# Construct Validity, Reliability and Cross-Cultural Equivalence of the LSP's Cognitive Dimension: A Rasch Analysis

Patrick C. Hardigan, Ph.D.
Executive Director of Assessment
Health Profession Division
Nova Southeastern University
3200 South University Drive
Ft. Lauderdale, Fl 33328
954-262-1524
954-262-2252 (FAX)
patrick@nova.edu

Date Completed January 12, 2010

**Abstract**

The Learning Style Profile (LSP) is one of the first assessment instruments that measures cognitive skills as well as affective and environmental preferences. Developed for use with sixth-through twelfth-grade learners, the LSP consists of 126 questions representing 24 sub-scales These 24 subscales represent four higher-order constructs: (1) Cognitive Style, (2) Perceptual Preferences, (3) Study Preferences, and (4) Instructional Preferences. Although the LSP was not designed for college students, research shows that it is being used with college-aged students. Past studies which looked at the LSP with college students were exploratory in nature and focused on the perceptual preferences, study preferences, and instructional preference subscales. No analysis has explicitly looked at LSP's cognitive dimensions and its cross-cultural equivalence among traditional-aged college students. The purpose of this study was to evaluate the construct validity, reliability and cross-cultural equivalence of the LSP's cognitive dimension among undergraduate college students using Rasch techniques. Results show that in most cases the number of acceptable items is not enough to cover each subscale's content domain. Where we do have enough items, the subscales suffer from content saturation. While we have evidence for unidimensionality and no differential item functioning we conclude that the LSP should not be used with undergraduate college students.

## Introduction

One of the most reliable and effective means of improving a student's academic performance involves the use of cognitive style assessment tools. The usefulness of the cognitive style assessment concept and various diagnostic approaches has been demonstrated in terms of student achievement, the inhibiting of dropout rates, and increasing students' satisfaction with instruction (Cross, 1983). Robert Smith (1983) states:

> *...knowledge of one's preferences and tendencies in learning and information processing can be helpful in making choices about what, when, where, and how to learn; it can also help pinpoint personal difficulties or explain problems with particular subjects, methods, or instructors. (p. 98).*

The Learning Style Profile (LSP), is one of the first assessment instruments which measures cognitive skills as well as affective and environmental preferences. The LSP adopted Charles Letteri's General Operations Model as the prototype for relating learning styles to cognitive information processing. The LSP was developed by the National Association of Secondary School Principals because of limited instrumentation to measure the wide variety of learning and cognitive styles (Melear, 1989). Developed for use with sixth-through twelfth-grade learners, the LSP consists of 126 questions representing 24-four sub-scales: analytic, spatial, discrimination, categorizing, sequential processing, and memory skill; visual, auditory, and emotive perceptual response; persistence and verbal risk orientation; verbal-spatial, manipulative, grouping, posture, mobility, sound, lighting, and temperature preference; and study time preference (early morning, late morning, afternoon, or evening). These

24 subscales represent four higher-order constructs: (1) Cognitive Style, (2) Perceptual Preferences, (3) Study Preferences, and (4) Instructional Preferences.

The Learning Style Profile, originally developed for sixth-through twelfth-grade students, is being used with students beyond the intended group. For example, the Profile has been used as part of a study examining cognitive processes among non-majors in college biology, the LSP has been used with undergraduate agriculture majors, educators use it a part of a training program with adult students, and it is offered online to community college students to assess their learning and cognitive style preferences. (Melear, 1989; Boyd & Murphery, 2004; James and Maher (2007), Keeley, 2008).

Although the LSP was not designed for college students, research shows that its strongest reliabilities are of those variables adopted from the Dunn and Dunn model and represent the perceptual, study and instructional preferences (Melear, 1989, Hardigan, 2000). The LSP's weakest measure are those based on cognitive dimensions (Melear 1989, Hardigan 2000). Past studies which looked at the LSP with college students were exploratory in nature and focused on the perceptual preferences, study preferences, and instructional preference subscales. No analysis has explicitly looked at LSP's cognitive dimensions (it's potentially weakest dimension) and its cross-cultural equivalence among traditional-aged college students. Therefore the purpose of this study is to evaluate the construct validity, reliability and cross-cultural equivalence of the Learning Style Profile's Cognitive dimension among undergraduate college students using Rasch techniques.

The LSP's cognitive dimension is composed of seven subscales:

*Analytic Skill*

The analytic skill subscale (AS) is modeled after the Embedded Figures Test (EFT) developed by Herman Witkin (Keefe & Monk, 1988). Five questions asking the student to locate a simple figure hidden within a larger, more complex form, were created to test this skill. This method is similar to the group version of the Embedded Figure Test (GEFT). Field-dependent people are strongly influenced by the surrounding field while field independents are capable of distinguishing between the parts of the field.

