



Libraries and Learning Services

University of Auckland Research Repository, ResearchSpace

Suggested Reference

Brown, G. T. (2016). Socio-demographic differences in the impact of beliefs upon achievement - Non-invariance by sex, age, and ethnicity in conceptions of assessment upon achievement. In *Compass Seminar Series*. University of Auckland.

Copyright

Items in ResearchSpace are protected by copyright, with all rights reserved, unless otherwise indicated. Previously published items are made available in accordance with the copyright policy of the publisher.

For more information, see [General copyright](#).

Socio-demographic differences in the impact of Beliefs upon Achievement: Non- Invariance by Sex, Age, and Ethnicity in Effect of Beliefs Upon Achievement

Compass Seminar, University of Auckland

13 May, 2016

Prof Gavin T L Brown, PhD



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

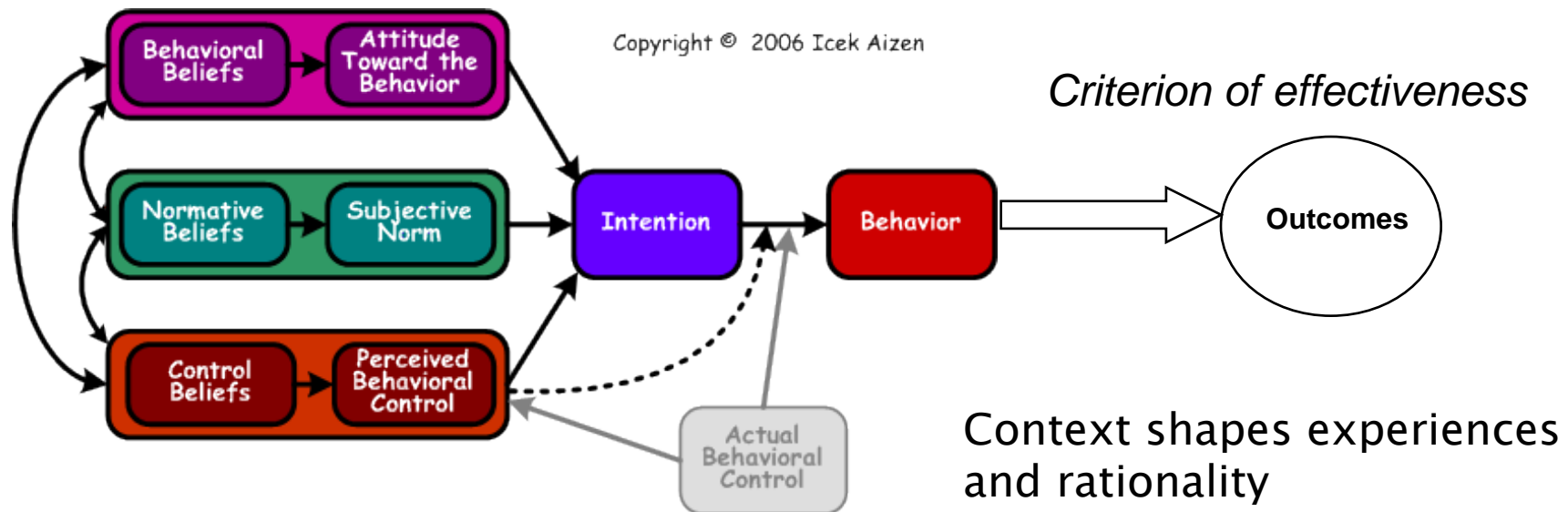
**EDUCATION AND
SOCIAL WORK**

Why Assessment?

- ▶ Assessment is powerful and present in teachers' & students' lives
- ▶ Classroom assessment supposed to lead to increased academic performance
- ▶ Need to understand what assessment means to teachers & students even if that is not what we want them to think...

General Framework

- ▶ Icek Ajzen: Reasoned or Planned Behaviour—Beliefs & Intentions relate to Practices & Outcomes

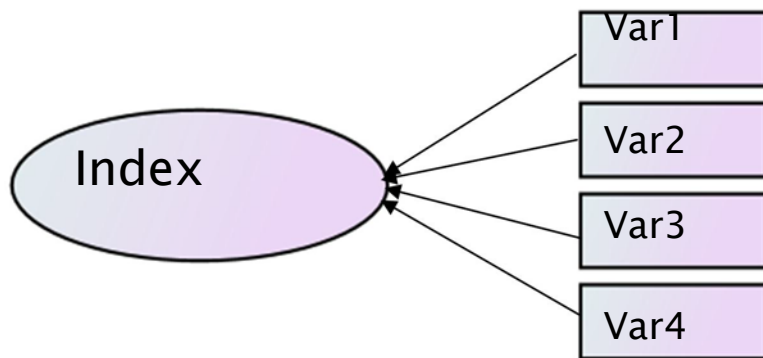


What do we know about what students think of assessment?

