The Myths and Realities of Psychometric Testing

Agenda

- **A values approach to product-based consulting.**

- **Some key questions around psychometric testing.**

- **The truth about psychometric assessment.**

- **The real state of the psychometric assessment industry.**

- **The 16 myths of psychometric testing that you need to know.**

Our Values

- **Practical Application**

- **Results Focus**

- **Integrity**

- **Long-Term Relationships**

- **Success**

Our Business

- **Assessment Tools**

- **Survey Solutions**

- **Training**

- **Consulting Services**

"Blending contemporary research with practical application"

Issues faced by OPRA

- The market operated at a transactional level.

- Consulting organisations were the keepers of psychometric data = high cost, limited access.

- Psychometric testing was being promoted as a specialist activity, leading to high costs and reduced uptake.

- Limited access to robust technical data to support the claims by test promoters.

- Test users weren't encouraged to do their own sourcing of independent information, and education was limited.

- The start of the Internet and the growth of commercialisation.

Psychometric Assessments

Q. How useful are psychometric assessments?

Q. What makes a psychometric assessment robust?

Industry Changes

Industry changes since 1997:

- **Serious commercialisation of psychometric testing companies.**

- **Psychometric testing companies become publicly listed in NZ.**

- **Growth of internet and internet based testing.**

- **Growing number of providers enter the market.**

BUT WHAT HAS CHANGED FOR THE END USER?

Industry Changes

"No technology of which we are aware- computers, telecommunications, televisions, and so on- has shown the kind of ideational stagnation that has characterized the testing industry. Why? Because in other industries, those who do not innovate do not survive. In the testing industry, the opposite appears to be the case. Like Rocky I, Rocky II, Rocky III, and so on, the testing industry provides minor cosmetic successive variants of the same product where only the numbers after the names substantially change. These variants survive because psychologists buy the tests and then loyally defend them (see preceding nine commentaries, this issue). The existing tests and use of tests have value, but they are not the best they can be…".

STERNBERG & WILLIAMS, 1998

Myth 1

Myth: Good quality psychometric tools must be expensive.

Reality:

- **Psychometric tools are expensive due to lack of competition.**

- **Who set the benchmark?**

- **Why do consumers accept this benchmark?**

- **Cost as a measure of quality?**

Myth 2

Myth: Being a dominant test provider means they are the best.

Reality:

- **Dominance comes from being first.**
- **Not a guarantee of quality.**

Myth 3

Myth: Predictive validity studies demonstrate the usefulness of a tool for all organisations.

Reality:

- **Predictive validity studies demonstrate the usefulness of a tool for the particular organisation in which they were conducted.**

- **Important to evaluate against own criteria.**

- **Meta-analysis demonstrate general usefulness.**

Questions on Validity

- Where is the evidence that tests vary greatly in their predictive power in similar settings measuring similar constructs?

- Where is the evidence that more new measurement methodologies provide any gain in predictive validity?

- Why do we persist using single scale linear correlation as a supporting evidence in tests and then ignore the results in practice?

- Where is the research that combines measurements to show true incremental gain?

- Where is the research to show competencies can be measured like traits?

