An Analysis of Clinical Data: Illustrating Equivalence of Unidimensional Item Response Theory and Cognitive Diagnosis Models

Kevin Carl P. Santos, Ph.D.

School of Statistics

University of the Philippines-Diliman

Jimmy de la Torre, Ph.D.

Faculty of Education

The University of Hong Kong

14th NCS National Convention on Statistics

1-3 October 2019 | Crowne Plaza Manila Galleria

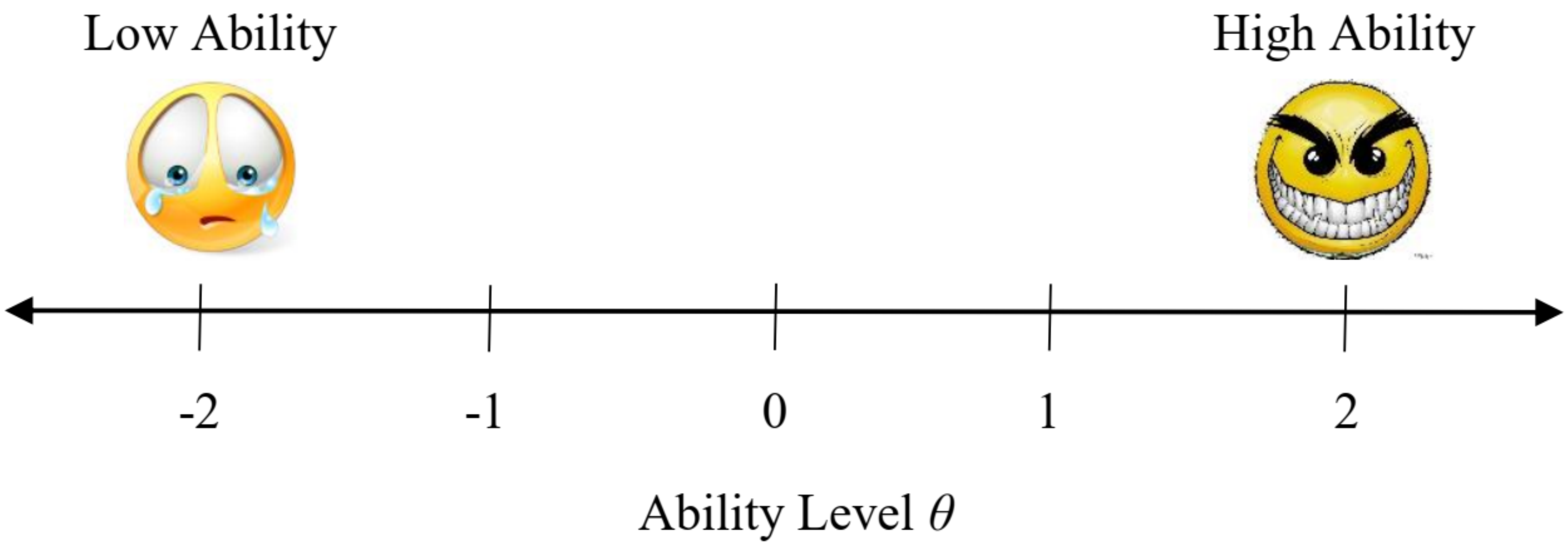
Organized by the Philippine Statistical System