- ▶ Assessment is used to identify what they do and don't know (have or haven't learned) so that they and the teacher can do something to improve the situation
- ▶ Assessment is used to hold them accountable for learning outcomes—certification, selection, qualifications, etc.
- ▶ Assessment is a bad thing and has negative consequences for people
- ▶ Assessment activities can be enjoyable or not

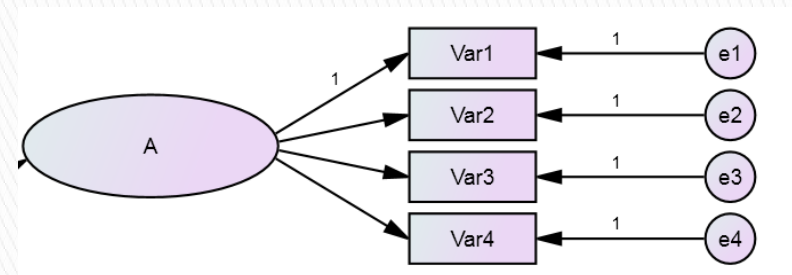
Latent & Manifest variables

- ▶ Indicators create an index
 - SES is a product of manifest vars such as income, education, prestige, etc.



Emergent

- ▶ Latent factors cause responses in a sample of possible indicators



Latent

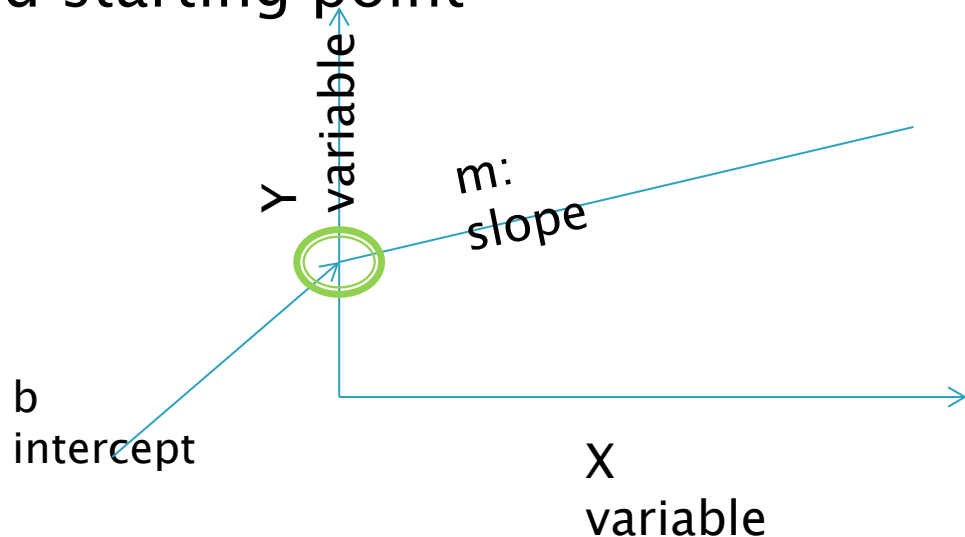
Self-report inventories: Latent trait theory

- ▶ Invisible traits explain responses & behaviours
 - Example: Intelligence (latent) explains how many answers (manifest) you get right on a test



Linear regression components

- ▶ This represents linear regressions
 - Increases in Latent (x) cause increases in Observed (y)
 - Slope is strength of association (usually standardised as β)
 - Intercept is biased starting point



Data vs Model



EDUCATION AND
SOCIAL WORK

Statement

- 29. Assessment fosters students' character.
- 22. Assessment cultivates students' positive attitudes towards life.
- 20. Assessment is used to provoke students to be interested in learning.
- 14. Assessment helps students succeed in authentic/real-world experiences.
- 13. Assessment ensures students pay attention during class.
- 34. Assessment measures students' higher order thinking skills.
- 27. Assessment allows different students to get different instruction.
- 24. Assessment stimulates students to think.
- 49. Assessment forces teachers to teach in a way against their beliefs.
- 31. Assessment interferes with teaching.
- 10. Assessment has little impact on teaching.
- 26. Assessment is an imprecise process.
- 23. Assessment results are filed & ignored.
- 45. Teachers conduct assessments but make little use of the results.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--------------|--------------|-------|--------|--------|-------|--------|
| 0.556 | 0.023 | -0.11 | -0.154 | -0.097 | 0.047 | -0.072 |
| 0.685 | -0.049 | -0.02 | -0.074 | -0.065 | 0.059 | -0.008 |
| 0.591 | 0.04 | 0.084 | -0.066 | -0.059 | -0.02 | 0.048 |
| 0.446 | 0.085 | 0.105 | -0.216 | 0.092 | -0.14 | -0.124 |
| 0.533 | 0.066 | 0.131 | -0.012 | 0.007 | -0.22 | -0.224 |
| 0.509 | -0.167 | 0.007 | -0.03 | -0.176 | -0.11 | 0.077 |
| 0.487 | 0.017 | 0.102 | -0.128 | 0.011 | 0.15 | 0.213 |
| 0.678 | -0.061 | 0.074 | 0.008 | 0.001 | -0.12 | 0.105 |
| -0.083 | 0.458 | -0.03 | 0.121 | -0.071 | -0.19 | 0.106 |
| -0.102 | 0.54 | -0.08 | -0.06 | 0.086 | -0.13 | 0.066 |
| 0.134 | 0.384 | -0.19 | -0.034 | 0.062 | -0.01 | -0.067 |
| -0.004 | 0.629 | 0.034 | 0.008 | 0.021 | 0.057 | 0.094 |
| -0.017 | 0.646 | -0.01 | -0.057 | -0.02 | 0.022 | -0.056 |
| -0.019 | 0.493 | 0.045 | -0.003 | -0.193 | 0.008 | 0.012 |

