The DAATS Model: Initial Psychometric and Statistical Findings
A Top Ten Illustration

Paper Presented at the Annual Meeting of the American Association of
Colleges of Teacher Education (AACTE)
New Orleans, LA.
February 2008

W. Steve Lang, Ph.D.
Professor, Measurement and Research
University of South Florida St. Petersburg
email: lang@stpt.usf.edu

Abstract

The INTASC Principles, when used as the basis for developing appropriate measurement instruments to assess teacher dispositions, provide a viable approach to the diagnosis and remediation of skill-related affective performance in teacher candidates and also to meeting NCATE requirements for Standard 1. In this symposium, the development and use of a battery of linked and calibrated disposition measures is discussed.   This paper highlights the psychometric findings of a pilot study of three instruments designed to assess dispositions. This paper follows the overview found in *The DAATS Model: Where it Comes from and What it is* (Wilkerson, 2008).

Statement of the Issue

Others have already stated (Hays, 1988), "It is unfortunate that in their search for quantitative methods researchers sometimes overlook the question of level of measurement and tend to read quite unjustified meanings into their results. However, the core problem of level of measurement lies outside the province of mathematics and statistics" (p. 71).   For our applications' purposes which include prediction of future behaviors, diagnosis of differential performance, job-related validity, and comparative measurement of growth, the instrument(s) development included development of an interval level scale from the raw scores.

Disposition assessment appears to pose a difficult assessment challenge according to some recent reviews (Diez, 2008).   Here, we differ from that viewpoint. In this report on the use of three (of a battery of five) assessment instruments, traditional item types are combined with modern construct modelling and analysis to demonstrate a valid and reliable approach to practical assessment. This short report is intended to illustrate some basic results to common questions that teacher educators might ask about such measurement as reported in graphical and numerical formats. Here, we hope to find reasonable and appropriate ways to generate practical applications for typical teacher education programs who wish to assess dispositions while maintaining psychometric integrity. Even though the construct basis (INTASC Principles) and item types (Thurstone, questionnaires, rating scales) are both common, the analysis and reporting are possibly new to some viewers.

Literature Review

In this short symposium, a few notable and outstanding findings are presented as important to the validity and utility of the current assessment development. Even though comprehensive results are not presented here, the analysis and basis for psychometric development and statistical reporting were based on guidelines and recommendations from the following sources:

Bond, T. & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nd).* Mahwah, NJ: LEA.
Linacre, J. M. (2003). *A user's guide to FACETS: Rasch-model computer programs.* MESA Press: Chicago.
Linacre, J. M. (2003). *A user's guide to WINSTEPS: Rasch-model computer programs.* MESA Press: Chicago.

Smith, E. & Smith, R., Eds. (2004). *Introduction to Rasch Measurement.* Maple Grove,
     MN: JAM.
Wilson, M. (2005). *Constructing measures: An item response modelling approach.*
     Mahwah, NJ: LEA.

<u>Method</u>

<u>Analyses</u>

All analyses were completed using three commonly available software programs:

1. *Winsteps:* Rasch measurement software for persons & items. Multiple choice
    questions, rating scales, and partial credit can be analysed with up to 255 categories
    per scale.

2. *FACETS:* Many-facet Rasch measurement for persons, items, judges, and
    tasks is software that assesses rater error and calibrates different variables on
    parallel scales.

3. *Systat 12:* SYSTAT is statistical analysis and graphical software.

<u>Sample</u>

Three public universities in Florida contributed data from the students in their teacher
education programs. At the time of this writing, there were 228 subjects included in a variety
of programs within the various colleges. For the purposes of this short report, each finding
in the results describes the demographics or characteristic of interest. Specific programs and
students are hidden.

<u>Results: The Top Ten</u>

In a short paper, it is a challenge to report all of the results. Here we will highlight the top ten
results we think will be of most relevance and interest to a policy-making audience. Will present each
result in a question and answer format.

1.) Are the three instruments, individually and collectively "reliable"?

The Rasch model computes a model reliability called Separation Reliability. Basic results
from Winsteps are included here. In this case, the numbers can be roughly interpreted as a
Cronbach's Alpha.   These results indicate that the entire scale and each different item type
discriminate between persons well. **The answer is "Yes".**

Table 1. *Rasch Reliability Estimates*

Instrument	Separation	Reliability	Number of Items
Total Battery	5.11	.96	90
BATS	3.83	.94	60
ETQ (rater unadjusted)	1.71	.75	10
ETQ (rater adjusted)	2.14	.82	10
SRA	2.73	.88	20

2.) What does the distribution of scores look like?

```
TABLE 1.1 Disposition Rating Analysis 08
INPUT: 228 persons, 90 items
--------------------------------------------------------------------------------

                          MAP OF persons AND items
MEASURE                                   |                           MEASURE
  <more> -------------------- persons -+- items   -------------------- <rare>
  100                                     +                            100
                                          |
                                          |   X
   90                                     +                             90
                                          |T
                                          |   X
                                          |
                                          |
   80                          X    +  XX                              80
                                     |   D
                         XXXX  T|
                          XXX    |    XXXX
                        XXXXXXX  |    XXXX
   70          XXXXXXXXXXXXX  S+  XX                                    70
             XXXXXXXXXXXXXXXX  |S XXXD
          XXXXXXXXXXXXXXXXXXXX  |   XDDD
        XXXXXXXXXXXXXXXXXXXXXXX M|   XXX
      XXXXXXXXXXXXXXXXXXXXXXXXX  |   XD
   60          XXXXXXXXXXXXXX  +  XX                                    60
             XXXXXXXXXXXX  S|   XDD
               XXXXXX    |    X
                  XX     |    XXD
                  XX  T|   DDD
   50            XXXX   +M DDDD                                         50
                         |   DD
                         |   DDDD
                   X     |   DD
                   X     |   DDD
   40                    +   DD                                        40
                         |   DDD
                         |   DDDD
                         |   DD
                        |S DDD
                         +   DDDD                                      30
                         |
                         |   DDDDDD
                         |   DD
                         |
                         +   DDDD                                      20
                         |
                         |
                         |
                        |T D
   10                    +  D                                          10
  <less> -------------------- persons -+- items   -----------------<frequent>
```

