Advanced Projects R&D Ltd.

Single Item Assessment of Psychological Attributes: Graphical Profiler, Dynamic Analog, and Visual Analog Scales

Paul Barrett, PhD

www.pbarrett.net
paul@pbarrett.net

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

April 20th 2011

UC UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Eliminate classical and modern psychometrics?

The facts about measurement imparted by Michell (1997, 2004, 2008, and 2011) have indicated that classical and modern psychometric test theory and methods may be unnecessarily restrictive for creating useful, robust, psychological assessment.

One sentence in his most recent article explains why: "*There is no evidence that the attributes that psychometricians aspire to measure (such as abilities, attitudes and personality traits) are quantitative.*"

What changes?

What matters now when designing/deploying assessments is establishing/evaluating:

1. Reproducibility as retest reliability
2. Accuracy as theory-mediated predictive accuracy
3. Validity as "*what you think is being assessed is actually being assessed*"
4. Bias as unfair discrimination

Why single-item assessment?

- Simple response interface.
- Efficient assessment.
- Fast completion times.
- Graphically interesting/novel for test-takers.
- Interesting response options and person-target profile possibilities.
- Straightforward validity tools and procedures

  (Use Gigerenzer's "toolbox" rather than highly restrictive psychometric and latent variable models).

Gigerenzer, G. (2004) Mindless Statistics. *The Journal of Socio-Economics*, 33, 587-606.

1 POP Questionnaire – Barrett and Paltiel, 1996

Asked a simple question ..

Is it possible to replace 8 items in a scale with a single item, and retain "good enough" comparative assessment?

We were just curious as to what principles might be required to do this, and whether or not it could actually be done.

- We reduced the 30 scales of the OPQ Concept 5.2, 240-item, questionnaire to just one item per scale.

- Both the OPQ concept 5.2 and the 30-item POP questionnaire were administered concurrently to 420 managerial applicants (359 male; 61 female) as part of a corporate selection exercise.

The 11-point response scale used for the POP items

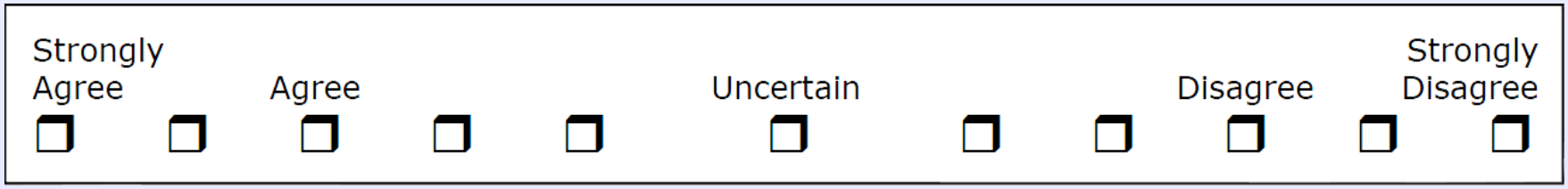
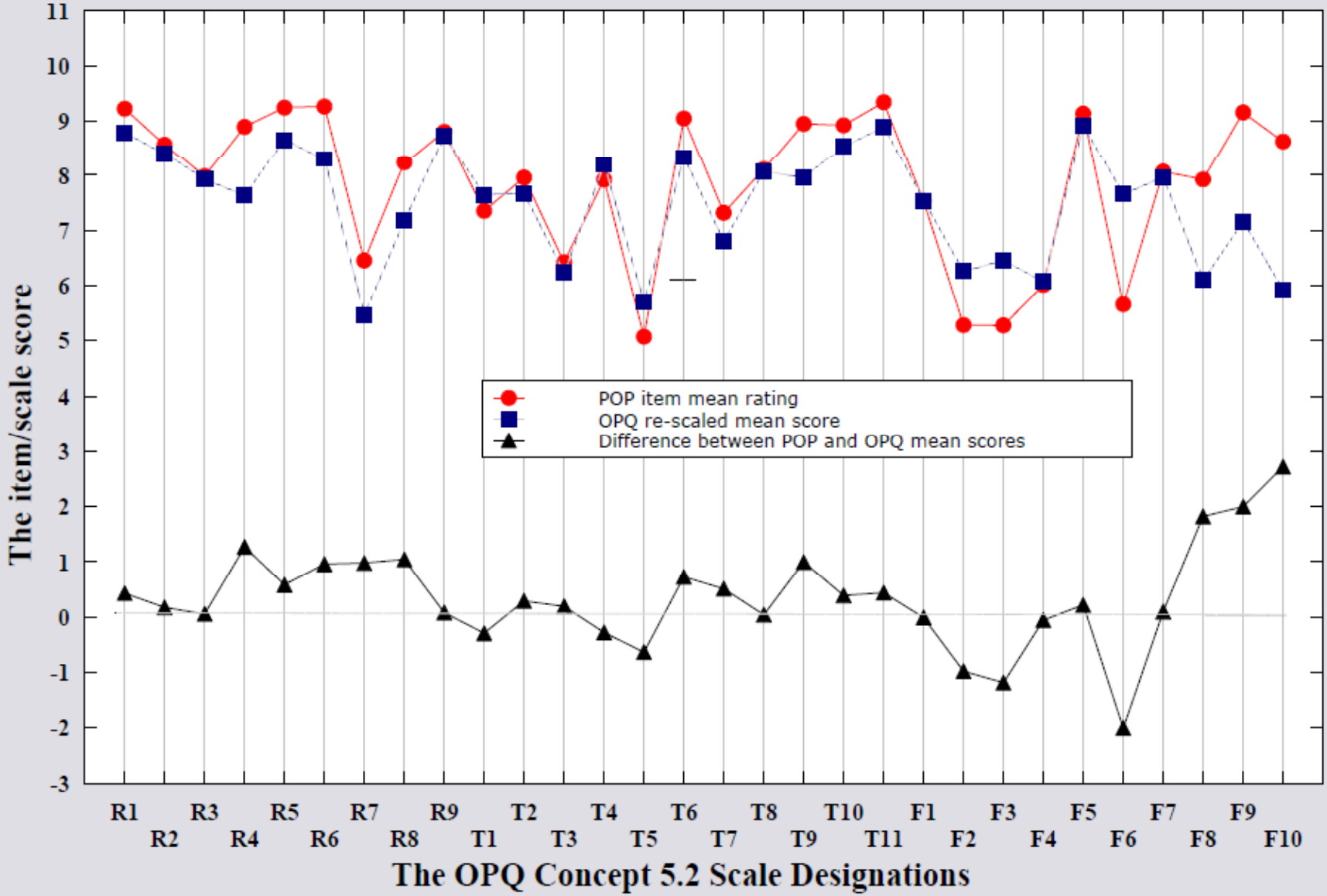