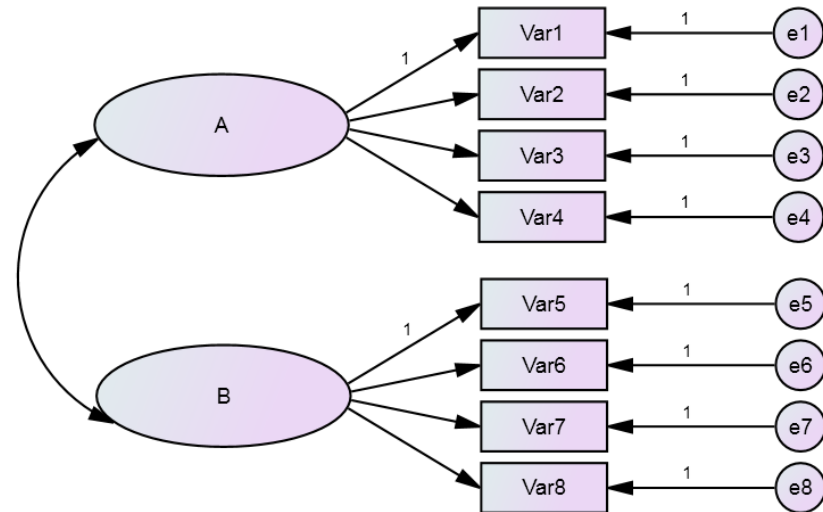
NB. This is the SPSS pattern matrix of regressions

Model turns the OFF values to ZERO
Does that simplification fit the data?

Note. Non-zero values on other factors, but all weak.

Confirmatory factor analysis

- ▶ CFA simplifies the data
 - Forces all non-specified paths to be ZERO
 - If it fits well, we work with latent value scores instead
 - All models are wrong, but hopefully our models are close to the real world



Analysis: Structural Equation Modeling

- ▶ Underlying models assume direction & linearity
- ▶ Model is theory driven
 - everything connects to everything but we want to find what is important and test our notions of that
- ▶ SEM
 - Uses correlations, covariances, means, variances, intercepts, AND residuals SIMULTANEOUSLY
 - Tests for fit of proposed model to data

Preparation: Estimation

- ▶ Maximum likelihood estimation of Pearson product moment correlations,
 - defensible for ordinal rating scales of five or more response categories (Finney & DiStefano, 2006).
 - Additional benefit: handles robustly moderate deviation from univariate normality (Curran, West, & Finch, 1996).
 - Esp. kurtosis up to 11.00
 - excessive kurtosis does not prevent analysis, it does result in reduced power to reject wrong models (Foldnes, Olsson, & Foss, 2012).

Evaluating Results: Fit of model to data

| | <u>Goodness of Fit</u> | | <u>Badness of fit</u> | |
|------------|------------------------|-------------------------|-----------------------|-------|
| Decision | p of χ^2/df | CFI <u>gamma hat</u> | RMSEA | SRMR* |
| Good | >.05 | >.95 | <.05 | <.06 |
| Acceptable | >.05 | >.90 | <.08 | <.08 |
| Marginal | >.01 | .85–.89 | <.10 | |
| Reject | <.01 | <.85 | >.10 | >.08 |

Note.

Report multiple indices but beware.....

CFI punishes **falsely** complex models (i.e., >3 factors)

RMSEA rewards **falsely** complex models with mis-specification

See Fan & Sivo, 2007

*AMOS only generates SRMR if NO missing data;

thus, important to clean up missing values prior to any analysis. Recommend expectation maximization (EM) procedure

Research Questions

- ▶ What do secondary school students think about the nature and purpose of assessment, especially when an inventory is used to capture all four major conceptions?
- ▶ What differences are there by social status of sex, ethnicity, and age?
 - Is the data collection inventory equivalent across groups?
 - Are the predictor paths to achievement identical?

Data collection

- ▶ Data collected Term 1, 2004 as part of the asTTle research and development program
 - Test papers included: 38 reading, 35 writing,
- ▶ SCoA questionnaires were assigned to a subset of asTTle tests which were randomly assigned to students participating in the program.

| Time | English | | | Māori | | |
|---------------|---------|---------|-------------|-------|----------|----------|
| | Reading | Writing | Mathematics | Pānui | Tuhituhi | Pāngarau |
| November 2000 | ☑ | ☑ | | | | |
| March 2001 | | ☑ | | ☑ | | ☑ |
| November 2001 | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |
| June 2002 | | | ☑ | ☑ | ☑ | |
| June 2003 | | | ☑ | | | ☑ |
| March 2004 | ☑ | ☑ | | ☑ | ☑ | |

Participants

- ▶ New Zealand secondary students ($N = 3,504$)
- ▶ **Sex:**
 - 54% female compared to the asTTle population of 49% female.
- ▶ **Grade Level:**
 - 1,462 Year 9; 967 Year 10; 449 Year 11; 591 in Year 12.
 - only the students in Year 12 had participated in the NCEA qualifications assessments.
- ▶ **Ethnicity:**
 - self-reported as New Zealand European/Pakeha ($n = 1,969$), Maori ($n = 474$), Pasifika ($n = 290$), Asian and other ($n = 736$).

