



Measurement Invariance (MI): a general overview

Eric Duku
Offord Centre for Child Studies

21 January 2015



Plan

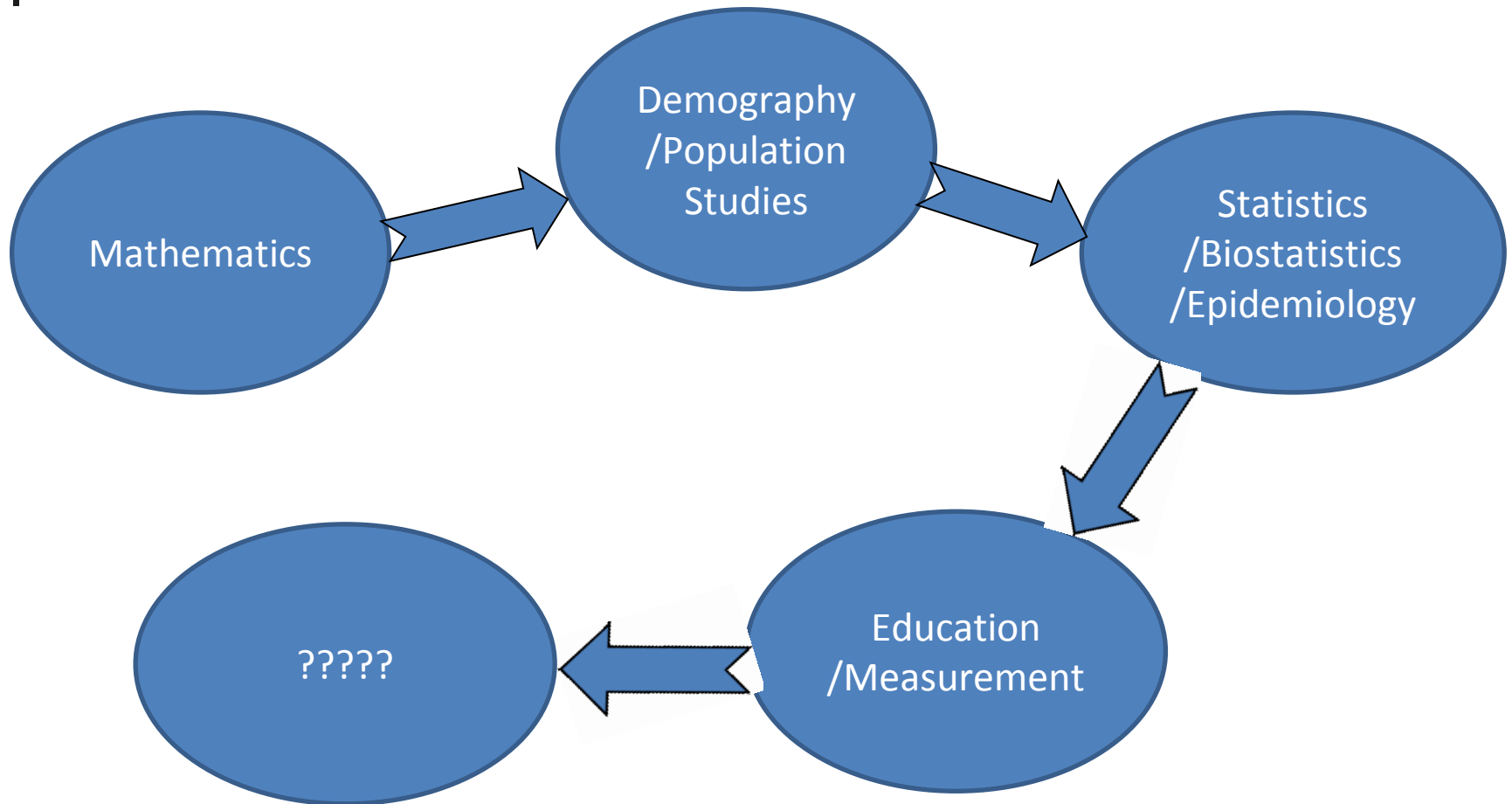
- Background
- What is Measurement Invariance
- Methodology to test MI
- Challenges with post-hoc MI analysis
- Summary
- Questions to ponder



“Mathematics is the foundation of many disciplines”...