An X on the right side is a rating scale item (ETQ or SRA), a D is a Thurstone (BATS) item. The middle range of the scale is 50; indicated by the M on the right.

An X on the left side is a person, The people appear to be normally distributed. The typical student has a scale score of about 65. A score of about 50 appears "low" and about 75 appears "high".

Notes:

We would expect most teacher educators to report dispositions consistent with INTASC principles. Most do even though there is variation between students and there is no a "ceiling effect" where everyone reports 100% are "target". This is important for precise measurement, improvement, and analysis. **The scale as described from 0 to 100 is an *interval level scale* which is appropriate for almost any statistical analysis. It is not a percent scale.**

3.) Is there evidence of growth within students?

Least Squares Means
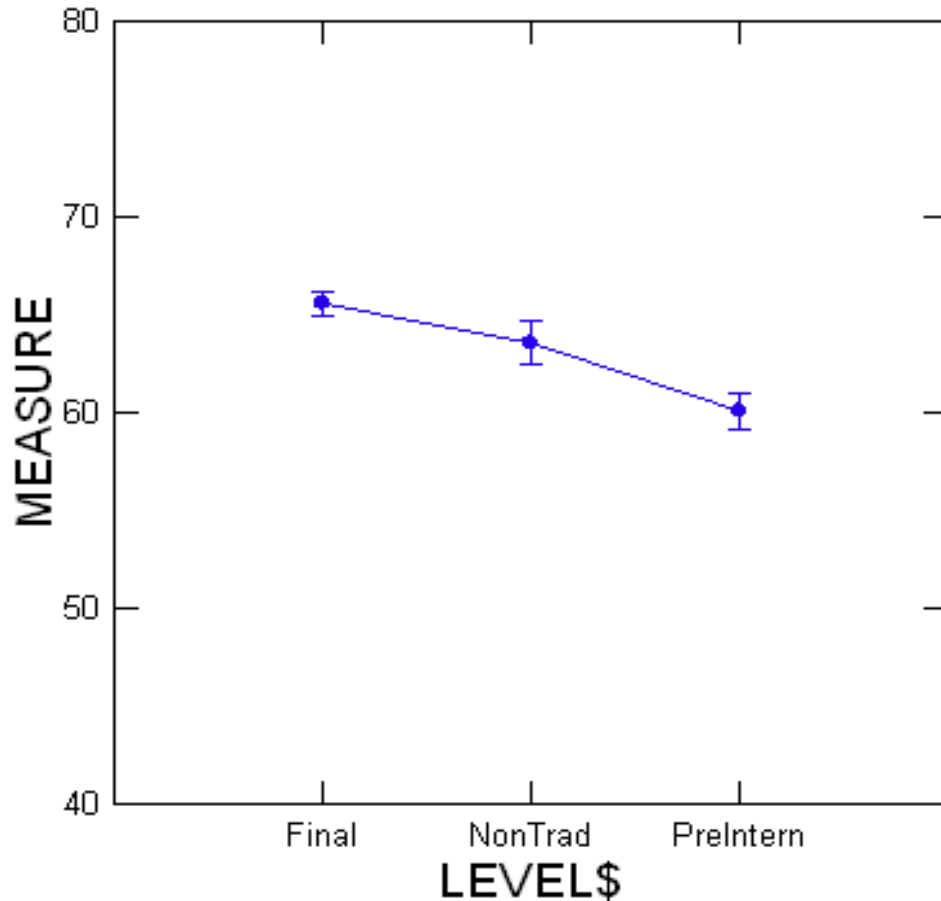
Notes: Yes, even with small samples and over the Junior and Senior years of teacher education. **These results indicate a change in consistency with INTASC principles that is meaningful.** Stated differently, during their teacher preparation programs, students appear to be acquiring increasing commitment to the skills of teaching, as operatinally defined in the INTASC Principles. This graph is also confirmed by an ANOVA and the points plotted are the least-squared means from that analysis. Also, the variability of teacher candidates as final interns is the smallest. The variablility of the nontraditional (alternative certification) candidates is the greatest. As students progress to final internship, they become more consistent and more homogeneous in their consistency with INTASC principles. This is evident in the standard errors indicated in the points plotted above by the brackets above and below each plotted point.

4.) Is there a correlation between disposition assessment and cognitive assessment?