Instruments

- ▶ SCOA
 - 2 different forms with total 29 items;
 - 11 items in common
- ▶ asTTle reading tests as outcome measure
 - Taking into account sex, year, & ethnicity

Response Format: Why not Likert?

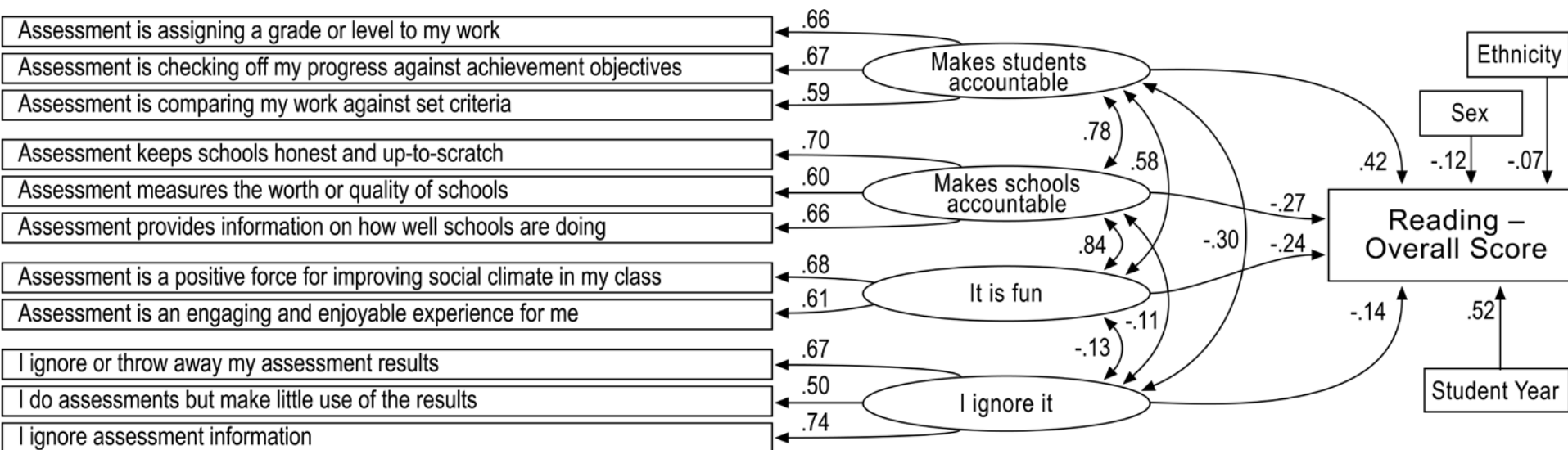
▶ Response Format

- 6 points, Positively–packed, agreement
 - Strongly & mostly **disagree**;
 - Slightly, moderately, mostly, & strongly **agree**
- Increases variance and precision
- Especially useful when participants inclined to agree with something that is socially desirable
- The adjectives used on the agree side are approximately equally spaced

Student Conceptions of Assessment (v2 abridged)

| Code | Conceptions and Items | <u>Item Statistics</u> | | |
|------|---|------------------------|-----------|----------|
| | | <i>M</i> | <i>SD</i> | <i>B</i> |
| | <u>Student Accountability</u> | | | |
| AC4 | Assessment is assigning a grade or level to my work | 4.05 | 1.37 | .66 |
| AC5 | Assessment is checking off my progress against achievement objectives | 3.92 | 1.28 | .67 |
| AC6 | Assessment is comparing my work against set criteria | 3.86 | 1.35 | .59 |
| | <u>School Accountability</u> | | | |
| AC8 | Assessment keeps schools honest and up-to-scratch | 3.37 | 1.36 | .70 |
| AC9 | Assessment measures the worth or quality of schools | 3.24 | 1.45 | .60 |
| AC11 | Assessment provides information on how well schools are doing | 3.52 | 1.40 | .66 |
| | <u>Assessment is Fun</u> | | | |
| IMP8 | Assessment is a positive force for improving social climate in my class | 3.11 | 1.38 | .68 |
| IMP9 | Assessment is an engaging and enjoyable experience for me | 2.77 | 1.39 | .61 |
| | <u>Assessment is Ignored</u> | | | |
| IR8 | I ignore or throw away my assessment results | 2.22 | 1.34 | .67 |
| IR9 | I do assessments but make little use of the results | 2.90 | 1.33 | .50 |
| IR10 | I ignore assessment information | 2.25 | 1.29 | .74 |

Secondary Students' Conceptions of Assessment



SCOA as an inventory

- ▶ Model is consistent with self-regulation and formative assessment.
 - The whole sample exhibited a trend that successful students were those who claimed to use assessment as feedback about their own abilities (Winne, 1995; Zimmerman, 1990).
- ▶ Measurement invariance across sex, year level, and ethnicity?
 - If the SCoA is not equivalent then comparisons are not valid

So what about these demographic variables?