Figure 2: Comparison between the POP item score and corresponding OPQ scale

(The OPQ scale scores have been re-scaled into the 1-11 range of the POP items)



1 POP Questionnaire – Barrett and Paltiel, 1996

More details:

http://www.pbarrett.net/publications.html download #58

Barrett, P.T. and Paltiel, L. (1996). Can a single item replace an entire scale? POP vs the OPQ 5.2. *Selection and Development Review*, 12, 6, 1-4

http://www.pbarrett.net/presentations.html#POP

The POP questionnaire - single item psychometrics and 16PF FormA vs 16PF5.

2 The Birth of the Graphical Profiler
Mariner7.com (Carter Holt Harvey), 2001-2002

- Commercial Application – Web Administered
- 2-dimensional assessment
    - Semantic-Opposite-Pair Work Preference Rating
    - How you like splitting your time between them during a working day.

- Person-Target candidate-job Profiling Application – using 2-dimensional profile matching.

2 The Innovation within the Graphical Profiler

- *Algorithmic Profiling* (invented the Kernel Smoothed Distance coefficient and weighted distance 2-D profile matching).
- We removed construct bipolarity.
- Simultaneous 2-Dimensional assessment.
- The candidates see what is to be said about them as they complete the test.
- Spatial positioning of attributes by the candidate in a 2D space as a response option.
- Mapped preferences into ONET 4.1 job categories

Reproducibility and Stability

Two small samples of data to date:

Adult Working Volunteers *(N=**61** 3-month long term sample and **10** x 5-day individuals)*

Auckland University undergraduates *(N=**25** x **5**-day short-term retest, **23** x 1-month individuals)*

SIOPSA Keynote: 2003

	5 days	1 month	3 months
N	35 29	23 17	61 47
Pearson r	0.65 0.85	0.35 0.77	0.53 0.83
ICC (intraclass)	0.64 0.84	0.34 0.76	0.52 0.82
MAD	11.18 7.34	15.0 8.26	13.26 7.85

*MAD = Mean Absolute Deviation (0-100 range)
Figures in RED are for "clipped" data

SIOPSA Keynote: 2003

2 The Birth of the Graphical Profiler
Mariner7.com (Carter Holt Harvey), 2001-2002

More details:

http://www.pbarrett.net/presentations/Graphical_Profiler_Assessment_NZIO_2002.pdf
Graphical Profiler Assessment? (2002)

http://www.pbarrett.net/presentations/NZIO.pdf
Single item psychometrics, can it be done? (2003)

http://www.pbarrett.net/presentations/SIOPSA_keynote_2003.pdf
Psychological Assessment and Data Utility: it's time to innovate.