Number of Observations: 24

Pearson Correlation Matrix		
	COGNITIVE	DISPOSITION
COGNITIVE	1.000	
DISPOSITION	0.098	1.000

Scatter Plot Matrix
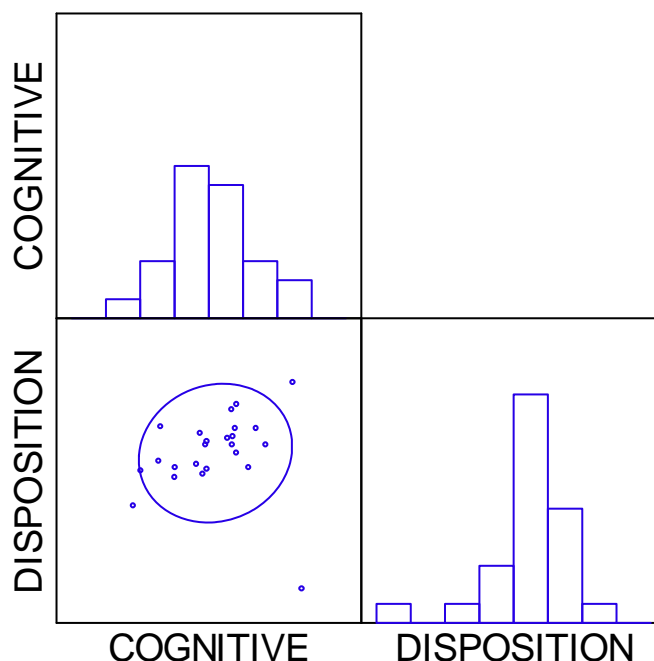
Notes: One subset of the pilot included a class of pre-internship education majors who completed the three disposition instruments (90 items). The same students completed three content tests (225 items). Both the three disposition instruments and the three content tests were calibrated separately using the Rasch model. The disposition instruments were framed by Krathwohl's affective taxonomy. The content instruments were framed by Bloom's cognitive taxonomy. Both sets of scores indicated a normal distribution. The correlation between the two sets of scores on the same students in this same class is .098. **This indicated an almost exclusively separate construct for the sample.**

Also, please notice the outlier student in the lower right corner with high cognitive performance and low disposition performance. That student has an A-minus GPA and is the "growling orange dog" student cited in the Wilkerson paper. When asked what he would say to a student showing his middle finger in a tantrum (SRA), this is the student who wrote: *"Well, I'd force myself to transform into a fierce orange dog, with large-cold eyes and long-white curved fangs. I'd leap onto his desk, raise the hair on my back, show my teeth, and growl, "would you like to repeat that?" Would I want to be his teacher? Why not. It doesn't bother my self-worth any. That child can't grow up doing that, so a swift, immediate response would end that behavior permanently in my eyes."*

The same outlier response to a question (ETQ) about keeping abreast of current developments in the field: *"To be honest, I simply read the newspaper everyday."*

This student was the fourth lowest of 228 students on the BATS assessment of consistency with INTASC.

© 2008 W. Steve Lang, Ph.D.

5.) Is it possible to give each student a score on each of the 10 separate INTASC principles?

```
+----------------------------------------------------------------------------+
| item     OBSERVATIONS     BASELINE     DPF     DPF     DPF     person       |
| CLASS   COUNT AVERAGE EXPECT MEASURE   SCORE MEASURE  S.E.    Number  Name   |
|----------------------------------------------------------------------------|
| 01        8    1.50   1.25   68.30     .25   81.23    9.78        1 **001*****|
| 02        8    1.13   1.37   68.30    -.25   58.47    6.42        1 **001*****|
| 03        7     .86   1.06   68.30    -.20   57.21    8.17        1 **001*****|
| 04        6    1.33   1.07   68.30     .26   89.10   19.11        1 **001*****|
| 05        7    1.43   1.34   68.30     .09   71.80    7.95        1 **001*****|
| 06        7    1.29   1.31   68.30    -.02   67.51    6.97        1 **001*****|
| 07        7    1.14   1.09   68.30     .05   70.19    7.40        1 **001*****|
| 08        7    1.00   1.07   68.30    -.07   65.56    7.39        1 **001*****|
| 09        8    1.13   1.06   68.30     .07   70.54    6.49        1 **001*****|
| 10        8    1.00   1.12   68.30    -.12   64.67    6.19        1 **001*****|
```