Spearheaded by the Philippine Statistics Authority

Introduction

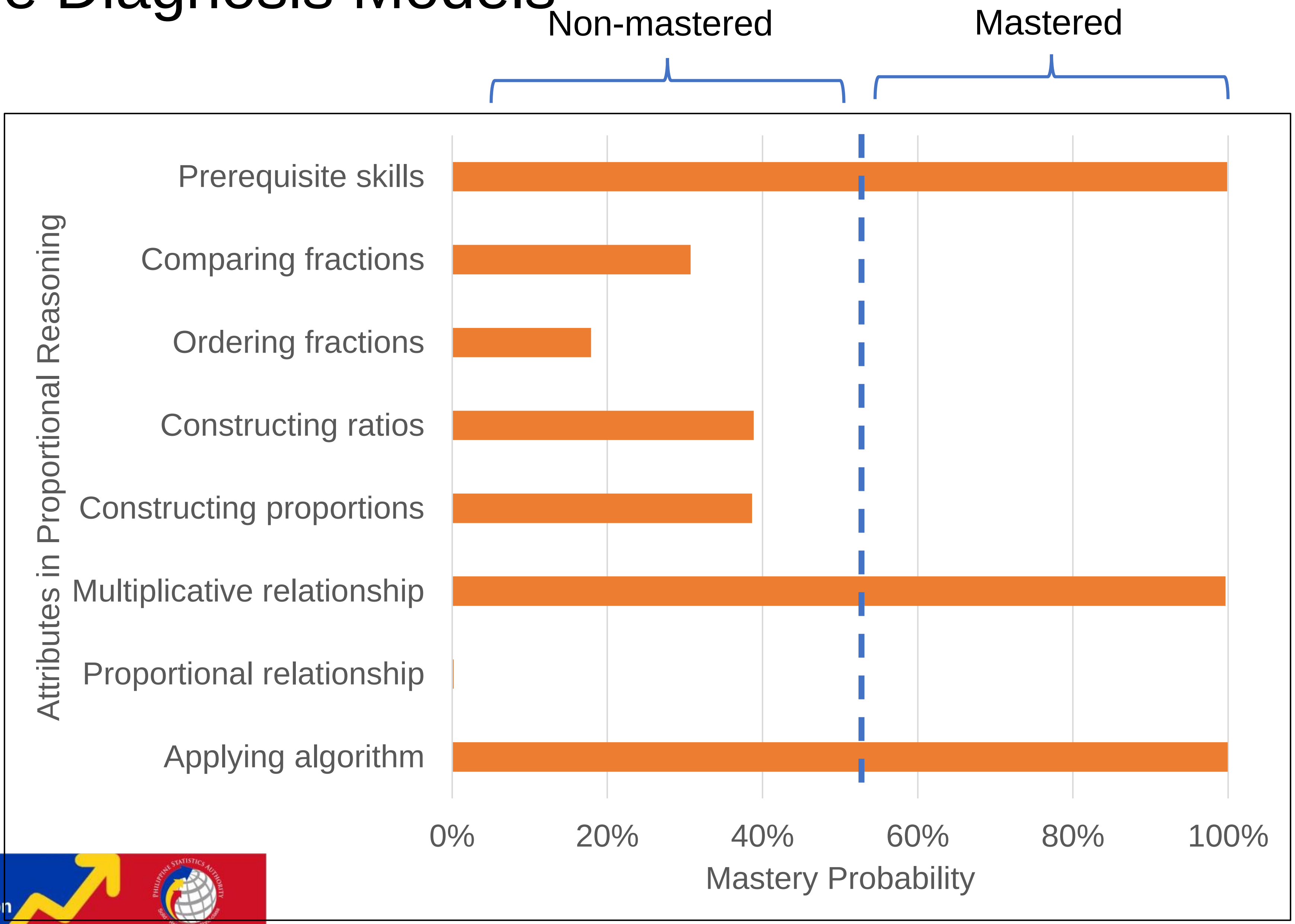
- At present, many existing educational assessments are developed and analyzed using unidimensional item response theory (IRT) models, which assume a single continuous latent variable
- To extract more diagnostic information, the same assessments have been retrofitted with cognitive diagnosis models (CDMs), which assume a multidimensional discrete latent variable
- However, it is not clear to what extent disparate psychometric frameworks can be used on the same data
- To address this issue, we propose a unifying framework for relating the two classes of model, as well as boundaries as to when this can be done

Unidimensional IRT Models



Proportional Reasoning

Cognitive Diagnosis Models



Equivalence of IRT and CDM

- In CDM, the marginal probability of x_j can then be written as

$$p(x_j) = \sum_{l=1}^{L} p(x_j|\alpha_l)p(\alpha_l),$$

where $p(x_j|\boldsymbol{\alpha}_l)$ is the *item response model* and $p(\boldsymbol{\alpha}_l)$ the *joint attribute distribution*