- ▶ If the societal experience of schooling achievement is not universal perhaps this self-regulating model is not identical across social status?
- ▶ Test for INVARIANCE in structural relations of SCoA to achievement
 - means that differences in the key statistical properties of group responses to a test or a questionnaire are so small that they are attributable to chance rather than to group characteristics
 - Indicates the groups have been sampled from the same overall population
 - lack of invariance demonstrates that the research tool (e.g. self-report questionnaire or knowledge test) triggers systematically different responding in each group, making score comparisons invalid since differing response mechanisms underlie group answers

MGCFA invariance testing

- ▶ CFA tests how well a simplified model fits data
- ▶ MG tests how well the same model fits 2 different groups
- ▶ If responses differ only by chance then the inventory works in the same way for both groups; they are drawn from one population
- ▶ If responses differ by more than chance then one set of factor scores cannot be used to compare groups
 - Different models and scores are needed

Testing for Invariance

- ▶ Every CFA produces a set of fit indices; if certain indices change within chance when the equivalence constraint is imposed on the model then that aspect of responding is invariant
 - Change in comparative fit index: $\Delta CFI < .01$ indicates equivalence
 - Change in χ^2 relative to change in df should have $p > .05$
- ▶ Equivalence is needed for
 - Configural (all paths identical)
 - Metric (all regression weights similar)
 - Scalar (all intercepts similar)
 - Each tested sequentially

- ▶ Ethnic groups used in NZ schools at time of study
 - New Zealand European/Pakeha (the dominant, largely British, group who colonized New Zealand in the 1800s),
 - Māori (the aboriginal Polynesian group colonized by the British),
 - Pasifika (immigrant Polynesian groups who settled in New Zealand largely after World War II),
 - Asian/Other (immigrant groups from all other places in the world, with a large proportion being from anywhere between Pakistan and Japan).
- ▶ Followed Stats NZ rules for classifying multiple selection

- ▶ NZ European & Asian generally higher academic performance in reading and mathematics
 - (in the order of 1 to 2 years equivalence) from about age 10 to the end of secondary schooling (Ministry of Education, 2006).
- ▶ Despite sharing similar socio-economic spaces, academic performance, and attitudes as Māori students ('Otunuku & Brown, 2007), it is unlikely that immigrant Pasifika students would have the same structural relations as Māori students, through the largely voluntary processes and consequences associated with migration.
- ▶ Research has suggested Māori students believe teachers are biased against them and if teachers address appropriately these pre-conceptions Māori students attain greater academic achievement (Bishop, Berryman, Tiakiwai, & Richardson, 2003).
- ▶ Hence, we might expect Māori ethnic identity to have significant impact on how students' conceptions of assessment relates to academic performance.

Ethnicity

ETHNICITY

Fit-Statistics for the Model with
constrained measurement weights:

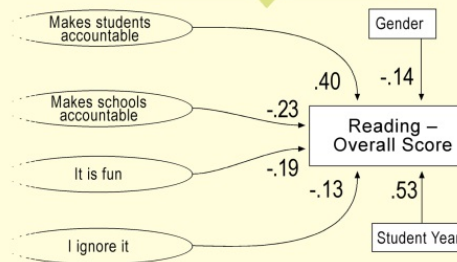
| χ^2 | df | RMSEA | TLI | CFI |
|----------|-----|-------|-----|-----|
| 647.37 | 218 | .026 | .94 | .95 |

Invariance test for the predictors
of achievement

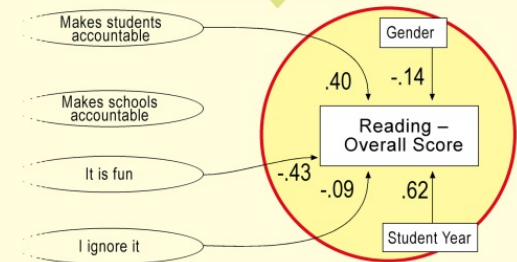
| df | CMIN | p | RFI rho-1 | TLI rho-2 |
|----|-------|------|--------------|--------------|
| 8 | 12.80 | 0.12 | -0.00 | -0.00 |