BACKGROUND

Will the real Eric stand up?



How did I get to be interested in MI



- Almost 10 years ago
- International EDI work
 - Adaptation, translation, implementation
- Other studies at OCCS
 - adoption/adaptation of instruments
- OCCS study group
- UOttawa
- The rest is history...or is it?



WHAT IS MEASUREMENT INVARIANCE (MI)?



Measurement Invariance (MI)

- Fundamental component of measurement
 - Similarity of understanding of concepts, instructions
- Measurement Invariance
 - Measurement equivalence, measurement invariance, factorial invariance
 - Cross-cultural validity /Cross-cultural equivalence

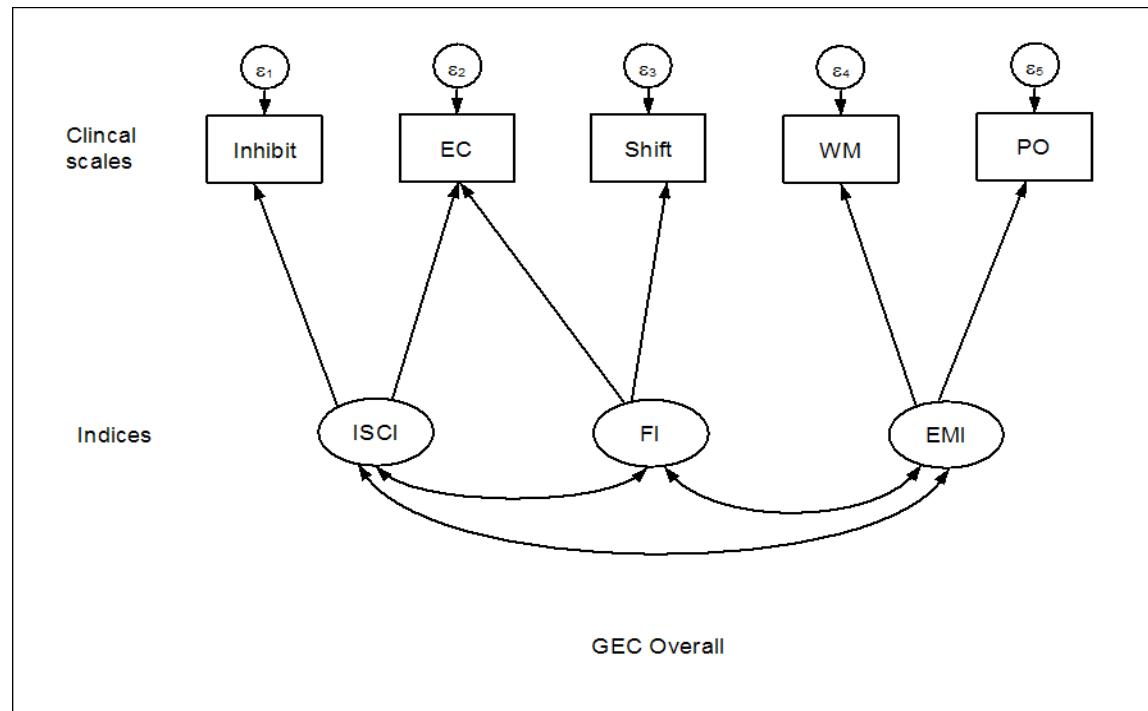


Measurement Invariance (MI)

Central principle of MI:

Relationships between items and traits (or latent constructs) are stable across groups & have the same psychometric properties across groups

Example: A Measurement model of Executive Function - BRIEF-P



Clinical scales: Inhibit, EC (Emotional Control), Shift, WM (Working Memory) and PO (Plan/Organize)

Indices: ISCI (Inhibitory Self-Control Index), FI (Flexibility Index) and EMI (Emergent Metacognition Index)

GEC = Global Executive Composite



What we know...

