscape	High
Elements within the view would be at odds with existing features in the landscape	Moderate
Elements and composition of the view would remain largely unaltered	Low
Project barely discernible	Negligible

While Clouston's description claims the dimension is "a quantitative assessment" there is no quantitative measure involved. So that statement is false. Comparing the definitions for the *High* and *Moderate* categories shows there is no clear distinction between them.

The reference in *High* to "greatly dominate existing features" is confounded with what is required for a *High Quantum of View* ("where the Project occupies the greater proportion of the view frame").

It can be seen that the authors are confused about what constitutes a *High* rating on this dimension. The dimension definition says "If the development radically changes the nature or composition of the elements in the view, the scale of change is high", i.e. it specifically refers to change in the ***nature*** of the view, rather than the geographic scope. Installation of any large scale industrial structure, including wind turbines, in a rural area necessarily radically change the nature of the view.

However, the definitions specifically given for the *High* category is that "Elements within the view would greatly dominate existing features". This is more a statement about physical magnitude than about nature of change. For instance, establishment of a pine plantation across the Jupiter project area would certainly "dominate existing features in the landscape", where it could be seen, but would not amount to anything like the nature of change brought by 173m industrial structures.

So the authors of the dimension present two different definitions of what constitutes a High rating on the scale. Any scale so ill-defined will be fundamentally flawed in its application.

Magnitude

Assessment:

- **A composite scale for which the formulation that generates scores on the scale has not been revealed**
- **Reported results include obvious inconsistencies demonstrating arbitrariness in application**
- **Failure to provide any validation for the formulation of the scale**
- **Inappropriate truncation created at upper end of scale**

Clouston's description of this dimension is:

> "The magnitude of the visual effects of the development within the landscape. A series of quantitative assessments are studied, including distance from development, quantum of view and duration of view and magnitude of change." [13]

[13] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 59.

There are no specific category definitions for this scale, which is wholly computed in some manner from a combination of scores on the scales:

- *Quantum of View*
- *Distance of View*
- *Period of View*
- *Magnitude of Change*

Like those other scales, it is denominated as four categories: *Negligible, Low, Moderate, High*.

Unknown Score Determination and Arbitrary Results

Since scores on this scale are wholly derived from other scales, the formula by which those scores are determined from the subscale scores is absolutely critical.

It cannot be done by any simple arithmetical means such as averaging because the subscales are at best ordinal numbers not interval numbers and *ordinal numbers cannot be validly used in mathematical calculations*.

There may, nonetheless, be a logical formulation to describe how to get from the subscale values to the "aggregated" scale. However, a determination formula is not provided in the EIS in order to explain how the *Magnitude* scores are derived.

There is, however, a table *Visual Impact Summary – Representative Viewpoints*[14] which records all the scale values for 20 viewpoints. There is a column for *Magnitude* and columns for each of the constituent subscales. Perusal of the data in that table shows what appear to be arbitrary determinations of *Magnitude* scores.

Viewpoints 10, 12, 13, and 20 each have *High* scores on two subscales and *Moderate* scores on two subscales. The *Magnitude* score is shown as *High* for v12 but *Moderate* for all the others.

Viewpoint 17 is *Low* on two subscales, *Moderate* on another and *High* on the fourth. The net result is shown as *Moderate* on *Magnitude*. But viewpoint 14 is also *Low* on two subscales, and *Moderate* on one and *High* on another. However, in that case the *Magnitude* is shown as *Low*.

It may be that because there are some differences between these viewpoints as to which subscales have the various scores, the actual result can be explained from a specific formula that is being applied. Unfortunately, since the formula is not described, the result simply appears arbitrary.

However, the problem is actually worse than that as the next examples show.

For viewpoint 3, two of the subscales scores are *High* and two are *Moderate*. The resultant *Magnitude* score is shown as *High*. Viewpoint 4 is also shown as *High* on the ***same*** two subscales (as for v3) and *Moderate* on the ***same*** two subscales as are *Moderate* for v3. Yet the *Magnitude* for v4 is shown as *Moderate*. So we have two viewpoints with ***identical*** subscale scores but ***different*** scores on the resultant scale.

[14] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 102, Table 4.8.

These are not the only instances. The same occurs for viewpoints 5 and 9 which also have ***identical*** subscale scores but ***different*** resultant Magnitude scores.