Note:
Structural Model
Invariant across
ethnicity groups
except for Maori

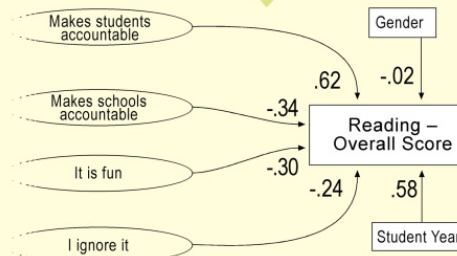
Pakeha n = 1969



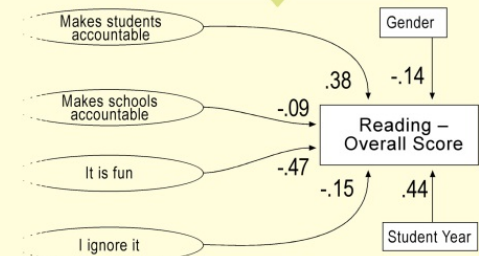
Māori n = 474



Pasifika n = 290



Asian/Other - n = 736



Path values without Māori

| | <u>Student Accountability</u> | | | <u>School Accountability</u> | | | <u>Assessment is Fun</u> | | | <u>Assessment is Ignored</u> | | |
|--------------------------|-------------------------------|-----|-----|------------------------------|------|------|--------------------------|------|------|------------------------------|------|------|
| Item | E | P | A | E | P | A | E | P | A | E | P | A |
| AC4 | .65 | .74 | .67 | – | – | – | – | – | – | – | – | – |
| AC5 | .66 | .69 | .65 | – | – | – | – | – | – | – | – | – |
| AC6 | .61 | .57 | .57 | – | – | – | – | – | – | – | – | – |
| AC8 | – | – | – | .68 | .76 | .74 | – | – | – | – | – | – |
| AC9 | – | – | – | .55 | .60 | .58 | – | – | – | – | – | – |
| AC11 | – | – | – | .74 | .70 | .68 | – | – | – | – | – | – |
| IMP8 | – | – | – | – | – | – | .63 | .69 | .67 | – | – | – |
| IMP9 | – | – | – | – | – | – | .56 | .54 | .62 | – | – | – |
| IR8 | – | – | – | – | – | – | – | – | – | .68 | .52 | .68 |
| IR9 | – | – | – | – | – | – | – | – | – | .55 | .22 | .45 |
| IR10 | – | – | – | – | – | – | – | – | – | .74 | .73 | .74 |
| Relations to achievement | .46 | .72 | .49 | –.30 | –.33 | –.12 | –.24 | –.40 | –.61 | –.12 | –.30 | –.13 |

Invariance without Māori

| Model | χ^2 ($\Delta\chi^2$) | df (Δdf) | p | CFI (ΔCFI) |
|---|-----------------------------|--------------------|-----|----------------------|
| Test 1: Baseline | 362.25 | 114 | .00 | .96 |
| Test 2: Invariant factor loadings | (14.07) | (14) | .45 | (.00) |
| Test 3: Invariant factor loadings , variances and covariances | (75.81) | (36) | .00 | (.01) |
| Test 3a: Invariant factor loadings and covariances | (55.62) | (26) | .00 | (.00) |
| Test 3b: Invariant factor loadings and invariant ignore–students and ignore–schools covariances | (17.89) | (18) | .46 | (.00) |
| Relations to achievement | | | | |
| Test 4: Baseline | 415.21 | 149 | .00 | .96 |
| Test 5: Invariant path coefficients | (14.85) | (8) | .06 | (.00) |
| BUT intercepts NOT equivalent | | | | |

Explaining ethnicity differences

- ▶ The relationship to outcomes seemed to be invariant across three ethnicities.
 - self-regulation constructs are equally important for all students (Marsh et al., 2006; Winne, 1995).
- ▶ However, substantially different relationship for Māori students.
 - There were more Māori than Pasifika students, so small sample size is not an adequate explanation.
 - More likely that the special experience of Māori in the New Zealand educational system is the source of this variability.
 - Although the SCoA captures some important and general aspects of students' conceptions about assessment, further research and development is needed to capture the uniquely Māori conceptions of assessment.

Sex

- ▶ Females consistently outperform males
 - in national secondary school qualifications (NZQA, 2006) and
 - on international assessments (e.g., PIRLS, PISA, TIMSS; Ministry of Education, 2006).
- ▶ Having higher grades might make it easier for girls to agree that assessment is about making students accountable for grades.

Sex

SEX

Fit-Statistics for the Model with
constrained measurement weights:

| χ^2 | df | RMSEA | TLI | CFI |
|----------|-----|-------|-----|-----|
| 743.06 | 136 | .035 | .93 | .94 |

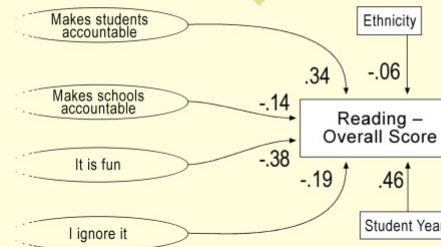
Invariance test for the predictors
of achievement

| df | CMIN | p | RFI rho-1 | TLI rho-2 |
|----|-------|------|--------------|--------------|
| 4 | 11.00 | 0.03 | -0.00 | -0.00 |

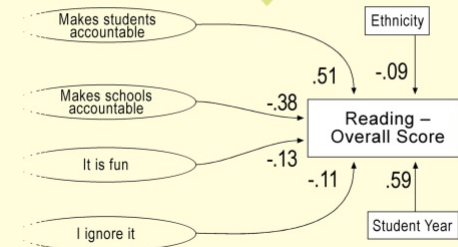
Note:
Structural Model
NOT Invariant

Boys show more effect
around **Accountability**;
Girls show more effect
for **Fun**

Boys n = 1873



Girls n = 1596



Path values by sex

| | Student Accountability | | School Accountability | | Assessment is Fun | | Assessment is Ignored | |
|---------------------------------|------------------------|------------|-----------------------|-------------|-------------------|-------------|-----------------------|-------------|
| Item | Females | Males | Females | Males | Females | Males | Females | Males |
| AC4 | .66 | .66 | – | – | – | – | – | – |
| AC5 | .64 | .70 | – | – | – | – | – | – |
| AC6 | .60 | .58 | – | – | – | – | – | – |
| AC8 | – | – | .70 | .71 | – | – | – | – |
| AC9 | – | – | .61 | .57 | – | – | – | – |
| AC11 | – | – | .66 | .66 | – | – | – | – |
| IMP8 | – | – | – | – | .67 | .68 | – | – |
| IMP9 | – | – | – | – | .63 | .59 | – | – |
| IR8 | – | – | – | – | – | – | .67 | .66 |
| IR9 | – | – | – | – | – | – | .51 | .49 |
| IR10 | – | – | – | – | – | – | .72 | .76 |
| Relations to achievement | .42 | .61 | –.11 | –.48 | –.51 | –.20 | –.19 | –.07 |