<u>Notes</u>: In the table above is the first of the 228 students in the analysis. The highlighted column labelled "DPF MEASURE" is the student's scale score for each of the 10 INTASC principles listed on the far left as "item CLASS". In reporting to a student in class, it is easier to provide a simple graph to facility student self-analysis. **Yes, it is possible to score subscales by INTASC.**
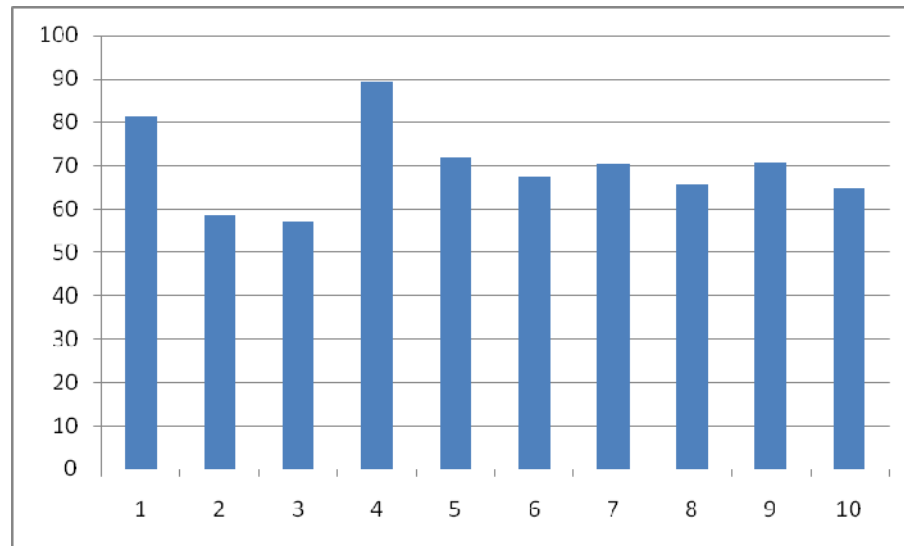
1. Content Knowledge
2. Development and Learning
3. Diversity
4. Critical Thinking
5. Learning Environment and Motivation
6. Communication
7. Planning
8. Assessment
9. Reflection and Continuous Improvement
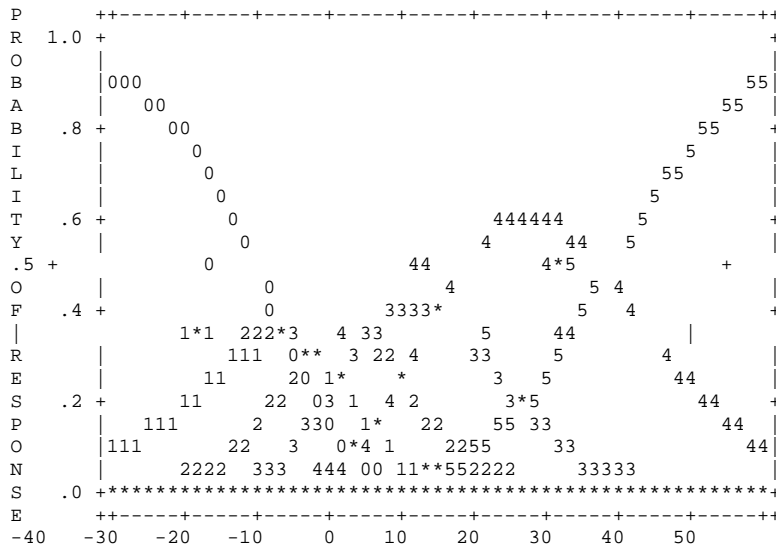10. Collegiality and Professionalism

6.) Was the Krathwohl Taxonomy a useful framework for teacher dispositions with evidence of construct validity present?

```
+-------------------------------------------------------------------------
|CATEGORY    STRUCTURE   |  SCORE-TO-MEASURE   | 50% CUM.| COHERENCE|ESTIM|
| LABEL    MEASURE  S.E. | AT CAT. ----ZONE----|PROBABLTY| M->C C->M|DISCR|
|-----------------------+---------------------+---------+----------+-----+
|   0       NONE         |(-30.43) -INF  -23.81|         | 69%  38%|     | 0
|   1      -15.88   .93  | -16.81 -23.81 -12.04|  -20.29 | 34%  29%|  .83| 1
|   2      -12.28   .75  |  -7.80 -12.04  -3.43|  -11.88 | 38%  72%| 1.37| 2
|   3       -5.12   .73  |   1.68  -3.43   8.32|   -4.08 | 40%  37%| 1.30| 3
|   4        4.94  1.05  |  17.64   8.32  30.51|    6.85 | 25%   1%| 1.15| 4
|   5       28.34  4.13  |( 40.01) 30.51  +INF |   29.21 |  0%   0%| 1.00| 5
+-------------------------------------------------------------------------
M->C = Does Measure imply Category?
C->M = Does Category imply Measure?


CATEGORY PROBABILITIES: MODES - Structure measures at intersections
   P       ++-----+-----+-----+-----+-----+-----+-----+-----+-----++
   R  1.0 +                                                         +
   O       |                                                        |
   B       |000                                                 55|
   A       |   00                                            55   |
   B   .8 +      00                                        55      +
   I       |         0                                    5        |
   L       |           0                                55         |
   I       |            0                              5           |
   T   .6 +              0                   444444      5         +
   Y       |               0                4        44  5         |
   .5 +                0                44          4*5            +
   O       |              0        0               4        5  4   |
   F   .4 +                  0           3333*            5    4   +
   |            1*1   222*3   4 33          5       44         |
   R       |             111   0**   3 22 4      33      5          4        |
   E       |            11     20 1*     *         3  5           44        |
   S   .2 +      11       22   03 1  4 2       3*5              44      +
   P       |  111        2    330  1*   22    55 33            44  |
   O       |111       22     3   0*4 1    2255      33              44|
   N       |      2222   333   444 00 11**552222     33333          |
   S   .0 +***********************************************************+
   E       ++-----+-----+-----+-----+-----+-----+-----+-----+-----++
   -40    -30    -20    -10     0     10    20    30    40    50
```

Notes: The categories here are the five Krathwohl affective stages plus an earlier stage defined as "pre-receiving." Here the data show that the six stages were distinct and correctly ordered across all instruments. A similar pattern was evident on the FACETS analysis of the ETQ with multiple judges rating the same questionnaires. **Yes, this supports construct validity:**

