# Preliminary Application of Cognitive Diagnostic Models (CDM) to Medical Education Assessment of Entrustment

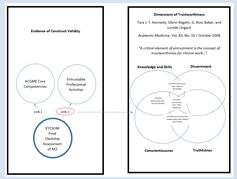
**Brock Mutcheson PhD, Tracey Criss MD, Richard C. Vari PhD**

***Virginia Tech Carilion School of Medicine***

## CONCEPTUAL FRAMEWORK

- The purpose of this proof-of-concept study was to apply a modern measurement model to student assessments to investigate the potential for the formative use of granular and categorical inferences into student entrustment. It is expected that when equipped with the more precise diagnostic information provided by a cognitive diagnostic model, faculty and administration may be able to more effectively and confidently identify strengths and areas of opportunity in guiding students to entrustment. This particular analysis included results from the preliminary empirical investigation into the construct internal structure using exploratory factor analysis (EFA).
- "A critical element of entrustment is the concept of trustworthiness for clinical work" (Kennedy, et al., 2008). Trustworthiness for clinical work consists of 4 dimensions:

1) knowledge and skill;
2) discernment of limitations;
3) truthfulness ;
4) conscientiousness.




*Kennedy, Reghr, Baker and Lingard (2008)*

- VTCSOM Administration and Faculty experts mapped the VTCSOM Final Clerkship Assessment of M3 Student to ACGME Core Competencies and the Entrustable Professional Activities. This work was institutionally approved by the BIC2 and MCC committees.
- The concept map (above) showed an example of how assessment items will often theoretically and empirically measure more than one element, attribute, or dimension of latent constructs such as "trustworthiness." Latent constructs are mental abstractions that can not be directly observed and must be measured using indicators. The map included above does not reflect all VTCSOM items, nor does it include the full text of any item. Assessment methodologies such as cognitive diagnostic models (CDM) (de la Torre, 2004) provide the ability to account for complex inter-item variance, and thus, they provide a more precise estimate of proficiency on each hypothesized dimension.

## CONCLUSIONS

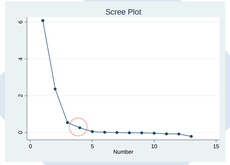
- The exploratory factor analysis with an oblique oblimin rotation yielded adequate assessment model fit results with a simple 4-factor loading structure solution. Although this commonly applied methodology provided valuable empirical information about the internal structure of the construct, its assumptions, and thus, its conclusions lacked the appropriate theoretical nuance expected in the measurement of a complex latent construct like entrustment.
- The information from this solution will be incorporated into the development of the diagnostic model, which is anticipated to prove more precise and useful in understanding the empirical nature of entrustment.

## HYPOTHESIS

- **Right:** This figure is an example of the Q-Matrix Hypothesis (Tatsuoka, 1983). A Q-matrix depicts which skills/attributes contribute to the probability of a higher score on an item. By accounting for complex loadings, items can be assigned to multiple skills, and thus, the more skills assigned to an item, the more skills that affect the probability of a higher response outcome on that item. Q-matrix indicator entries are binary in that a skill affects the probability of a more positive response or it does not. In this exploratory analysis, we are investigating the internal structure in order to inform the Q-matrix for this study.

| Assessment Item | Attributes: Dimensions of Entrustment (Kennedy et. al.) | | | |
|---|---|---|---|---|
| | #2: Knowledge and Skill | #2: Discernment | #3: Conscientiousness | #4: Truthfulness |
| #1 Medical Knowledge… | 1 | 0 | 0 | 0 |
| #2 Interpretation Skills… | 1 | 1 | 0 | 0 |
| #3 Data-Gathering Skills… | 1 | 0 | 1 | 0 |
| … | … | … | … | … |
| #11 Recognition of Biases and Diversity… | 0 | 1 | 1 | 1 |
| … | | | | |

## PRELIMINARY RESULTS



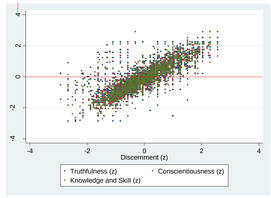

- Results from steps 1-3 were displayed. Eigenvalues were scree plotted and a 4-factor solution was retained. The primary estimated rotated factor loadings were included below. Factor reliability (mean α=0.8) was assessed. Factor scores were calculated, standardized, plotted, and reviewed to better understand multidimensional relationships and inform Q-matrix construction for future model specification.