3 The 1-dimensional Personality Profiler Barrett and Nina Ebberling - 2003

Asked another simple question ..

Is it possible to assess an individual's personality" as one might with a conventional questionnaire that assesses traits … without using such a questionnaire?

Nina completed a research study as part of her I/O master's – co-funded by Mariner7 and a NZ government enterprise scholarship.

The steps we followed:

1. Construct a conventional personality trait questionnaire with known item and scale properties, and high internal-consistency (> 0.75).

2. Using the free 5-factor personality model item bank at the International Personality Item Pool http://ipip.ori.org/ipip/

Specifically using 10 facets taken from the AB5C 45-facet personality questionnaire as “typical” personality test scales, spanning 106 questionnaire items in total …
http://ipip.ori.org/ipip/newAB5CTable.htm

3. Extract the meaning of all the items in a scale, and compose a single rating statement that seems to best encompass the meaning not only of the scale name, but that embodied within the items.

4. Use a dynamic, graphical, computer-based technology. Create an assessment where an individual chooses how much a single statement applies to them – on a continuous response scale (no numbers, no Likert points, just anchors).

5. Ask 100 university students (*Canterbury University*) to complete both the questionnaire and graphical profiler assessment and ask them which they prefer to do, and which they felt most accurately represented their “personality”.

3 The 1-dimensional Personality Profiler
Barrett and Nina Ebberling - 2003

Questionnaire Completion Time = 20 minutes.
Graphical Profiler = **5** minutes.

93% of individuals preferred completing the Profiler assessment over the paper and pencil Questionnaire

The Results

	Actual Correlation	Disattenuated
Friendliness	0.71	*0.77*
Leadership	0.54	*0.59*
Talkativeness	0.70	*0.75*
Efficiency	0.63	*0.68*
Purposefulness	0.56	*0.61*
Organization	0.60	*0.65*
Orderliness	0.75	*0.80*
Calmness	0.54	*0.62*
Impulse-Control	**0.28**	*0.32*
Happiness	0.64	*0.69*