No reputable academic journal would allow someone to publish results with such arbitrary score determination. Nor would any legitimate profession allow its members to apply to patients or clients advice based on scores with such arbitrariness.

Missing Validation

A somewhat related matter is that when combining subscales to produce scores on some composite scale, it is normal to engage in a process of validation to demonstrate that the method of aggregation does indeed produce a result that is a reasonably accurate representation of the dimension meant to be measured via the resultant scale.

No evidence of validation of the *Magnitude* scale has been reported in the EIS. Without that the scale is simply an arbitrary assembly of pseudo-numbers which has no validity.

Viewer Sensitivity[15]

Assessment:

- **The operationalisation of the scale is inconsistent with the definition**
- **The definition refers to the sensitivity of individual people, while the category definitions are in terms of the characteristics of locations**
- **Operationalisation of the scale is confounded with magnitude of exposure**
- **Distinction between the *High* and *Moderate* categories is subjective**
- **Inappropriate truncation created at upper end of scale**

Clouston's description of this dimension is:

> "Each visual receptor type has an inherent and varied sensitivity to change in the visual scene based on the personal context in which their view is being experienced. This sensitivity has a direct bearing on the perception of visual impact experienced by the receptor and qualifies the quantitative impacts.
>
> Number of viewers also has a bearing on sensitivity. Viewpoints have a varied number of potential receivers depending on whether the viewpoint is public or private, the popularity of the viewing location and its ease of accessibility. Views from public reserves and open space are often given the highest weighting due to the increased number of viewers affected."

The four categories are

Public open space, public reserves, living areas or gardens/balconies of residential properties with direct views of Project.	High
Residential properties with limited views, commercial properties, scenic public roads (eg official tourist routes).	Moderate
Minor roads, service providers.	Low
Vacant lot, uninhabited building, car park.	Negligible

[15] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60, Table 4.6.

Note that the description refers to "visual receptors", i.e. people, each of whom the description allow have personal sensitivity **BUT** that the category descriptors refer to locations without any indication of the sensitivity of the people involved.

Clouston's description of the dimension is itself reasonable but is at variance with the category definitions for *High* and *Moderate* as they apply to people in residential properties. For those two categories, in relation to residences, the difference is whether they have "living areas or gardens/balconies with direct views of Project" (*High*) or "limited views" (*Moderate*).

These distinctions confound sensitivity with magnitude of exposure. If someone is sensitive to pollen, or loud noise, or peanut butter, the level of their ***sensitivity*** is wholly unrelated to whether any of the noxious matter is currently available in their vicinity. The ***impact*** of the material is a combination of the ***amount*** to which exposed ***and*** their ***sensitivity*** to that material.

The conceptual description of visual impact in the VIA distinguishes sensitivity from magnitude of exposure but then the operationalisation of a scale for sensitivity includes both sensitivity and magnitude of exposure. It is a seriously faulty operationalisation.

Within the overall schema, the combination of *Quantum of View* (supposed to measure the "openness of the view") and *Distance of View* (physical distance from turbines) are supposed to capture the overall scale of the development visible to particular viewers.

Yet the distinction between *High* and *Moderate* categories for residential property *Viewer Sensitivity* is based essentially on directly including *Quantum of View* in the *Viewer Sensitivity* scale. That means that if a viewer is judged to have a *Moderate Quantum of View*, their *Viewer Sensitivity* is likely to be rated down to *Moderate* as well.

That is simply cheating. The dimensions are defined as separate but then in application the score on one of them is adjusted down based on the score on another. That makes it easier to surreptitiously engineer something less than a *High* overall *Visual Impact* rating.

All residences should have *High Viewer Sensitivity*. The overall *Visual Impact* rating then logically depends on the magnitude of exposure. If there is zero exposure, for whatever reason, then overall visual impact is zero. There is no reason to attempt to make resident *Viewer Sensitivity* itself subject to the separate dimension of *Magnitude* other than an attempt to covertly engineer lower overall impact ratings for more residences.

There is also no rational reason to differently categorise "scenic public roads" (*Moderate*) and minor road (*Low*) except for the intellectually flawed assumption that what matters is that fewer people travel on any instance of the latter than on the former.

However, as we saw with *Period of View*, that issue arises only because the VIA is using specific public locations rather than actual people as the unit of analysis. When the focus is on people (i.e. those whose experience constitutes the actual *Visual Impact*) then most of the locals will be found to travel on both major and minor roads around the wind farm and experience cumulative *Visual Impact* from the combined exposure.