Model = ?,?B,?B,KRATHWOLH (Rating or Partial Credit Scale)
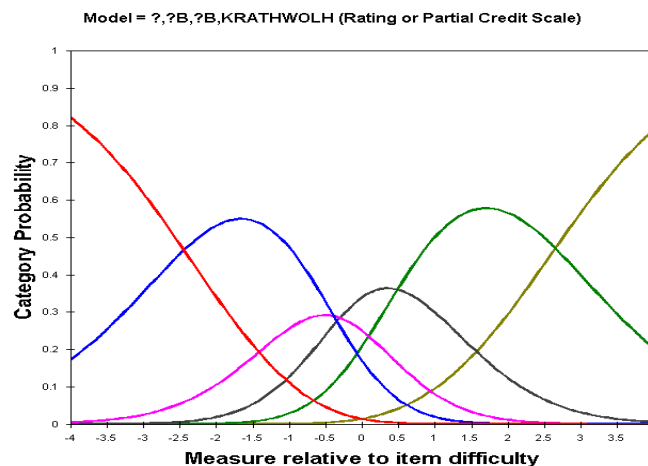
7.) Does the individual judgment of raters differ enough to cause reliability problems?

```
Experiential Teacher Questionnaire  02-01-2008 17:55:27
Table 6.0  All Facet Vertical "Rulers".

Vertical = (1A,2A,3A) Yardstick (columns,lines,low,high)= 60,15,-1,1
-------------------------------------------------------------------------------
|Measr|+candidate      + ETQ Question INTASC                        |+Rater|KRATH|
-------------------------------------------------------------------------------
+   1 +           +                                                 +  (5) +
|     | 17  6     |                                                 |  --- |
|     | 18        |                                                 |      |
|     |           |                                                 |      |
|     | 4         |                                                 |      |
|     |           |                                                 |      |
|     |           |                                                 |      |
|     |           | Learning Environment & Motivation               |      |
|     | 22        |                                                 |    3 |
|     |           |                                                 |      |
|     | 9         | Development & Learning                          |      |
|     |           | Collegiality & Professionalism, Planning        |      |
|     | 21  23  7 |                                              | 1 |      |
|     | 12  3     |                                              | 2 |      |
*   0 * 11        * Diversity                                    * 3 *  --- *
|     |           |                                              | 3 |      |
|     | 19 20  5  8| Reflection & Continuous Improvement             |      |
|     | 14        | Assessment, Critical Thinking                   |      |
|     |           | Communication, Content Knowledge                |      |
|     | 10 15 16  |                                                 |      |
|     |           |                                                 |    2 |
|     |           |                                                 |      |
|     | 13        |                                                 |      |
|     |           |                                                 |      |
|     |           |                                                 |      |
+  -1 +           +                                                 +  (0) +
-------------------------------------------------------------------------------
|Measr|+candidate    |+Question                                     |+Rater|KRATH|
-------------------------------------------------------------------------------
```

> Each judge is calibrated on a ruler of rater leniency on the *same* interval scale that of the measures.

> Each student is calibrated on a ruler of disposition measures of consistency with INTASC. These scores remove the judge bias in the computation.

Notes: This is a FACETS (Rasch model) analysis of the rater effects for the ETQ ratings from one institution. The rulers can be reversed for interpretation. In this case, Students 17 and 6 are the most consistent with INTASC. Meanwhile, Learning Environment & Motivation were the most difficult for students to agree with INTASC on the ETQ questionnaire. Judge number 1 was the "harshest" and had the highest standard for consistency with INTASC as a rater (even though the judges were relatively close). Krathwolh stages range from a rating of 1 (Receiving) to 5 (Characterization). The Rasch model produces a linear transformation of Cohen's Kappa except the expected result is 0 instead of 1. In this case the Rasch Kappa values are:
Judge 1's inter-rater Kappa = .26, Judge 2's inter-rater Kappa = .20, Judge 3's inter-rater Kappa = .28. Since 0 is the expected value, this is a moderate, but not excellent set of values. It is likely that rater training or rubric improvement would result in better consistency. **FACETS analysis of the Rasch model allows instant and easy understanding of rater effects. It is possible to have an excellent inter-rater reliability and fail to show that judges are too lenient or too harsh.**

The Rasch model provides a FAIR AVERAGE for each assessed student without rescoring the questionnaire.   This is produced by removing the rater effects on the scale score.

8.) Is it possible to provide a score for each student on Krathwohl's Taxonomy?

Score report for Joy: 1st term junior, elementary education
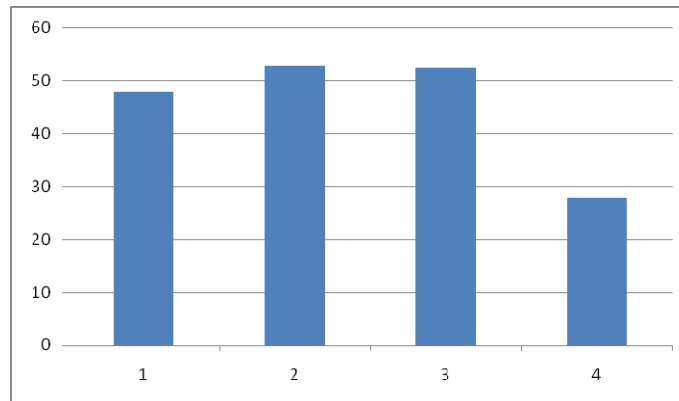
1. Receiving
2. Responding
3. Valuing
4. Organizing
5. Characterizing (not shown)