- One way of specifying $p(\boldsymbol{\alpha}_l)$ is to use a unidimensional higher-order latent trait, such that

$$p(\alpha_l|\theta) = \prod_{k=1}^{K} p_k(\alpha_{lk}|\theta),$$

where $p_k(\boldsymbol{\alpha}_{lk}|\theta)$ is the *attribute mastery function* (AMF)

Equivalence of IRT and CDM

- Hence,

$$p(x_j) = \sum_{l=1}^{L} \int_{\theta} p(x_j|\alpha_l)p(\alpha_l|\theta)p(\theta)\partial\theta$$

- For greater generality, the CDM can be represented by the **generalized *deterministic inputs*, *noisy*, “and” *gate* (G-DINA; de la Torre, 2011) model:**

$$P(\alpha_{lj}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk}\alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'}\alpha_{lk}\alpha_{lk'} + \cdots + \delta_{j12\ldots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk},$$

where δ_{j0} is the baseline probability, δ_{jk}s the main effects, δ_{jkk}s the two-way interaction effects, and $\delta_{j12\ldots Kj*}$ the highest order interaction effect

Equivalence of IRT and CDM

- To compare unidimensional IRT models and CDMs, we can express the CDM success probability on item *j* as a function of θ, as in,

$$p(x_j|\theta) = \sum_{l=1}^{2^K} p(x_j, \alpha_l|\theta) = \sum_{l=1}^{2^{K_j^*}} p(x_j|\alpha_{lj}^*)p(\alpha_{lj}^*|\theta)$$

- We need to further re-write $p(x_j|\theta)$ to better understand its properties
- For notational convenience, we write $p(\alpha_k = 1 \mid \theta) = p_k(1 \mid \theta)$ as p_k
- When only one attribute is required, $p(x_j|\theta)$ simplifies to

$$\begin{aligned} p(x|\theta) &= \sum_{\alpha_1=0}^{1} p(x|\alpha_1)p_1(\alpha_1|\theta) \\ &= \delta_0 + \delta_1 p_1. \end{aligned}$$

Equivalence of IRT and CDM

- When K_j^* attributes are required, $p(x_j|\theta)$ can be expressed as

$$
\begin{aligned}
p(x_j|\theta) &= \sum_{\alpha_1=0}^{1} \cdots \sum_{\alpha_{K_j^*}=0}^{1} p(x_j|\alpha_1,\ldots,\alpha_{K_j^*})p(\alpha_1,\ldots,\alpha_{K_j^*}|\theta) \\
&= \delta_0 + \sum_{k=1}^{K_j^*} \delta_k p_k + \sum_{k=1}^{K_j^*-1} \sum_{k'=k+1}^{K_j^*} \delta_{kk'} p_k p_{k'} \\
&\quad + \cdots + \delta_{1\ldots K_j^*} \prod_{k=1}^{K_j^*} p_k
\end{aligned}
$$

- We refer to this as the **reformulated HO-GDINA (RHO-GDINA) model**

Equivalence of IRT and CDM

Sufficient Conditions for Monotonically Nondecreasing $p(x_j|\theta)$

- For $p(x_j|\theta)$ to be monotonically nondecreasing, the following sufficient conditions need to be met:

(1) The AMF of p_k, $k = 1,\ldots,K$ should be of the form

$$p_k = \frac{\exp[\beta_k(\theta - \lambda_k)]}{1 + \exp[\beta_k(\theta - \lambda_k)]},$$

where β_k and λ_k are the discrimination and difficulty parameters with respect to attribute k

(2) $p(x_j|\boldsymbol{\alpha}_l^*) \leq p(x_j|\boldsymbol{\alpha}_{l'}^*)$ whenever $\boldsymbol{\alpha}_l^* \leq \boldsymbol{\alpha}_{l'}^*$ (monotonicity property)