Visual Impact

Assessment:

- **A number of the category descriptors are identical with those of other categories, just described with slightly different wording**
- **In application, the difference between categories, especially adjacent ones, is highly subjective**
- **There is no sense of magnitude associated with the "impact" referred to by each of the categories and therefore no sense of quantum of difference between them**
- **The scale is a rehash of the Magnitude of Change dimension (i.e. it is not a distinct scale)**
- **The scale is conceptually confused and fails to actually reflect Visual Impact as something perceived by viewers**
- **Scale attenuation is achieved through creative labelling which serves the interest of developers rather than parties actually subject to Visual Impact**

So the scale is in no way fit for purpose, unless the purpose is to achieve minimisation of the number of parties counted as subject to a significant degree of Visual Impact.

The VIA uses a scale for visual impact with the following 6 categories[16]

The Project becomes the dominant feature of the scene to which other elements become subordinate, significantly affecting and changing the visual character.	High
The Project is a discernible feature of the scene, altering the character of the view.	Moderate/High
The Project may form a visible and recognisable new element within the overall scene that affects and changes its overall character.	Moderate
Whilst discernible, the Project does not dominate the visual scene and has only slight impacts on the character of the view.	Moderate/Low
The Project constitutes only a minor component of the view, which might be missed by the casual observer or receptor. Awareness of the proposal would not have a marked effect on character of the view.	Low
None or only an insignificant part of the Project is discernible.	Negligible

Confused and Subjective Categories

It is readily apparent that the category definitions are generally not distinct from adjacent categories and, in some instances, from categories further apart.

Consider *Moderate/High*. The first part of the description "is a discernible feature of the scene" is no use as a distinguishing descriptor since being "a discernible feature of the scene" applies to every category except perhaps *Negligible*.

That leaves us with "altering the character of the view". How is that different from "affects and changes its (*i.e the overall scene*) overall character" for *Moderate*?

Likewise there is no meaningful difference in the descriptors for the *Moderate* and *Moderate/Low* categories. One refers to having "only slight impacts on the character of the view" while the other refers to "not having a marked effect on character of the view". These are simply statements that mean the same thing.

[16] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 59, Table 4.5.

Note that the category description for *Moderate/Low* is that the change "not dominate the visual scene" and have "only slight impacts on the character of the view". Since the special characteristics for the *High* category are that the change be "dominant" and "significant", this wording serves to distinguish the *Moderate/Low* category from the *High* one.

But what distinguishes the *Moderate/High* category from the *High* category is that the former, by implication, does not involve a change that is "dominant" or "significant". That then leaves no distinction between the descriptors for *Moderate/High* and *Moderate/Low*.

Thus, many of the category descriptors do not describe different things. That is a huge problem in scale construction when the scales depend on descriptions rather than meaningful numbers.

Even when the descriptors are arguably different, the application in any particular situation is subjective.

For instance, the descriptor for *Moderate/High* "altering the character of the view" equally applies to the *High* category, except that the latter requires the alteration to be the "dominant" feature and "significantly" affecting the view. So any distinction depends on the wholly subjective judgement as to whether the alteration in character is "significant" or not.

Rehash of Magnitude of Change Dimension

The description of the *Magnitude of Change* dimension is

> "Scale of change is a quantitative assessment of the change in compositional elements of the view. If the proposed development is largely similar in nature and scale to that of existing elements in the vicinity, the scale of change is low. If the development radically changes the nature or composition of the elements in the view, the scale of change is high. Distance from the development would accentuate or moderate the scale and variety of visible elements in the overall view and hence influence this rating." [17]

That turns out to be a comprehensive statement of the category descriptions presented for *Visual Impact*. Consider the *High* category descriptors in the two cases.

> ***Magnitude of Change***: Elements within the view would greatly dominate existing features in the landscape
>
> ***Visual Impact***: The Project becomes the dominant feature of the scene to which other elements become subordinate, significantly affecting and changing the visual character.

The descriptor for *Visual Impact* is a little more repetitiously wordy (by definition, if one thing is dominant, others are subordinate; and if you have added something that is now dominant in the landscape it will inevitably have "significantly affected and changed the visual character") but otherwise identical to that for *Magnitude of Change*.