**Key Rotated Factor Loadings (Pattern Matrix) and Unique Variances**

| VTCSOM Assessment | ACGME Competency | EPA | $\lambda_1$ Truthfulness | $\lambda_2$ Conscientiousness | $\lambda_3$ Knowledge and Skills | $\lambda_4$ Discernment |
|---|---|---|---|---|---|---|
| 1: Medical Knowledge/Basic Science | Medical Knowledge | EPA 2,3,7 | - | - | 0.61 | - |
| 2: Interpretation Skills | Medical Knowledge | EPA 1,2,3,5,7,10 | - | - | - | 0.520 |
| 3: Data-Gathering Skills | Patient-Care | EPA 1 | - | - | 0.625 | - |
| 4: Clinical Reporting Skills | Patient-Care | EPA 1,5,6 | - | - | - | 0.616 |
| 5: Procedural Skills | Patient-Care | EPA 12 | - | - | - | 0.230 |
| 6: Communication Skills with Patients, Families, Team Members, and Staff | Interpersonal Skills | EPA 6,8,9,11 | - | - | - | 0.460 |
| 7: Relationships with Patients and Families | Interpersonal Skills | EPA 3,4,5,6,9,11,12 | - | 0.750 | - | - |
| 8: Professional Relationships | Professionalism | EPA 9 | - | 0.803 | - | - |
| 9: Educational Attitudes | Professionalism | EPA 7 | 0.636 | - | - | - |
| 10: Dependability and Responsibility | Professionalism | EPA 5,13 | 0.642 | - | - | - |
| 11: Recognition of Biases and Diversity | Professionalism | EPA 1,2,4,5,6,11 | 0.451 | - | - | - |
| 12: Systems-based practice | Systems-Based Practice | EPA 3,4,5,9,11,12,13 | - | 0.292 | - | - |
| 13: Management Skills | Practice-Based Learning and Improvement | EPA 1, 2, 3, 4, 5, 7, 8, 10, 11, 12 | - | - | 0.647 | - |

**Factor Correlation Matrix**

| | 1: Truthfulness | 2: Conscientiousness | 3: Knowledge and Skill |
|---|---|---|---|
| 1: Truthfulness | 1 | | |
| 2: Conscientiousness | 0.847 | 1.000 | |
| 3: Knowledge and Skill | 0.805 | 0.844 | 1.000 |
| 4: Discernment | 0.786 | 0.805 | 0.838 |

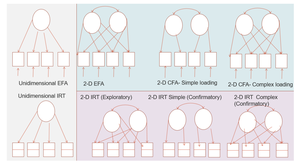



**Factor Scores**

| | N | Mean | Min | Max |
|---|---|---|---|---|
| Truthfulness | 1,681 | 3.9 (0.49) | 2.0 | 5 |
| Conscientiousness | 1,681 | 3.9 (0.44) | 1.3 | 5 |
| Knowledge and Skill | 1,681 | 3.8 (0.42) | 2.6 | 5 |
| Discernment | 1,681 | 3.8 (0.48) | 1.5 | 5 |

## METHODOLOGY

- **Step 1:** *Literature Review*
- **Step 2:** *Data Collection*: Aggregate six years of clerkship data (n = 1681)
- ***Step 3:*** *Empirical validation approach I*: Explore dimensionality using traditional eigenvalue-based exploratory factor analysis (EFA) with maximum likelihood estimation method and oblique oblimin rotation (Liu, Douglas & Henson, 2009). Demonstrate EFA simple solution comparing Eigenvalues and residual variances. Further investigate quality of measured factors and indicators, eliminating poor measures and factors. Determine model fit. Visualize scree plot and interpret factors.
- **Step 4:** *Expert input:* Further develop hypotheses to explore and test complex factor loadings accounted for in Q-matrix hypothesis using CDM.
- **Step 5:** *Model Specification and Fit:* Determine model, compare model fit statistics and item parameters




***Left:*** *Concept illustration comparisons of common assessment approaches to investigate latent construct dimensionality and internal structure.*
***Below:*** *The proposed cognitive diagnostic model to account for complex loading structures in this study.*

**The Compensatory Re-Parameterized Unified Model (C-RUM)**

$$\pi_{ic} = P(X_{ic} = 1 \mid \alpha_c) = \exp(\lambda_{i,0} + \sum_{a=1}^{A} \lambda_{i,1,(a)} \alpha_{ca} q_{ia}) / 1 + \exp(\lambda_{i,0} + \sum_{a=1}^{A} \lambda_{i,1,(a)} \alpha_{ca} q_{ia})$$

where P was the probability of a positive response (i.e. a response of 1),
exp (.) was the exponential function,
$\pi_{ic}$ is the probability of positive response to item i in latent class c,
$X_{ic}$ was the observed response for item i in latent class c,
$q_{ia}$ was the indicator from the Q-matrix indicating whether attribute a is measured by item i,
$\alpha_{ca}$ was the attribute mastery indicator for attribute a in latent class c,
$\lambda_{i,0}$ was the intercept parameter for item i, and $\lambda_{i,1,(a)}$ was the slope of parameter for item i and attribute a.

## LIMITATIONS AND FUTURE DIRECTION

- The research team also intends to investigate methods to account for the variance attributed to experiential differences to increase internal validity of the applied model. One important source of this variance occurs as a result of "clerkship order." More specifically, students experience VTCSOM core clerkships in different orders than other students and thus scores must be horizontally equated across points in time. Sample sizes and conditions likely permit this calculation and so once this is further understood using appropriate methods, more precise solutions can be explored. Inter-rater reliability is another key source of variance to be further investigated.
- Using this information, several diagnostic models will be tested and compared to the results from the simple loading structure and the compensatory re-paramterized unified model.