The Extent of Our Knowledge

- **Smarter people are more likely to perform well (Schmidt & Hunter, 2004).**

- **Those who work hard, are goal oriented, and have an eye for detail tend to perform well (Barrick, Mount & Judge, 2001).**
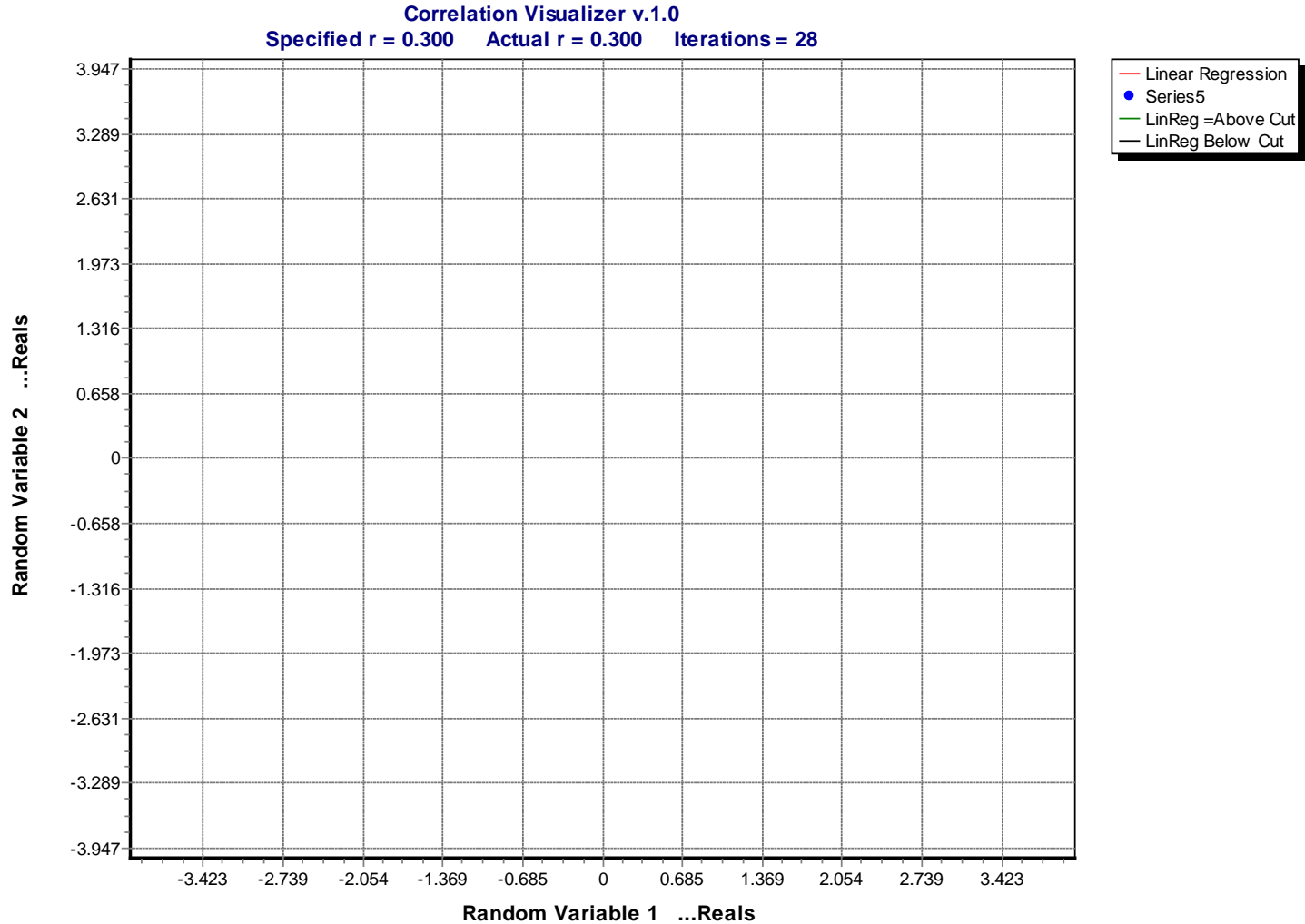
Well that is incredible!!!