<u>Notes</u>: **Just like the INTASC principles, the Rasch model can produce calibrated results on the same interval scale on multiple frameworks!** This might include: state requirements, chronological progress, degree program, demographics, or any other identified variable as long as the variable is logical. Notice that Joy (above) is below a typical consistency score (50) is only barely at the responding and valuing level common for Final Internship students in education. Contrast that report with Kelly who is in a Final Internship and is already scoring in the scale range (60's) for Receiving, Responding, and Valuing.
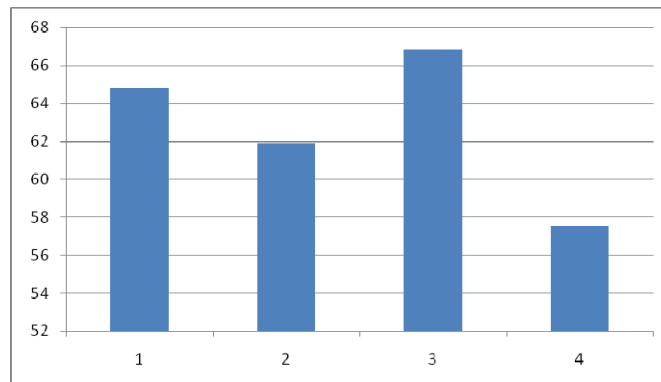
1. Receiving
2. Responding
3. Valuing
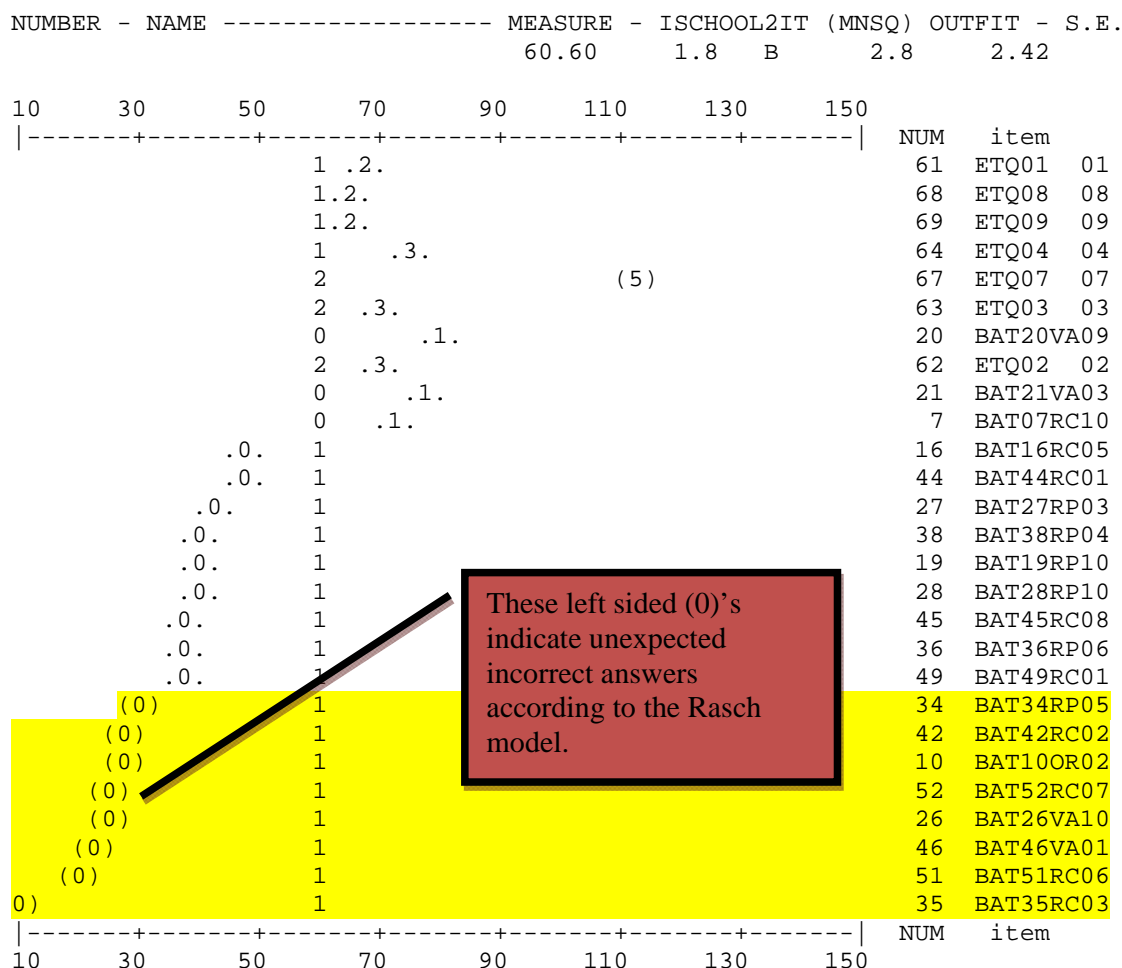4. Organizing
5. Characterizing (not shown)

