

# CONSTRUCT VALIDITY ISSUES AFFECTING THE UTILITY OF A DEVELOPMENT ASSESSMENT CENTRE –

Vuyo Nako and Antoni Barnard

# AC: A RECIPE FOR SUCCESS?

- AC continues to gain popularity for its use in selection and development
- Scientific and pragmatic usefulness of AC - validity unitarian framework:
  - Content = similarity between content of the measure and job content
  - Construct = relationship between measure and construct it measures
  - Criterion related = relationship between measure and performance indicator
- AC utility research: good predictive validity vs inconsistent construct validity (Arthur et al., 2000; Bowler & Woehr, 2006; Lievens & Conway, 2001; Lance et al., 2004; Thornton and Rupp, 2006)
- Task-based vs dimension-based ACs? (Arthur & Woehr, 2003 vs Lance, 2008; 2012)
- MTMM approach to investigate convergent and divergent validity as indication of construct validity
- Convergent validity = same dimension correlation across exercises
- Divergent/Discriminant validity = no correlation between diff dimension within exercises
- AC Construct validity = strong correlations between same dimension ratings across exercises (convergent) & weak correlations between different dimensions within exercise (divergent)

# WHAT'S THE PROBLEM?

- Question utility of AC in auditing firm
- The AC has been in operation since 2006
- The AC is used for developmental purposes (DC)
- Confirmed utility will impact on continued use or redesign of the DC
- No performance criterion data
- Thus choose to investigate construct validity
- Especially considering that construct validity in the development of ACs is a contentious issue worldwide
- Construct validity is fundamental to fair assessment, especially in a multicultural society like South Africa where assessments were characterised by unfair practices
- Resultantly, the labour and employment legislation requires assessments to be based on job-related competencies, valid and fair

# RESEARCH OBJECTIVE

## *Specific:*

- To extend AC construct validity research in the South African context
- Conduct a literature review on AC as a competency-based assessment in a multicultural context
- Determine if different dimensions measured within exercises are significantly discriminated from one another (divergent validity) in the Development Centre (DC)
- Determine if there are significant correlations between similar dimensions across exercises in the DC (convergent validity)

## *General:*

- To provide the firm with an objective assessment of an individual's ability to function at a leadership level, and as a means of providing guidance to the participant and the business unit on job-related strengths and development areas to be addressed towards partnership admission.

# THE DC BROTH: EXERCISES, DIMENSIONS etc

- 21 competency dimensions
- BARS rated on a 5-pt scale per competency
- **3 Simulation Exercises:**
  - Presentation Exercise
  - In-basket
  - Analysis Exercise – written report
- **2 Interviews:**
  - Competency-Based Panel Interview
  - One-on-one Executive Interview
- **Psychometrics:** (CPP & OPQ32i)

# ASSESSORS

- 3 partners/directors and 3 business unit managers
- One business unit manager was an industrial psychologist
- Received half day training in:
  - meaning of dimensions;
  - behavioural observation; pitfalls of bias and mental errors;
  - behaviour classification according to the BARS;
  - feedback
- Assessors were rotated to observe different participants in different exercises.
- Across exercise rating approach (AER) was combined with the within exercise rating method (WER)

# SAMPLE

- DC data were gathered over four years 2006-2010
- The sample consisted of the results of 138 candidates

Table 3.1: Race and gender breakdown of DC participants (n=138)

	Asian	Black	Coloured	White	Total
Female	8	10	5	26	49
Male	15	20	5	49	89
Total	23	30	10	75	138

# METHODOLOGY

- Quantitative, explorative study
- Variables were relationally analysed with no sample randomisation, manipulation of variables or use of control groups
- Pearson product-moment correlation coefficients were calculated between dimensions within exercises and between same dimensions across exercises.
- Composite dimension ratings were subjected to the principal component analysis (PCA) to cluster dimensions together per exercise and in the DC as a whole.



# CORRELATION RESULTS WITHIN EXs

- Statistical significant correlations between all dimensions within *Presentation Exercise*, *Executive Interview* and *In-basket* = poor discriminant validity
- Of the intercorrelations between 9 dimensions measured in *Panel Interview*, only 3 were statistically non-significant
- Of 21 dimensions measured in *Written Analysis Ex*, 43% (89 of 209) were non-significant 57% of correlations were significant indicating no discrimination
- *Panel interview* & *Written Analysis Ex* some/little discriminant validity was found

# CORRELATION RESULTS ACROSS Exs (convergence)

- 70% of same dimension correlations across ex = non-significant (no convergence)
- 12% moderate correlations (5 from 43)
- 19% statistically significant correlations (8 from 43)
- 7 of the statistically significant correlations was on dimensions between *OPQ* and *Panel interview*
- *Written Communication* correlated between *Exec Interview* and *Presentation Ex*
- For simulation exercises thus no convergence found

# PCA RESULTS

- Composite dimension scores calculated per individual for each exercise
- Inter-correlations reveal a 2 factor structure
- Theoretically the 2 factor structure did not make any sense at all
- Assessors weren't able to differentiate between competency dimensions measured in the overall AC as theoretically no logical pattern structure emerged from the PCA

## FACTOR MATRIX OF COMPOSITE SCORES ON 21 DIMENSIONS

Dimension	Component		h2
	1	2	
Client orientation	0.79		0.77
Applied judgment and insight	0.81		0.81
Business acumen	0.82		0.78
Knowledge- up-to-datedness	0.78		0.77
Practice Leadership	0.79		0.78
Winning business	0.76		0.71
Product and service knowledge	0.54		0.36
Creativity and innovation	0.73	0.52	0.80
Written communication		0.73	0.78
Oral communication		0.80	0.84
Personal impact		0.82	0.84
Personal integrity	0.68	0.61	0.83
Interpersonal skills	0.60	0.52	0.63
Achievement drive	0.72		0.62
Team work	0.73		0.68
Respect for diversity		0.89	0.91
Negotiation	0.60		0.60
Resilience		0.88	0.89
Flexibility			0.78

# LIMITATIONS

- Reactive research – no prior impact on AC design
- Data obtained over period 2006 – 2010
- Standardisation wrt DC method, assessors and administration unconfirmed
- The findings cannot be generalised to other ACs and across other organisations because:
  - Relatively small samples size (n=138)
  - Ratings from a single DC of one organisation and the dimensions reflected the organisation's values and strategies.
  - Cahoon et al., 2012 call for single case validity studies

# TOO MANY COOKS SPOIL THE BROTH

- Number of competencies
- Ambiguous, overlapping competencies
- Assessor cognitive overload despite training efforts
- DC participants' overload and ambiguous expectations
- Feedback 'mush' not useful

# RATING METHODOLOGY

- Exercise rating methodology has significant impact on AC utility
- WER = rating the task/exercise problem
- AER = recommended for improved construct validity and overall AC utility
- Problem WER = MTMM construct validation methodology
- Combination AER + WER?

# RECOMMENDATIONS

- Dimensions:
  - Reduce number acc to AC purpose (5-6 - DC & 6-8 - selection)
  - Conceptually distinct dimensions
  - Transparency of dimensions for participants
- Exercise development
  - SA focus on job relatedness of CBA; TAT equally important
- Assessors:
  - Psychologists & managers (involve psychologist in AC design)
  - Assessors : Assessee ratio - 2:1 (objectivity) and 1:2 (overload)
  - Quality of assessor training and post evaluation of skills
- Rating process
  - Rotate assessors
  - Integrate AER & WER
- SA Research



**QUESTIONS?**

**THANK YOU**