[17] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60, Table 4.6.

Conceptual Confusion about what Constitutes Visual Impact

Visual impact is something experienced by ***people***, not by houses, roads or public spaces. It occurs ***at*** houses, roads and public spaces but since an entity must have vision in order to experience visual impact, it is restricted to humans and animals.

Quite clearly a wind farm will have no visual impact for a blind person. Much of Clouston's plausible "mitigation" strategy is to plant trees, whose purpose is to prevent people from seeing wind turbines. The turbines would still be added to the landscape. They would still be "at odds" with the existing landscape, but on Clouston's theory the visual impact would be removed so long as the trees grow as Cloustons wants and the people involved stand in the right place so they can no longer see the turbines.

Thus the essential element for Visual Impact is that a person see the wind turbines and that they cognitively and emotionally react in some way.

Cloustons gives the occasional glimmer of understanding this, before retreating from the very inconvenient (for their client and thus for them) consequences.

Thus they give this definition of *Visual Amenity*:

> "the measure of the visual quality of a site or area experienced by residents, workers or visitors. It is the collective affect of the visual components which make a site or an area pleasant to be in." [18]

Visual amenity is of course what is reduced by *visual impact*. Since in Clouston's definition *visual amenity* is an affect (i.e. a psychological feeling or emotion), *visual impact* must be a change (either positive or negative) to an affect. But the *Visual Impact* scale makes no reference to "affect" in any form.

In their definition of *Viewer Sensitivity*, one of the two components determining *Visual Impact* according to their formulation, they say:

> "Each visual receptor type has an inherent and varied sensitivity to change in the visual scene based on the personal context in which their view is being experienced. This sensitivity has a direct bearing on the perception of visual impact experienced by the receptor and qualifies the quantitative impacts." [19]

Notice the dimension is ***Viewer*** *Sensitivity* and they refer to "visual receptor(s)", their jargon for what the rest of us call people. They refer to *visual impact* being something perceived by people and differing between individuals.

And in their description of *Period of View* they refer to "the length of time the visual receptor (i.e. person) is exposed to the view" and that "the longer the exposure the more detailed the impression of the proposed change in terms of visual impact".

This essential component of personal experience is also expressed in the *Visual Assessment Bulletin* recently published by the NSW Department of Planning, which says:

---

[18] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 18.
[19] *Jupiter Wind EIS_ Appendix F_ Landscape and Visual Part 1*, p. 60.

> "Generally, the visual impact of a wind energy project will depend upon the characteristics and values of the existing landscape, the extent to which the existing landscape is changed by the project and how these changes are perceived by individuals and the broader community." [20]

So, as with Clouston's formulation, it is given by the combination of the physical changes (nature and magnitude) and how those are perceived by actual people.

This has been detailed at length by what is currently the gold-standard review of VI assessment methodologies, which explicitly states:

> "The scientific literature on landscape perception repeatedly concludes that human perception of the landscape, including visual quality and visual impacts, is a transactional process. That is, perceived visual quality is the result of interaction between the landscape and people. Visual quality is the product of a relationship between the environment and viewers. Understanding both the affected landscape and the affected population of viewers is necessary for determining visual quality and visual impacts. . . . Aesthetic qualities are not intrinsic in the landscape, and beauty is not merely in the eye of the beholder; rather, the landscape and viewers operate within a system to generate perceptions of visual quality. The landscape is but one component of that visual system that also includes people." [21]

Not only is visual quality dependent on the viewers as well as the landscape, but viewers differ in how each responds to a particular landscape. There are also differences between groups of people and importantly:

> "The differences between what professionals value and what the public values is profound." [22]

The NCHRP Report makes clear that individual perceptions of visual quality are not simply abstract perceptions. They depend on the extent and way in which individuals interact with a particular landscape and, by definition, residents in a locality interact with its landscape, over long periods, in ways that transients do not. Thus, for the purpose of visual impact assessment:

> "Existing visual quality is the value placed on the existing landscape by those people who currently have views of the environment." [23]

It is worth paraphrasing those quotes. While there are some features that people commonly find attractive in landscapes (e.g. land-form relief; vegetation, particularly woodland presence; water bodies; apparent naturalism of land use; length or area of view[24]) in any particular situation there is not some objectively definable level of visual quality in the landscape. The actual value comes from the combination of what physically exists and the way individual viewers relate to it.