Profiler
-vs-
questionnaire
scale-scores

Correlations

N=99 cases

Some more detailed results for the Leadership scale

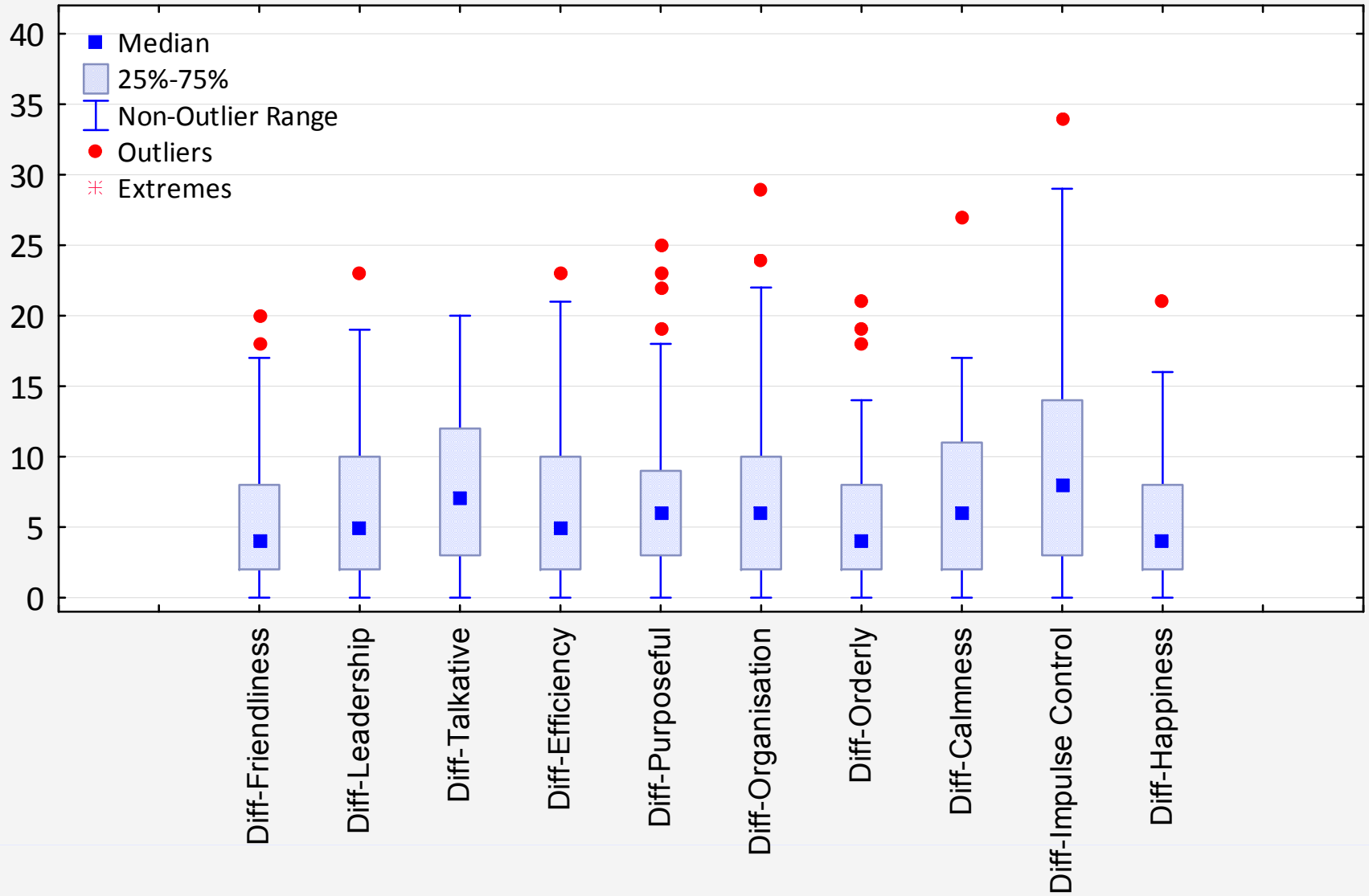
The Gower similarity coefficient of **0.84** indicates that relative to the maximum possible absolute (*unsigned*) discrepancy between them (41), the observations agree to within **84%** of each other's values.

RESCALING IMPLEMENTED

<Vector 1: GP Leadership>
has been rescaled into the metric of
<Vector 2: Q - Leadership>
using the minimum and maximum values specified for Vector 2.

COEFFICIENT	Value
Pearson (Monotonicity)	0.5362
Gower	0.8382
DSE-s	0.7961
KSD-s Sharp (range/6)	0.6371
KSD-s Smooth (range/3)	0.8532
ICC-1	0.4820
ICC-2	0.4946
ICC-3	0.5199
DIAGNOSTICS	
No. of valid cases	99
Original Minimum Vector 1	8.00
Original Maximum Vector 1	100.00
Original Minimum Vector 2	16.00
Original Maximum Vector 2	49.00
Uses Minimum Vector 2	10.00
Uses Maximum Vector 2	50.00
KSD-Sharp - sd	6.6667
KSD Smooth - sd	13.3333
mean absolute discrepancy	6.4707
mean squared discrepancy	66.5459

Absolute Differences between GP and Questionnaire "common-metric" scores



single item rewords? … facet = **Impulse Control**

H381	Keep my emotions under control.
H162	Let others finish what they are saying.
H525	Demand attention.
H802	React intensely.
E98	Talk even when I know I shouldn't.
H523	Often make a fuss.
H1192	Shoot my mouth off.
E19	Am easily excited.
H1176	Blurt out whatever comes into my mind.
H764	Barge in on conversations.
E68	Like to gossip.

NZ Psych Conference – Aug/Sept 2003

single item rewords? ... facet = **Impulse Control**

The Rating Statement:

I rarely express my emotions to others, always careful to think about what I'm going to say before I say it. I really dislike the kind of people who seem unable to properly control their emotions.

3 The 1-dimensional Personality Profiler Barrett and Nina Ebberling - 2003

More details:

http://www.pbarrett.net/presentations/personality_&_gp.pdf
Personality Assessment via Graphical Profiler (2003)

4 The 2-dimensional Graphical Profiler
StaffCV then QuietAgent.com 2005-2011

- Commercial Application – Web Administered
- 2-dimensional assessment
  - Semantic-Opposite-Pair Work Preference Rating
  - How you like splitting your time between them during a working day.
- Person-Target candidate-job Profiling Application – using 2-dimensional profile matching.