9.) Can institutions or programs be compared for consistency with INTASC?

```
TABLE 33.1 Disposition Rating Analysis 08          ZOU143ws.txt Feb  1 10:21 2008
INPUT: 228 persons, 90 items  MEASURED: 154 persons, 88 items, 12 CATS      3.57.2
--------------------------------------------------------------------------------
                        Institution differences by Krathwolh
+------------------------------------------------------------------------------+
| person  DIF    DIF   person  DIF    DIF      DIF    JOINT            item     |
| CLASS  MEASURE S.E.  CLASS  MEASURE S.E.  CONTRAST  S.E.   t  d.f. Prob. CLASS|
|------------------------------------------------------------------------------|
| SCHOOL1  .93    .54  SCHOOL2  2.95  1.15     -2.02  1.27  1.60 INF .1098 RECEIVING
| SCHOOL1  .93    .54  SCHOOL3 -6.81  1.20      7.73  1.31 -5.89 INF .0000 RECEIVING
| SCHOOL2 2.95   1.15  SCHOOL3 -6.81  1.20      9.75  1.66 -5.88 INF .0000 RECEIVING
| SCHOOL1 -.49    .90  SCHOOL2  2.88  1.90     -3.36  2.10  1.60 INF .1096 RESPONDING
| SCHOOL1 -.49    .90  SCHOOL3  -.55  1.73       .07  1.95  -.03 INF .9730 RESPONDING
| SCHOOL2 2.88   1.90  SCHOOL3  -.55  1.73      3.43  2.56 -1.34 538 .1816 RESPONDING
| SCHOOL1  .31   1.02  SCHOOL2  3.95  2.08     -3.64  2.32  1.57 INF .1163 VALUING
| SCHOOL1  .31   1.02  SCHOOL3 -5.25  2.14      5.56  2.37 -2.35 INF .0190 VALUING
| SCHOOL2 3.95   2.08  SCHOOL3 -5.25  2.14      9.20  2.98 -3.08 493 .0022 VALUING
| SCHOOL1-4.21   3.67  SCHOOL2  1.93  7.28     -6.14  8.15   .75 232 .4519 ORGANIZING
| SCHOOL1-4.21   3.67  SCHOOL3  7.78  4.21    -11.99  5.59  2.15 230 .0329 ORGANIZING
| SCHOOL2 1.93   7.28  SCHOOL3  7.78  4.21     -5.85  8.41   .70  88 .4886 ORGANIZING

                    Institution differences by INTASC Principle
+--------------------------------------------------------------------------------+
| person    DIF    DIF  person    DIF    DIF      DIF    JOINT             item   |
| CLASS   MEASURE  S.E. CLASS   MEASURE  S.E.  CONTRAST  S.E.   t  d.f. Prob. CLASS|
|--------------------------------------------------------------------------------|
| SCHOOL1    .71    .80 SCHOOL2  -1.89  1.73      2.59  1.90 -1.36 838 .1734 01   |
| SCHOOL1    .71    .80 SCHOOL3  -1.01  1.45      1.72  1.65 -1.04 846 .2991 01   |
| SCHOOL2  -1.89   1.73 SCHOOL3  -1.01  1.45      -.87  2.25   .39 326 .6987 01   |
| SCHOOL1    .06    .86 SCHOOL2   1.51  1.83     -1.45  2.02   .71 843 .4749 02   |
| SCHOOL1    .06    .86 SCHOOL3  -1.26  1.54      1.32  1.76  -.75 851 .4557 02   |
| SCHOOL2   1.51   1.83 SCHOOL3  -1.26  1.54      2.76  2.39 -1.16 328 .2487 02   |
| SCHOOL1   -.53    .81 SCHOOL2   -.39  1.80      -.14  1.97   .07 842 .9453 03   |
| SCHOOL1   -.53    .81 SCHOOL3   1.88  1.45     -2.40  1.66  1.45 851 .1481 03   |
| SCHOOL2   -.39   1.80 SCHOOL3   1.88  1.45     -2.27  2.31   .98 327 .3260 03   |
| SCHOOL1   -.02    .79 SCHOOL2   1.18  1.68     -1.20  1.86   .64 725 .5192 04   |
| SCHOOL1   -.02    .79 SCHOOL3  -1.06  1.61      1.04  1.79  -.58 724 .5627 04   |
| SCHOOL2   1.18   1.68 SCHOOL3  -1.06  1.61      2.23  2.32  -.96 271 .3370 04   |
| SCHOOL1    .13    .78 SCHOOL2    .25  1.70      -.12  1.87   .06 834 .9505 05   |
| SCHOOL1    .13    .78 SCHOOL3   -.73  1.48       .86  1.68  -.51 841 .6077 05   |
| SCHOOL2    .25   1.70 SCHOOL3   -.73  1.48       .98  2.26  -.43 327 .6655 05   |
| SCHOOL1    .04    .81 SCHOOL2   1.13  1.84     -1.09  2.01   .54 833 .5889 06   |
| SCHOOL1    .04    .81 SCHOOL3   -.98  1.54      1.02  1.74  -.59 844 .5564 06   |
| SCHOOL2   1.13   1.84 SCHOOL3   -.98  1.54      2.11  2.40  -.88 325 .3802 06   |
| SCHOOL1    .76    .78 SCHOOL2  -2.84  1.54      3.60  1.72 -2.09 823 .0370 07   |
| SCHOOL1    .76    .78 SCHOOL3   -.18  1.55       .94  1.73  -.54 819 .5884 07   |
| SCHOOL2  -2.84   1.54 SCHOOL3   -.18  1.55     -2.66  2.18  1.22 316 .2236 07   |
| SCHOOL1  -1.07    .77 SCHOOL2   2.78  1.57     -3.85  1.75  2.20 833 .0281 08   |
| SCHOOL1  -1.07    .77 SCHOOL3   1.62  1.52     -2.69  1.70  1.58 841 .1139 08   |
| SCHOOL2   2.78   1.57 SCHOOL3   1.62  1.52      1.16  2.19  -.53 328 .5961 08   |
| SCHOOL1   -.16    .78 SCHOOL2    .31  1.66      -.47  1.83   .26 835 .7961 09   |
| SCHOOL1   -.16    .78 SCHOOL3    .32  1.54      -.48  1.72   .28 844 .7805 09   |
| SCHOOL2    .31   1.66 SCHOOL3    .32  1.54      -.01  2.26   .00 327 .9975 09   |
| SCHOOL1    .02    .76 SCHOOL2  -1.52  1.70      1.54  1.87  -.83 842 .4080 10   |
| SCHOOL1    .02    .76 SCHOOL3   1.03  1.48     -1.01  1.66   .61 849 .5425 10   |
| SCHOOL2  -1.52   1.70 SCHOOL3   1.03  1.48     -2.56  2.25  1.13 327 .2574 10   |
+--------------------------------------------------------------------------------+
```