Millon Clinical Multiaxial Inventory-III

- For MCMI-III has been indicated by clinicians as being one of the most frequently used self-report instruments for clinical assessment
- To illustrate the IRT and CDM equivalence, we analyzed the responses of 1,210 subjects to 130 statements of the Dutch version of the MCMI-III
- The statements measure 16 clinical disorders, namely,

α_1	Depressive	α_9	Somatoform
α_2	Sadistic	α_{10}	Bipolar
α_3	Negativistic	α_{11}	Dysthymia
α_4	Masochistic	α_{12}	Drug Dependence
α_5	Schizotypal	α_{13}	Post Traumatic Stress
α_6	Borderline	α_{14}	Thought Disorder
α_7	Paranoid	α_{15}	Major Depression
α_8	Anxiety	α_{16}	Delusional Disorder

Millon Clinical Multiaxial Inventory-III: Method

- The model fit of the saturated and higher-order (1-parameter logistic or 1PL and 2PL) G-DINA models, and the four unidimensional IRT models (i.e., 4PL, 3PL, 2PL, and 1PL) were compared

- The AIC and BIC were employed for relative fit evaluation

- The correlations of the different $\hat{\theta}$s were calculated

- To compare the IRT and CDM estimates, the number of disorders were plotted against the latent trait estimates

Millon Clinical Multiaxial Inventory-III: Results

Model	AIC	BIC
4PL	161130	163781
3PL	159581	161569
2PL	**159433**	**160759**
1PL	165675	166343
Saturated	282580	621740
2PL-GDINA	**155332**	**160533**
1PL-GDINA	156154	161278

- Among the IRT models, the 2PL obtained the lowest AIC and BIC
- Based on AIC and BIC, the 2PL-GDINA model fitted the data better than 1PL-GDINA or saturated GDINA model
- It can be noted as well that when compared with the four IRT models the 2PL-GDINA model had the lowest AIC and BIC

Millon Clinical Multiaxial Inventory-III: Results

Correlation	2PL-GDINA	1PL-GDINA
4PL	0.98	0.97
3PL	0.96	0.95
2PL	0.96	0.96
1PL	0.95	0.97

- The latent trait estimates obtained from the HO-GDINA and the unidimensional IRT models are highly correlated
- This is consistent with the results of the simulation study and real data analysis on proportional reasoning assessment previously conducted by the authors

Millon Clinical Multiaxial Inventory-III: Results

Disorder	β_k	λ_k	Disorder	β_k	λ_k
α_1	4.29	-0.36	α_9	2.41	-0.27
α_2	1.05	-0.74	α_{10}	1.31	-0.32
α_3	3.19	0.46	α_{11}	3.40	0.40
α_4	4.67	-0.90	α_{12}	0.10	-1.26
α_5	3.49	-1.23	α_{13}	1.72	-0.46
α_6	2.79	-0.98	α_{14}	4.12	0.35
α_7	1.48	0.00	α_{15}	2.75	-0.29
α_8	4.24	0.31	α_{16}	1.58	-1.29