Ideal Profiler

Ideal Profiler Math
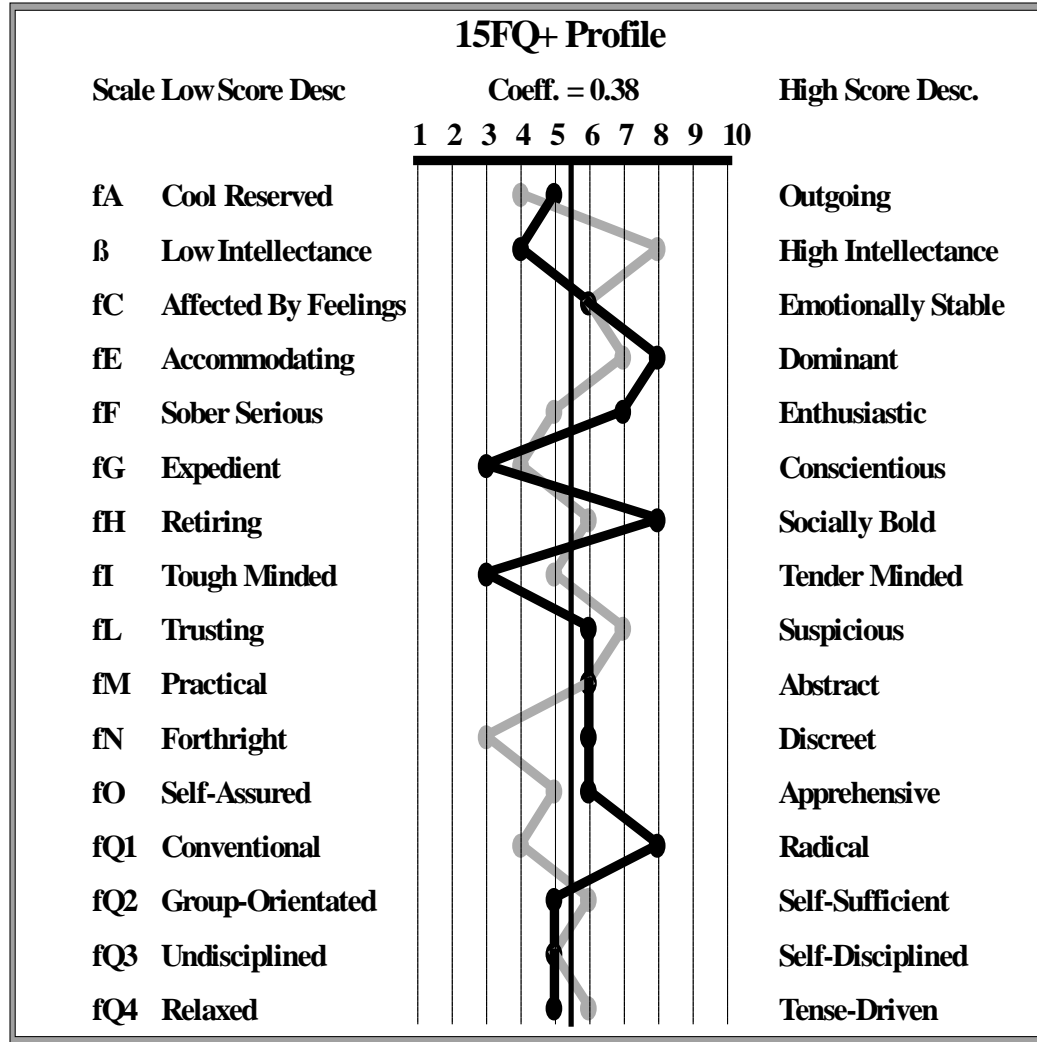
The Kernel Distance Profile Similarity coefficient used within GeneSys (Barrett, 2005)