- History of MI is extensive beginning with admission tests and moving to other areas of research
- MI rarely or explicitly examined in ECD research

Vandenberg 2002; Lance & Vandenberg, 2000; Byrne & Watkins, 2003 ; Knight & Zerr, 2010;



**WHY DO WE NEED TO CHECK
FOR MI?**



Why we need to check for MI

- In ECD most research is comparative, longitudinal, or cross-sectional
 - e.g., Children regularly compared on constructs by group using differences in means and multivariate methods
- Accuracy of inferences has been shown to depend on measurement invariance of constructs over time and across groups
 - need to examine the underlying measurement structure of the constructs we are interested in within our sample, and subgroups within our sample



Why we need to check for MI

Measurement non-invariance usually arises when responses are based on perceptions rather than observable and measurable quantities

Meredith & Teresi, 2006

- Critical in ECD research where responses are based on parent or teacher perceptions



Possible contributors to Measurement non-invariance

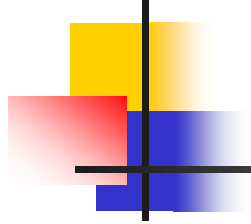
- Culture
 - shared language, ideas, beliefs, values, and behavioural norms (ethnicity or nationality)
- Age (development) - longitudinal
 - e.g. pre-grade 1 versus post grade 1
- Sample used
 - clinical versus community
- Informant/respondent (context)
 - Parent versus teacher



Contributors to Measurement Non-invariance

Other contributors

- instrument adaptation (e.g., translation),
- familiarity with item response formats, and
- many other socio-cultural factors e.g. taboos, gender, social desirability
- items having different meanings across groups; or
- instrument being administered under different conditions for different groups



METHODOLOGY TO TEST MI



Methodology used...

- Psychometric validity
 - typical & incomplete
 - internal consistency, construct & discriminant validity
- Rasch measurement model
 - Software - RUMM2030
- Multi-group Confirmatory Factor Analysis (MGCFA)
 - Software: Mplus, AMOS, STATA, LAVAAN, Mx, etc.



Methodology – Rasch analysis

- Based on the Rasch model, *G. Rasch, 1960*

probability of a given respondent affirming an item is a logistic function of the relative distance between the item location and the respondent location on a linear scale

e.g. person's level of anxiety and level of anxiety expressed by item
- Data should fit the model
- Iterative analyses

Overall fit to Rasch model; local dependency; extreme item and person residuals; disordered thresholds; unidimensionality; differential item functioning; invariance



Methodology – multi-group CFA

- Confirmatory Factor Analysis
 - most widely used method to test for MI
 - based on the original work by Karl Joreskog (1971)
 - permits testing for MI by setting cross-group constraints and comparing more restricted models with less restricted models



Process – multi-group CFA ?

- Is there a single measurement model established? Are there competing measurement models?
- Test fit of measurement model(s) in sample; and in each subgroup of interest
- Test for invariance; use constraints to test for differences between models
- Significant change in fit statistics => reject hypothesis of invariance



Methodology – multi-group CFA

Hierarchical iterative process:

- Configural invariance (or construct equivalence)
 - common factor model across groups
- Metric invariance (or measurement unit equivalence)
 - same meaning across groups
 - can compare factor variances/covariances
- Scalar invariance (or full scale equivalence)
 - can compare group means
- Uniqueness – unique variance of observed items
- Partial invariance



Implications of non-invariance

- Back to the drawing board
- Examine instrument adoption/adaptation process
- Figure out a new measurement model using EFA

A typical instrument creation or adaptation process

Investigation of conceptual and item equivalence	Literature review Discussion with experts in the field and members of target population
Original instrument translated	Translator I: Fluent in target language, good understanding of original language Translator II: Fluent in target language, good understanding of original language
A synthesized translated version	Translator III: Fluent in target language, good understanding of original language
Back-translations	Back-translator I: Fluent in original language, good understanding of target language Back-translator II: Fluent in original language, good understanding of target language
A synthesized back-translated version	Back-translator III: Fluent in original language, good understanding of target language
Expert committee	
Instrument pretested	
Revised instrument	
Investigation of operational equivalence	Literature review Discussions with experts in the field and members of target population
Main study	
Exploratory and confirmatory analysis	
Final instrument	