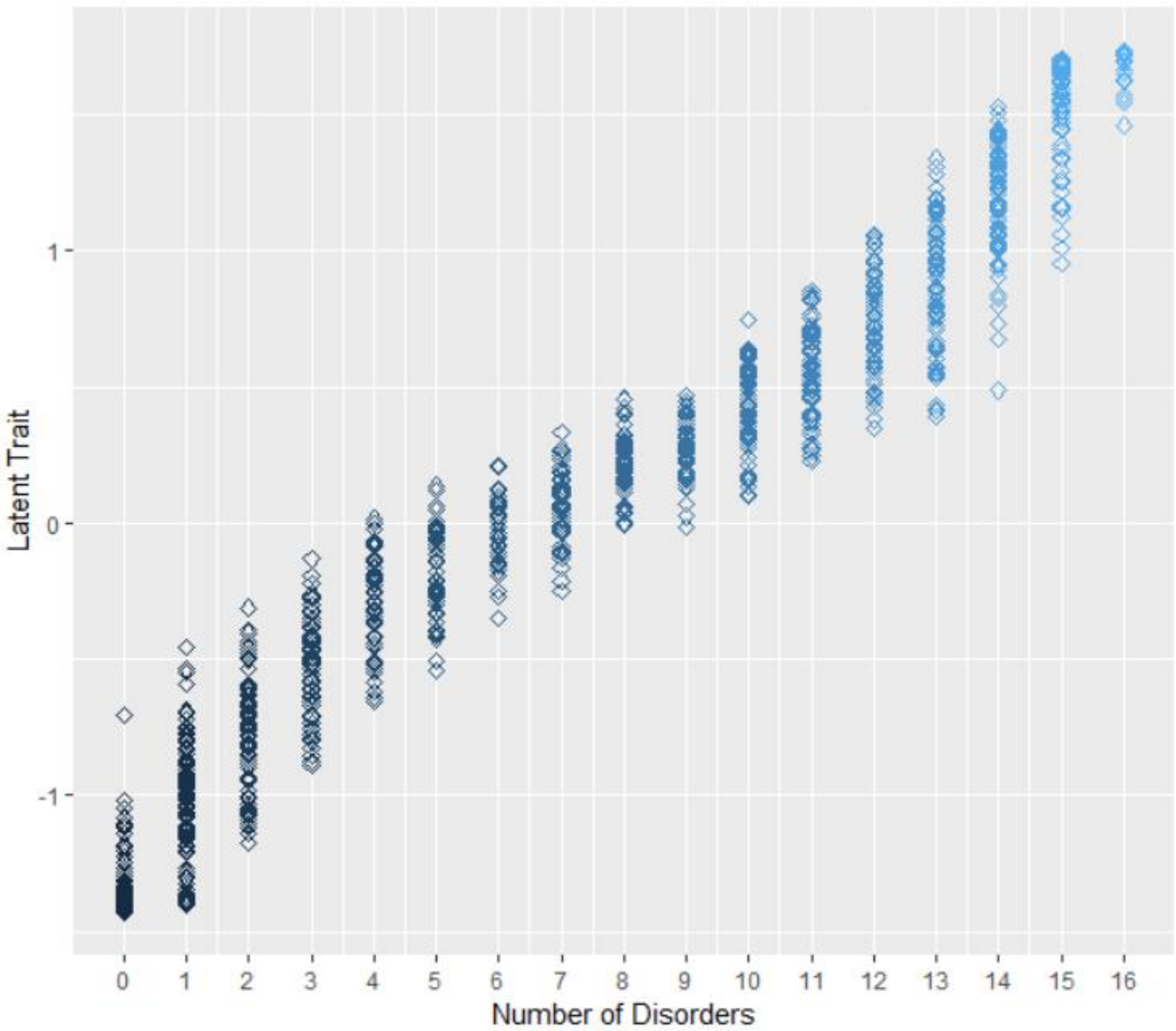
- Except for drug dependence, all slope coefficients had very values, suggesting a strong relationship of each disorder to the general latent trait (i.e., an indication of unidimensionality)