Notes: Above it appears that three institutions are very similar in consistency with INTASC principles with a few (possibly random-Type I) "significant" differences. On the other hand, on institutions differs from the other two with regard to Krathwolh development at the Receiving, Valuing, and Organizing level. *Whether this is "good" or "bad" is unknown*, but it seems that the type of student, the development of the student, and the differences in groups of student could lead to improvement if faculty are aware of the student dispositions. **Without measures aggregated results, important iSCHOOL2ormation will never be known nor addressed.**

10.) Can the Rasch model identify students who need attention?

```
NUMBER - NAME ----------------- MEASURE - ISCHOOL2IT (MNSQ) OUTFIT - S.E.
                                   60.60    1.8   B    2.8    2.42

10    30    50     70     90    110    130    150
|------+------+------+------+------+------+------+------|  NUM   item
                  1 .2.                                      61   ETQ01  01
                  1.2.                                       68   ETQ08  08
                  1.2.                                       69   ETQ09  09
                  1     .3.                                  64   ETQ04  04
                  2                      (5)                 67   ETQ07  07
                  2  .3.                                     63   ETQ03  03
                  0      .1.                                 20   BAT20VA09
                  2  .3.                                     62   ETQ02  02
                  0      .1.                                 21   BAT21VA03
                  0    .1.                                    7   BAT07RC10
            .0.   1                                          16   BAT16RC05
            .0.   1                                          44   BAT44RC01
          .0.     1                                          27   BAT27RP03
         .0.      1                                          38   BAT38RP04
         .0.      1                                          19   BAT19RP10
         .0.      1                                          28   BAT28RP10
        .0.       1                                          45   BAT45RC08
        .0.       1                                          36   BAT36RP06
       .0.        1                                          49   BAT49RC01
       (0)        1                                          34   BAT34RP05
       (0)        1                                          42   BAT42RC02
       (0)        1                                          10   BAT10OR02
      (0)         1                                          52   BAT52RC07
      (0)         1                                          26   BAT26VA10
     (0)          1                                          46   BAT46VA01
    (0)           1                                          51   BAT51RC06
 0)               1                                          35   BAT35RC03
|------+------+------+------+------+------+------+------|  NUM   item
10    30    50     70     90    110    130    150
```

These left sided (0)'s indicate unexpected incorrect answers according to the Rasch model.

Notes: Here is a student with a "misfitting" score who had a number of unexpected inconsistent answers according to the Rasch model's estimate of the student overall pattern. Here are a subset of the unexpectedly inconsistent responses. Is it possible that this student has a sensitivity to backgrounds different from their own? **This pattern would likely never be observed without a statistical analysis of calibrated assessment items!**

- I really enjoy meeting the parents and family of my students and seeing them inadvertently outside of the school setting.

- Teacher's who express their personal values should know that it makes some students with different backgrounds uncomfortable.

- I constantly ask students to describe their point of view so that I can understand their different perspectives.

Conclusions and Recommendations

       This presentation represents the proverbial tip of the iceberg. With carefully constructed instruments and data analysis, disposition assessment has the possibility of iSCHOOL2orming teacher education in ways are self-correcting for students, program revealing for faculty, institutionally meaningful for accreditation, and job-related predictions for teacher certification.

References

American Educational Research Association, American Psychological Association, and National Council of Measurement in Education (1999). *Standards for educational and psychological testing.*

Bond, T. & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nd).* Mahwah, NJ: LEA.

Council of Chief State School Officers (1992). *Model Standards for Beginning Teacher Licensing, Assessment, and Development: a resource for state dialogue.* Washington, D.C.: Author. Retrieved April 24, 2005, from http://www.ccsso.org/projects/Interstate_New_Teacher_Assessment_and_Support_Consortium/Publications

Diez, M. (2008). Assessing dispositions: Context and questions. In Diez, M. & Raths, J. (Eds.) *Dispositions in teacher education.* Charlotte, NC: IAP.

Hays, (1988). *Statistics.* (4th ed.). Forth Worth: Holt, Rinehart and Winston.

Krathwohl, D., Bloom, B. & Masia, B. (1956). *Taxonomy of educational objectives. Handbook II: affective domain.* New York, McKay.

Linacre, J. M. (2003). *A user's guide to FACETS: Rasch-model computer programs.* MESA Press: Chicago.

Linacre, J. M. (2003). *A user's guide to WINSTEPS: Rasch-model computer programs.* MESA Press: Chicago.

Smith, E. & Smith, R., Eds. (2004). *Introduction to Rasch Measurement.* Maple Grove, MN: JAM.

Thurstone, L.L. (1928). Attitudes can be measured. *American Journal of Sociology.* 33, 529-554.

Wilkerson, J. & Lang, W. S. (2007). *Assessing teacher competency: Five Standards-based steps to valid measurements using the CAATS model.* Thousand Oaks, CA: Corwin.

Wilson, M. (2005). *Constructing measures: An item response modelling approach.* Mahwah, NJ: LEA.