EXAMPLE



Example

THE JOURNAL OF CHILD
PSYCHOLOGY AND PSYCHIATRY

ACAMH THE ASSOCIATION FOR
CHILD AND ADOLESCENT
MENTAL HEALTH

Journal of Child Psychology and Psychiatry 54:12 (2013), pp 1346–1355

doi:10.1111/jcpp.12103

Measurement equivalence of the autism symptom phenotype in children and youth

**Eric Duku,¹ Peter Szatmari,¹ Tracy Vaillancourt,² Stelios Georgiades,¹ Ann Thompson,¹
Xiao-Qing Liu,³ Andrew D. Paterson,^{4,5} and Terry Bennett¹**

¹Offord Centre for Child Studies & Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada; ²Faculty of Education and School of Psychology, University of Ottawa, Ottawa, ON, Canada;

³Department of Obstetrics, Gynaecology and Reproductive Sciences, University of Manitoba, Winnipeg, MB, Canada;

⁴Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada; ⁵Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada



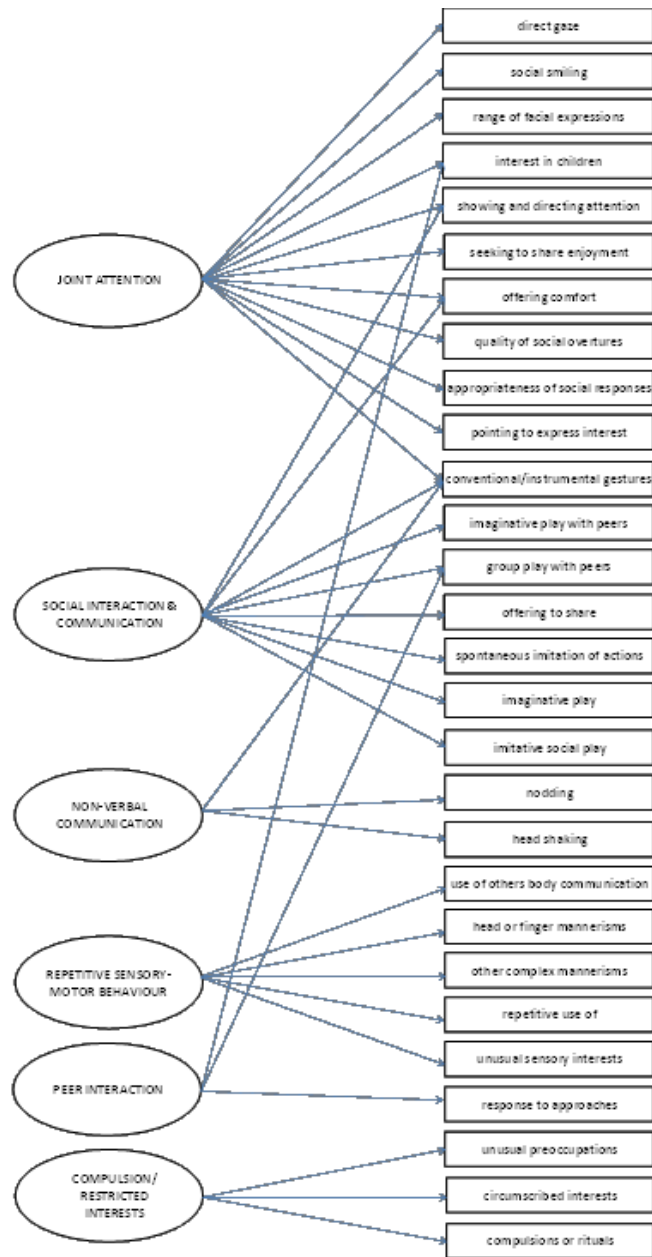
Background

Autism Diagnostic Interview - Revised (ADI-R)

Rutter, Le Couteur & Lord, 2003

- Semi-structured parent interview; 93 items, 34 used for diagnosis of ASD
 - 0=not present, 1=present in abnormal form, 2=definite abnormal, 3=extreme severity
- Measures impairment in 3 domains
 - Social interaction, Communication & Repetitive restricted behaviour
- 3628 individuals
 - Data on 28 algorithm items, child age, child sex and verbal status from Autism Genome Project (AGP) database
- Examine competing measurement models of the autism symptom phenotype using CCFA