Millon Clinical Multiaxial Inventory-III: Results

Higher Order $\hat{\theta}$

versus

Number of Disorders

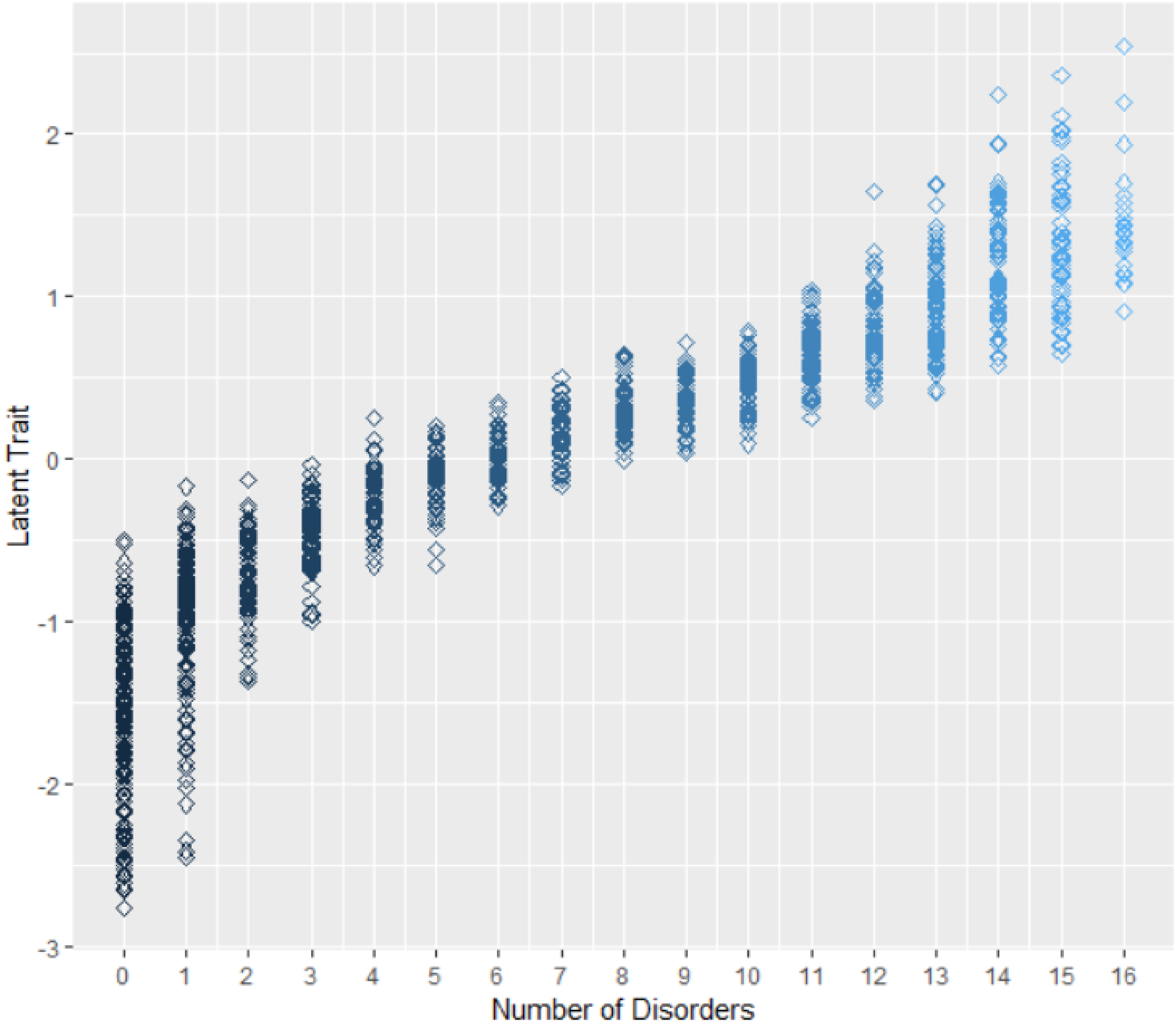


Millon Clinical Multiaxial Inventory-III: Results

IRT $\hat{\theta}$

versus

Number of Disorders



Millon Clinical Multiaxial Inventory-III: Discussion

- The subjects with higher latent traits possess more disorders
- When the number of disorders was fixed, the corresponding values of the latent trait varied and overlapped with different number of disorders
- Hence, using the latent trait estimate solely would be insufficient in targeting the specific disorders that need to be addressed
- Nevertheless, this work provides a framework for relating the two classes of psychometric models

Millon Clinical Multiaxial Inventory-III: Discussion

- Under certain conditions (e.g., AMF slope is large), the HO-GDINA model can be approximated by IRT models

- As shown in the real data analysis, almost all slope coefficients were very large, producing high correlation of the HO-GDINA and IRT model latent trait estimates

- Thus, in addition to finer-grained attributes, estimating the overall ability (or general latent trait) is also reasonable

- When the HO assumption is reasonable, IRT models can be fitted to CDM data to obtain an approximation of the HO ability

Fin.

Thank you very much!

kpsantos1@up.edu.ph