Consequently *Visual Impact* has to be measured as the experience of the people who are regular users of the affected landscape.

---

[20] Wind Energy: Visual Assessment Bulletin, NSW Department of Planning, December 2016, p. 3.
[21] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, p. 139.
[22] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, p. 139.
[23] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, p. 142.
[24] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, p. 140.

According to the NCHRP Report:

> "Without exception, peer-reviewed literature reviews characterize visual quality as an interaction between viewer and landscape. This characterization contrasts with artistic characterizations of landscape based on assumptions of intrinsic landscape qualities." [25]

Yet it is clear from the category descriptions for Clouston's *Visual Impact* scale that those categories relate to "assumptions of intrinsic landscape qualities" rather than what is experienced by individual viewers. They are, consequently, in direct conflict with all modern peer-reviewed literature relating to visual amenity and visual impact.

In summary, Cloustons suggests *Visual Impact* is determined by what is experienced by individuals; DPE's VI Bulletin concurs; and the NCHRP Report tells us that is supported by all peer-reviewed literature. Despite that, Cloustons then operationalised a scale for *Visual Impact* which totally ignores the change as experienced by those affected.

Consequently, aside from the other deficiencies noted, the scale has no conceptual validity.

Scale Truncation – Manufacturing a Desired Result

Scientifically valid scales need sufficient granularity to adequately express significant diversity within the population of interest. It is therefore important to avoid scale attenuation, i.e. compressing the scale into too few categories or truncating it at one end.

In fact the *Visual Impact* scale imposes truncation by labelling the top category *High*, rather than extending beyond that as is commonly done where scale compression is not encouraged by someone's interest in financial advantage.

The Department and industry-paid VI consultants generally support the idea that only landowners experiencing a "*High*" VI warrant significant compensation. Consistent with that, the Clouston's *Visual Impact* scale has a structure which diminishes the number of assignments to that category. It does so by expanding the scale to six categories (all other scales in the VIA have four) and capping the top one as *High*.

Consider what would have happened if, instead of the categories:

> *Negligible, Low, Moderate/Low, Moderate, Moderate/High, High*

Cloustons had used

> *Negligible, Low, Moderate, High, Very High, Extreme*

The latter categorisation is in no way unusual. For instance, the NCHRP Report refers to a scale for the retention of landscape scenic integrity[26]. That scale had six categories, ranging from "Very High" to "Unacceptably Low". Since high on that scale referred to the retention of the pre-existing landscape, if expressed in terms of adverse impact on the landscape, the scale direction would be reversed and would consist of "Very Low", "Low", "Moderate", "High", "Very High", and "Extreme".

---

[25] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, p. 44.

[26] *Evaluation of Methodologies for Visual Impact Assessments, NCHRP Report 741*, Transportation Research Board of the National Academies, Washington DC, 2013, p.23.

The Clouston's *Visual Impact* categories have been labelled in words, not numbers, and those words have commonly understood interpretative implications.

According to the Matt Knight table of wind farm visibility, "WTG (are) likely to dominate the field of view and appear large scale" up to 3 kms (independent research indicates 9 kms for 173 m turbines but, for the sake of argument, we will stick with the Matt Knight range).

Consider two properties occupied by people with similar landscape values and with an unimpeded view of 25 173m turbines. In one case the nearest turbine is 1 km away. In the other, it is 2.9 kms away. Is Cloustons really going to claim the *Visual Impact* is identical in these two cases?

Certainly in both cases it would be shocking to the occupants, but putting them both in the same category mislabels one or both. When the most extreme category is labelled *High*, then inevitably the very worst affected cases must end up there. Then other cases that are badly affected but not to the same extent get pushed into lower categories and if Departmental action is tied to whether or not a residence is in the *High* category, then those depressed below that level through this psychological trick of linguistic anchoring are unfairly treated.

We should be completely clear about the matter. The purpose of VI Assessment is not some interesting academic exercise. It is purely about determining whether some parties would be adversely affected sufficiently by the proposed project that either the project should be blocked, modified or some form of compensation required.

So the structure and interpretative labelling of scales used for VI Assessment has a financial impact for the parties involved and is not some arm's length, dispassionate undertaking by disinterested consultants.

Whether intentionally or otherwise, Cloustons has chosen descriptive names for the *Visual Impact* categories whose effect is to truncate the scale and thereby render it invalid.