$$KDC\_score = \frac{\sum_{i=1}^{N}\left[\frac{1}{s\sqrt{2\pi}}e^{-\left[\frac{(p_i-t_i)^2}{2s^2}\right]}\right]\cdot\left(100\cdot\left(s\cdot\sqrt{2\pi}\right)\right)}{N}$$
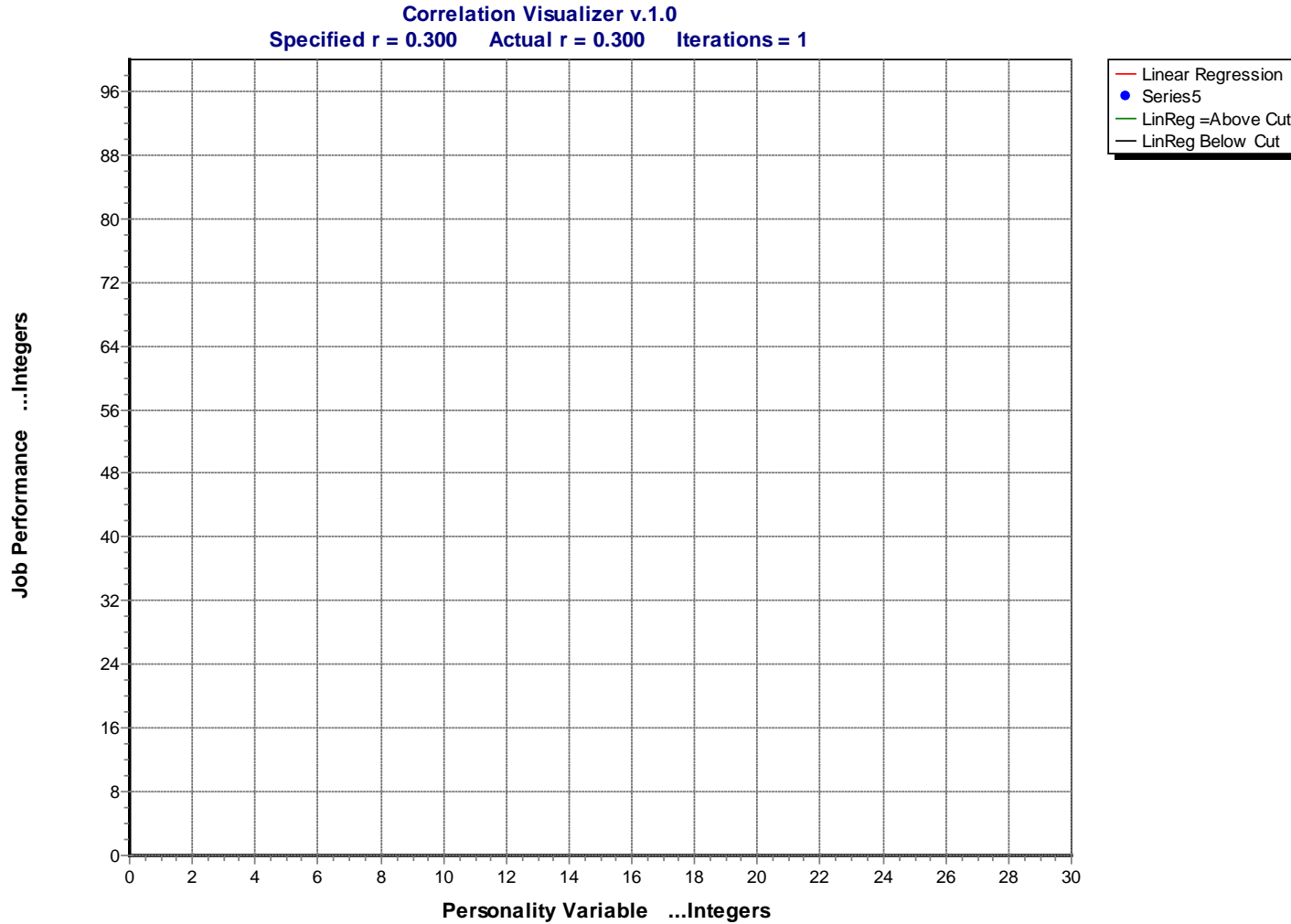
Where: *s* = the standard deviation (smoother) parameter
**   *p* = the comparison score for an attribute**
**   *t* = the target score for an attribute**
**   *N* = the number of attributes in the target profile**

Correlation r = 0.3 (Normal bivariate)

Correlation Visualizer v.1.0
Specified r = 0.300   Actual r = 0.300   Iterations = 28

"Blending Research with Application"

Correlation r = 0.3 (Test Scores)

Correlation Visualizer v.1.0
Specified r = 0.300   Actual r = 0.300   Iterations = 1

Legend:
- Linear Regression
- Series5
- LinReg =Above Cut
- LinReg Below Cut

Y-axis: **Job Performance   ...Integers**
X-axis: **Personality Variable   ...Integers**

"Blending Research with Application"

Correlations

Amount of variance accounted for by different correlations of psychometric tools

Correlation (r)	Variance
0.1	1%
0.2	4%
0.3	9%
0.4	16%
0.5	25%
0.6	36%
0.7	49%
0.8	64%
0.9	81%
1.0	100%