Same attributes as the original Mariner7 profiler, but with substantial changes to the interface.

5 The Dynamic Analog Scale – Grice et al, 2011

- Instead of comparatively rating multiple attributes for a single person in the Graphical Profiler, you comparatively rate multiple people on a single attribute in the DAS.

- Comparisons between a Big Five questionnaire and DAS measures showed near identical predictive accuracy of several criterion behavioural outcomes.

- Median same-day retest reliability was **0.83** (Pearson). All attribute retest coefficients were above 0.70.

5 The Dynamic Analog Scale – Grice et al, 2011

More details: james.grice@okstate.edu

Grice, J., Mignogna, M., & Badzinski, S. (2011) The Dynamic Analog Scale: A generic method for single-item measurement. *Personality and Individual Differences*, 50, 4, 481-485.

Brown, E., & Grice, J. (submitted) One is enough: Single-item measurement via the Dynamic Analog Scale. *Journal of Individual Differences*.

6 The 1-dimensional Visual Analog Scale

International Infrastructure Company- Safety Assessment

- Commercial Application – Web Administered (paper and pencil *compromise* version)
- 1-dimensional assessment
- Assessing 14 or 9 Safety-Related attributes/preferences, subdivided into two employee levels, across 60,000+ international workforce.
- Initial rollout 3,000+ UK employees first

Relevant data so far:

- predicting actual "*self-caused*" reportable incidents recorded over a three year period, using **threshold-category attribute scoring** (*a semantically-driven but computational-algorithmic form of scoring a test*) produced a **68%** overall classification accuracy, with balanced False-Positive/False negative rates, working with a 0.07 Base Rate.

- **7min** median completion time (14 attributes).

- Cross-validation on a new incident and trialling sample is underway (June, 2011).

Ongoing – being designed/constructed:

7 The 1-dimensional Visual Analog Scale (p&p) OMR response forms

Indonesian Bank – Competency Profiler for Graduate entrants

8 The 2-dimensional Graphical Profiler

For an International online ATS/job-match provider – Personality Profiler for all current job candidates (6million+)

9 The Dynamic Ability Item

For an International online ATS/job-match provider –

all current job candidates (6million+)

* A more comprehensive series of these new-design tests is being created for low-medium through to very high-ability assessment applications (as within the Asian/Chinese market where many standard ability tests show ceiling effects).

- One single evolving graphics video stimulus, multiple questions.
- Test-taker can replay/stop-start the stimulus as many times as they like.
- Difficulty a function of the evolving complexity of information content on screen.

- The primary assessment goal – to measure "**gestalt reasoning**" – the complete functioning of the cognitive system, including the dynamic interplay of temperament, persistence, motivation, attention, and reasoning ability.

- "Intelligence in life is all about the detection of covariations" (Bob Hogan). That's what this test is designed to assess; not by using static images/items but dynamically evolving patterns.

- The test is not designed to measure "discrete entities/abilities" as do the typical IQ, management, & graduate reasoning tests, but "how well your entire cognitive system copes dynamically with evolving-over-time increasing complexity".

Methodology R&D

- James Grice's Observational Oriented Modeling *(the book is now published April 18th, 2011)*
  http://psychology.okstate.edu/faculty/jgrice/personalitylab/methods.htm

- Custom algorithmic scoring methods allied to cross-validated theory-relevant predictive accuracy.

- Response-Pattern-Based algorithmic Adaptive Test construction (no more IRT).

- Relative Order, then magnitude, among attributes as both target profile and assessment method (*2-D scoring: configural similarity and magnitude-location*)

Key References

- Michell, J. (1997) *Quantitative science and the definition of measurement in Psychology*. British Journal of Psychology, 88, 3, 355-383.
- Michell, J. (2004) *Item Response Models, pathological science, and the shape of error.* Theory and Psychology, 14, 1, 121-129.
- Michell, J. (2008) *Is Psychometrics Pathological Science? Measurement: Interdisciplinary Research & Perspective*, 6, 1, 7-24.
- Michell, J. (2011) *Qualitative research meets the ghost of Pythagoras*. *Theory and Psychology*, 21, 2, 241-259.
- Grice, J. **Idiogrid software** (*for repertory grid and DAS scale construction, administration, and analysis*). Free of charge and downloadable from: http://www.idiogrid.com/