Model Invariance

| Model | $\chi^2 (\Delta\chi^2)$ | df (Δ df) | p | CFI (Δ CFI) |
|--|-------------------------|-------------------|-----|---------------------|
| Test 1: Baseline | 385.79 | 76 | .00 | .96 |
| Test 2: Invariant factor loadings | (34.11) | (7) | .40 | (.00) |
| Test 3: Invariant factor loadings, covariances and variances | (32.43) | (17) | .01 | (.00) |
| Test 3a: Invariant factor loadings, covariances and all but one variance | (21.36) | (16) | .17 | (.00) |
| Structural Models | | | | |
| Test 4: Baseline | 445.69 | 106 | .00 | .96 |
| Test 5: Invariant path coefficients | (11.27) | (4) | .02 | (.01) |
| NOT Equivalent at all | | | | |

Explaining sex differences

- ▶ While the social aspect (i.e., assessment is fun) was more indicative for the girls, the school accountability conception was more important for boys' achievement.
 - Among the girls, this may reflect a greater tendency for extravert females to socialize and generate a subsequent counterproductive consequence for academic performance (Chamorro-Premuzic & Furnham, 2005).
 - In contrast, the importance of accountability as a predictor of achievement may be related to the greater tendency of boys to have an external locus of control relative to girls (Chubb, Fertman, & Ross, 1997).

Year



**EDUCATION AND
SOCIAL WORK**

- ▶ In Years 9 and 10, all assessments are school-based and done for formative purposes, with results normally being reported to parents and used by teachers to inform possible changes to learning and instruction.
 - no high-stakes, external examinations or qualifications-related assessments until the third year of schooling (i.e., Year 11).
- ▶ From Year 11 onwards, the nationally moderated National Certificate of Educational Achievement (NCEA), administered by the external New Zealand Qualifications Authority (NZQA).
- ▶ Students earn credits toward the NCEA Level 1 in Year 11 through a combination of in-school assessments and end-of-year external examinations.
- ▶ Thus, it may be that students from Year 11 on would have generally different conceptions of assessment than those in the younger years where assessment is fundamentally low-stakes.

Year

YEAR

Fit-Statistics for the Model with constrained measurement weights:

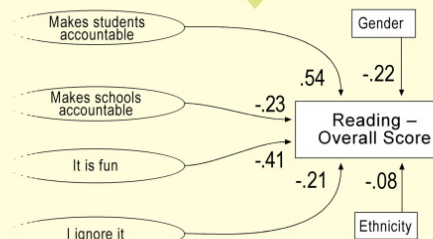
| χ^2 | df | RMSEA | TLI | CFI |
|----------|-----|-------|-----|-----|
| 959.75 | 293 | .026 | .91 | .93 |

Invariance test for the predictors of achievement

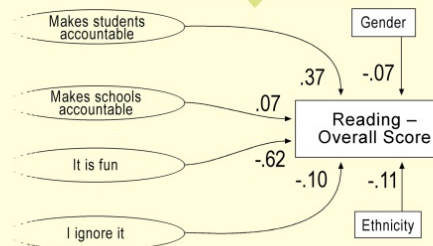
| df | CMIN | p | RFI rho-1 | TLI rho-2 |
|----|-------|------|-----------|-----------|
| 12 | 41.49 | 0.00 | -0.00 | -0.00 |