Myth 4
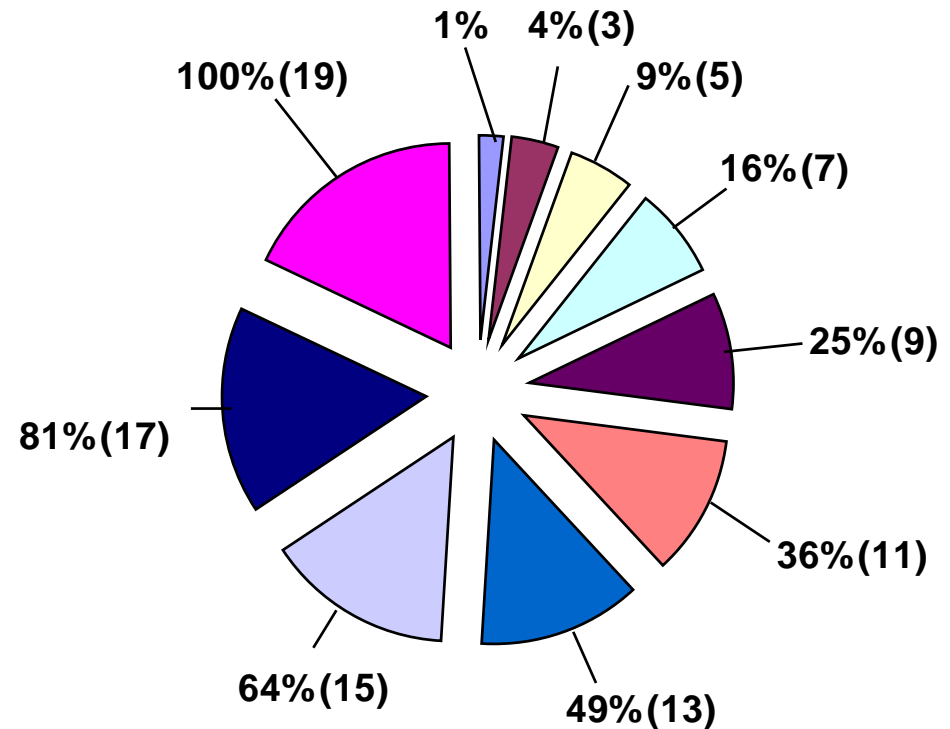
Myth: Psychometric testing is a transactional service.

Reality:

- **Psychometric testing is a strategic initiative.**
- **Use collected data to model performance.**

Myth: Psychometric testing is the domain of psychologists whose main interest is in furthering the discipline of psychology.

Reality:

- **Psychometric testing is too often the domain of non-psychologist business people interested solely in making a profit.**

- **Test producers should provide information on the psychometric properties of their tests.**

Myth 6

Myth: Good psychometric tests are made by psychologists.

Reality:

- **Good psychometric tests are made by psychometricians.**

Myth 7

Myth: Putting a test on the Internet is difficult and that is why few people offer it.

Reality:

- **Putting a test on the Internet is easy.**
- **People restrict its use for ethical and reliability reasons.**
- **Internet testing is convenient and has popular appeal.**
- **Psychometric test interpretation relies on standardised testing conditions.**

Myth: People have a 'work personality'.

Reality:

- **Work-personality negates the whole concept of personality.**

- **Work is not itself a single construct.**

- **Work personality creates a situational stability to behaviour that does not exist.**

- **Meta-analysis suggest that a test which provides a good measure of the 'big five' personality traits does predict performance regardless of the setting.**

Myth: It doesn't matter how a tool is constructed.

Reality:

- **The effectiveness of a tool depends primarily on how well it has been constructed.**

- **Factor analysis is generally regarded as the most robust statistical process for ensuring the rigor of a psychometric tool.**

Test Construction

- **An alternative method of test construction to Classical Test Theory is Item Response Theory (IRT).**

- **IRT allows us to investigate questionnaires at an item level based on two properties:**

  - Difficulty: the difficulty level of an item.

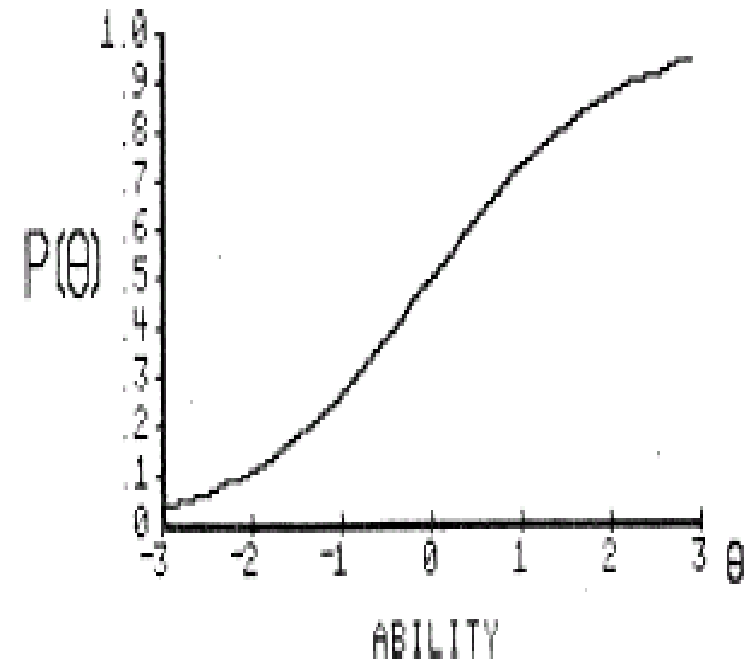 - Discrimination: an items ability to discriminate between individual test takers varying abilities.