Measurement model of autism symptom phenotype



Model	$\chi^2(\text{d.f.}) =$ statistic	CFI	TLI	RMSEA
Six-factor (Liu et al., 2011)	$\chi^2(220)=$ 2286.084	0.937	0.977	0.051
2 nd -order DSM-5	$\chi^2(224)=2$ 463.441	0.932	0.975	0.052



Results

- 2nd order 2-factor DSM 5 measurement model provided a good fit
- Autism symptom phenotype best characterized by six-factor 1st order measurement model
 - more parsimonious
 - MI across subgroups of ASD children – age, sex and verbal status



To be done...

- Examine MI across other subgroups of children e.g., IQ
- Examine the 2nd order 2-factor DSM-5 measurement model in other larger samples
- Replicate findings using other measures of the autism symptom phenotype e.g., Autism Diagnostic Observation Schedule (ADOS)

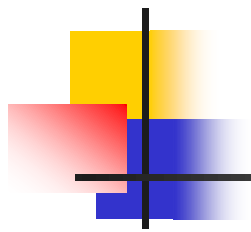


CHALLENGES WITH POST- HOC MI ANALYSES



Challenges with post-hoc MI analyses

- Availability of useable and accessible data
- Scope of data collected within each of the studies
- MI across subgroups beyond those in the data collected
- Cooperation of developers to have their already established instruments under scrutiny



SUMMARY



To recap...

- Credibility of inferences regarding the developmental process or developmental differences is affected by the absence of established measurement invariance
- MI needs to be examined when making comparisons across groups in ECD research since most scores cannot be directly measured
- recommended approach to testing and establishing MI is iterative and very comprehensive



Almost there...

- Importance to assess & verify MI prior to performing group comparisons and interpreting group comparisons in comparative ECD research
- Testing for MI provides useful information for the refinement of existing, widely used instruments
- **Commentary** (Hus et al.; JCPP, 2013)
 - Provides validation of the importance of this work by independent researchers on instruments with cooperation of developers



Almost final words...

- MI should not only be an a priori consideration, but also an aspect of critical appraisal of studies using and adapting instruments
- Assumption of universal applicability of instruments should be challenged and tested



Some questions to ponder...

- How do we bridge the gap between practitioners and academics?
- How do we make this methodology accessible to practitioners?
- How do we inform & educate ECD researchers to be aware of the implications of not examining MI a priori and as part of research plans?



Thank you...

- Members (too many to list 😊) of the Offord Centre for their support



Last slide...

“All models are wrong, but
some are useful”

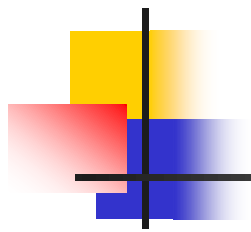
Box & Draper

Very last slide...



SEE, THEY ASKED HOW MUCH MONEY I SPEND ON GUM EACH WEEK, SO I WROTE, "\$500." FOR MY AGE, I PUT "43." AND WHEN THEY ASKED WHAT MY FAVORITE FLAVOR IS, I WROTE "GARLIC/CURRY."





RANDOM THOUGHTS???



Future random thoughts???

- Missing data
- Data Integration and Integrative data analysis
- Changing scales of measure



References used

- Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology, 34*, 155-175.
- Gjersing L., Caplehorn, J. R. M., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Medical Research Methodology, 10*:13.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical care, 44*(11), s78-s94
- Knight, G. P., & Zerr, A. A. (2010). Informed theory and measurement equivalence in child development research. *Child Development Perspectives, 4*(1), 25-30.
- Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*(11), s69-s77.
- Teresi, J. A. (2006). Overview of quantitative measurement methods equivalence, invariance, and differential item functioning in health applications. *Medical Care, 44*(11), s39-s49.
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: review of practice and implications. *Human Resource Management Review, 18*, 210–222.
- Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods, 5*(2), 139-158.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-69.