Note:
Structural Model
NOT invariant
by Year

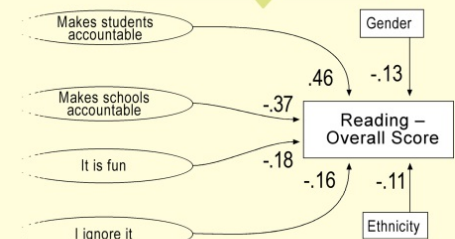
Year 9 $n = 1462$



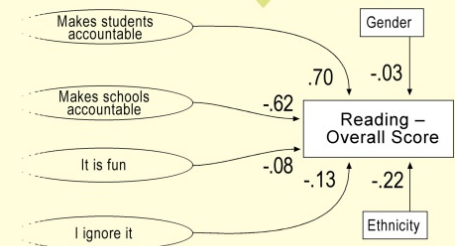
Year 11 $n = 449$



Year 10 $n = 967$



Year 12 $n = 591$



Path values by year



**EDUCATION AND
SOCIAL WORK**

| Item | Student | | | | School | | | | Assessment is | | | | Assessment is | | | |
|--------------------------|----------------|-----|-----|-----|----------------|------|------|------|---------------|------|------|------|---------------|------|------|------|
| | Accountability | | | | Accountability | | | | Fun | | | | Ignored | | | |
| | 9 | 10 | 11 | 12 | 9 | 10 | 11 | 12 | 9 | 10 | 11 | 12 | 9 | 10 | 11 | 12 |
| AC4 | .64 | .65 | .68 | .71 | - | - | - | - | - | - | - | - | - | - | - | - |
| AC5 | .63 | .68 | .70 | .73 | - | - | - | - | - | - | - | - | - | - | - | - |
| AC6 | .60 | .52 | .61 | .66 | - | - | - | - | - | - | - | - | - | - | - | - |
| AC8 | - | - | - | - | .67 | .68 | .77 | .76 | - | - | - | - | - | - | - | - |
| AC9 | - | - | - | - | .59 | .58 | .63 | .58 | - | - | - | - | - | - | - | - |
| AC11 | - | - | - | - | .67 | .58 | .74 | .69 | - | - | - | - | - | - | - | - |
| IMP8 | - | - | - | - | - | - | - | - | .65 | .67 | .72 | .72 | - | - | - | - |
| IMP9 | - | - | - | - | - | - | - | - | .64 | .56 | .63 | .61 | - | - | - | - |
| IR8 | - | - | - | - | - | - | - | - | - | - | - | - | .66 | .59 | .74 | .72 |
| IR9 | - | - | - | - | - | - | - | - | - | - | - | - | .52 | .45 | .47 | .59 |
| IR10 | - | - | - | - | - | - | - | - | - | - | - | - | .72 | .78 | .75 | .75 |
| Relations to achievement | .60 | .46 | .38 | .70 | -.30 | -.34 | -.09 | -.56 | -.40 | -.24 | -.67 | -.18 | -.23 | -.16 | -.10 | -.12 |

Invariance by Year Level

| Model | χ^2 ($\Delta\chi^2$) | df (Δ df) | p | CFI (Δ CFI) |
|--|-----------------------------|-------------------|-----|---------------------|
| Test 1: Baseline | 495.70 | 152 | .00 | .96 |
| Test 2: Invariant factor loadings | (23.26) | (21) | .33 | (.001) |
| Test 3: Invariant factor loadings , variances and covariances | (84.20) | (51) | .01 | (.004) |
| Test 3a: Invariant factor loadings and covariances | (56.00) | (39) | .04 | (.001) |
| Test 3b: Invariant factor loadings and covariances except the School–Ignore covariance | (48.93) | (36) | .07 | (.002) |
| Test 3c: Invariant factor loadings, covariances, and variances except the School–Ignore covariances and the variances of School and Ignore | (52.96) | (42) | .12 | (.002) |
| Relations to achievement | | | | |
| Test 4: Baseline | 599.82 | 222 | .00 | .96 |
| Test 5: Invariant path coefficients | (46.89) | (12) | .00 | (.004) |
| Test 5a: Invariant paths without school and ignore | (15.31) | (6) | .02 | (.001) |

Explaining year differences

- ▶ A recent meta-analysis into applicants' reactions to selection procedures (Hausknecht, Day, & Thomas, 2004) found that the strength of perception-performance relations was moderated by the context of the study (authentic vs. hypothetical).
 - Perhaps the higher the stakes, the more sensitive applicants become towards features of the testing environment.
 - Hence, stronger links between perceptions and performance are created.
- ▶ Accordingly, the Year 12 students in this study who had already experienced the high-stakes qualifications assessment system most clearly exhibited a trend toward personal accountability raising academic achievement ($\beta = .70$).

Explaining year differences

- We note that these differences might also reflect improved learning strategies that make the self-regulation efforts of Year 12 students more effective.
- We cannot determine whether the changes in strength of relations to achievement are a function of
 - increased experience with qualifications,
 - greater metacognitive self-regulation, or an interaction of both factors.

Methodology considerations

- ▶ A fundamental problem in the study of factorial invariance is the different conclusions one might draw based upon either χ^2 or CFI as indices of model fit.
 - While the CFI did not change beyond recommended thresholds in any of our comparisons, χ^2 was sensitive enough to detect slight differences.
 - Since it was possible to meaningfully interpret the variant parameters across the groups, it seems likely that relying on changes to CFI alone to interpret the results might have obscured important insights.
- ▶ The effect of different input matrices and estimation procedures also needs to be studied to ensure comparability of results with different statistical packages.
 - For example, it is generally argued that ordinal variables should be analysed using polychoric correlations as input in combination with WLS estimation (Flora & Curran, 2004) but the most widely cited simulation on fit indices in MGCFA by Cheung and Rensvold (Cheung & Rensvold, 2002) used AMOS, which does not allow for polychoric correlations as input, in combination with ML estimation.

Conclusion

- ▶ We didn't expect to find differences within these demographic groups so our explanations are post hoc and speculative
 - Systematic study is needed to determine whether the explanations are valid
- ▶ However, MGCFA invariance testing found that the SCoA was stable, but how the SCoA impacted achievement differed by more than chance according to sex and ethnicity.
 - The world of achievement is different

Some sources

- ▶ Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy and Practice*, 15(1), 3–17. doi: 10.1080/09695940701876003
- ▶ Brown, G. T. L., Harris, L. R., O'Quin, C. R., & Lane, K. (2015). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: Identifying and understanding non-invariance. *International Journal of Research and Method in Education*. Advance online publication. doi: 10.1080/1743727X.2015.1070823
- ▶ Hirschfeld, G. H. F., & Brown, G. T. L. (2009). Students' conceptions of assessment: Factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30–38. doi: 10.1027/1015-5759.25.1.30