Figure 1. Item Characteristic Curve

Item Response Theory (IRT)

Advantages

- Assessment of measurement equivalence across groups – determining item bias.

- IRT also allows theoretical justification for equating scores from one test to another e.g. GRT1 Verbal score and GRT2 Verbal score.

- Ability to deliver computer adaptive or tailored testing.

- Increasing measurement precision. By measuring assessments on an item level, the specific contribution of specific items can be assessed as they are added and removed from an assessment.

Myth 10

Myth: Research material should only be given to current test users.

Reality:

- Research material should be made available to everyone to both further the worldwide knowledge base and to allow for informed consumer decisions.

Myth 11

Myth: Ipsative tests (forced choice) are good for making selection decisions.

Reality:

- **Ipsative tests have been criticised by psychometricians as being inappropriate for use in selection.**

- **Cannot be normed.**

- **Results cannot be factor analysed.**

- **Subject to input response biases.**

"In sum, the standards required of tests used for employee selection are quite strict with regards to validity and reliability of the selection instruments. As such, the limitation inherent with ipsative measures pose too great to threat to the validity of the selection tools to make it a useful instrument for selection on a trait-by-trait basis."

MEADE (2004)

Myth 12

Myth: Psychometric tools should only be interpreted by a psychologist.

Reality:

- Psychometric tools can be interpreted by anyone who has had the relevant training.

- Psychometric tools are built to be interpreted in a standardised way.

Myth 13

Myth: If tests are objective anyone can interpret them and therefore training is unnecessary.

Reality:

- You need to be trained to make psychometric tools really useful.

- Training is necessary for these reasons: ethical, standardisation, legal, utility, psychological and HR guidelines.

Myth: The size of a norm group is often promoted as the most important norm criteria.

Reality:

- The relevance and distribution is often the most important norm criteria.

- We must compare like with like.

- Goodwin & Leech (2006). Understanding correlation: Factors that affect the size of r.

Myth: There needs to be an additional charge for reporting.

Reality:

- **You need only be charged once for testing.**

- **Test producers look at various means of extracting additional money from client organisations.**

- **Once the test data is inputted into a scoring system, no additional time is required for a report to be automatically generated.**

Myth: Predictive validity is the only research that counts.

Reality:

- **Tests construction determines the cross validation of tests.**

- **A core issue in the NZ context is bias and this is a test construction not predictive validity issue.**

- **Predictive validity is ultimately limited by the robustness of the development. Therefore, test development is of more elementary importance than predictive validity.**

References

- **Meade, A. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection.** *Journal of Occupational and Organizational Psychology, 77,* **531-552**

- **Sternberg, R. J., & Williams, W. (1998). Your proved our point better than we did: A reply to our critics.** *American Psychologist, 53 (5),* **576-577**

- **Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r.** *Journal of Experimental Education, 74 (3),* **251-266**

References

- **Schmidt, F. L., & Hunter, J. (2004). General Mental Ability and the World of Work: Occupational Attainment and Job Performance. *Journal of Personality and Social Psychology, 86 (1),* 162-173**

- **Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of a new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9 (1-2),* 9-30**

"Blending Research with Application"

References

- **Barrett, P.T. (2005). Person-Target Profiling. In A. Beauducel, B. Biehl, M. Bosnjak, W. Conrad, G. Schönberger, & D. Wagener (Eds.) *Multivariate Research Strategies: a Festschrift for Werner Wittman*. Chapter 4, pp 63-118. Aachen: Shaker-Verlag**

Questions/Comments?