*Spatial Skill*

The spatial skills subscale includes two components of general spatial reasoning: (1) pattern recognition-asks the subject to identify identical but different-sized designs within larger identical frameworks, and (2) spatial rotation-asks the subject to identify a two dimensional shape as a three dimensional object. Five items on the LSP perform this function. Spatial ability is defined as the ability to form and use visual spatial images (Slater, 1940). The two components of spatial ability-pattern recognition and spatial rotation-have been assessed and validated by a number of researchers (Smith, 1964). Pattern recognition is an element of a broader spatial ability called orientation (Keefe & Monk, 1988). Spatial rotation represents a subject's capacity to manipulate and/or twist objects.

*Discrimination Skill*

The discrimination skill subscale is based on the focusing control conceptualization (Keefe & Monk, 1988). The subscale measures a student's ability to focus on the important elements of the task. Five items on the LSP measure this construct. Subjects are presented with five circles and a sample-circle. Subjects are asked to compare the size of the five circles against the size of the sample. Two ideas form the basis of the discrimination construct: (1) narrowing of attention and (2) separating information (Schlesinger, 1954). Narrowing of attention refers to size estimation. Bruner & Rodrigues (1953) examined the intricacies of size estimation based on placement of shapes, the effects of light, and possible attitudinal differences. They discovered that the attitudinal or psychological differences in perception had the greatest influence on size estimation. The second idea, separating of information, explored the concept of scanning as opposed to focusing. The scanner is easily distracted and inclined to concentrate on irrelevant details while the focuser is capable of ignoring irrelevant detail and focusing on the task at hand (Keefe & Monk, 1988).

*Categorization Skill*

The categorization skill subscale is based on the notion of 'equivalence range' (Keefe & Monk, 1988). Equivalence range can be subdivided into groups: (1) narrow categorizer who employed a detailed scheme to sort objects into groups with that represented accurate groups, and (2) broad categorizers who used less exact criteria to sort groups which resulted in broader categories. Eight question on the LSP measure this construct. Subjects are given a series of category averages and are asked to estimate the largest and smallest members of each category (Keefe & Monk, 1988). Only averages are presented so that subjects must make decisions based on incomplete data. Gardner (1953) first studied categorization or 'equivalence range.' He performed a study where subjects were (1) asked to sort objects based on size, (2) to match shapes and (3) equate the brightness of patches of light. He discovered that persons are characterized by consistent differences in what they will accept as similar or identical in a variety of tasks. Gardner (1953) hypothesized that it is important for narrow categorizers to know the world around them in terms of reducible, classifiable features. Broad categorizers' knowledge of the exact nature of the outer world is relatively less important as they rely more on feelings in the process of reality testing. Narrow categorizers are better equipped for academic assignments which require complete, accurate and organized categories of easily accessible information (Keefe & Monk, 1988).

*Sequential and Simultaneous Processing*

Sequential and simultaneous processing are the ways humans process information. Sequential processing is defined as the processing of information in serial order (Das, Kirby & Jarman, 1979). Simultaneous processing is the synthesis of separate elements into groups. The NASSP developed separate tests to measure sequential and simultaneous processing. Six questions were developed for the sequential processing scale. The sequential processing subscale asks the subject to determine whether a series of geometric shapes is present or missing in a set of simple puzzles (Keefe & Monk, 1988). Five questions were developed to measure the simultaneous subscale. Here, students are given a form and four possible variations of that form. Students are asked which one of the four variations comes from the original. It is important to understand that sequential and simultaneous processing refers to the ways individuals work mental tasks and not the nature of the tasks themselves (Keefe & Monk, 1988).

*Memory Skill*

The memory skill subscale is a variation on a series of tests designed to assess the cognitive control of leveling versus sharpening (Keefe & Monk, 1988). Sharpening refers to a propensity to maximize differences, an attunement to small gradients of difference between figure and ground (Holzman, 1954). Sharpeners prefer complex to simple organization, and if required, can sustain an organization intact over time. Levelers organize a field in a simple or diffuse manner (Holzman, 1954). Levelers do not easily sustain a single organization over time; familiar fades unless continually affirmed by external sources (Holzman, 1954). Eighteen items on the LSP were developed to measure memory skill. Subjects are given a series of figures on separate pages, and are asked to determine whether each succeeding figure is identical to the preceding one.

## Method

*Sample*

The LSP was administered to a convenience sample of eight hundred and sixty–two undergraduate college students. One hundred and twenty–three were removed from the data analysis. Reasons included: (1) missing or incomplete data (66), (2) student is of graduate standing or student had completed requirements for a bachelor degree (16), student possessed a learning disability (11), or student was not enrolled in six or more credit hours (30). Therefore, seven hundred and thirty–seven students formed the sample for all statistical analysis.

*Statistical Analysis*

To assess the validity and unidimensionality of the subscales Analytic Skill (AS), Spatial Skill (SS), Discrimination Skill (DS), Sequential Processing (SQPS), Simultaneous Processing (SMPS), and Memory Skill (MS) we examined the first-order statistics R1c and the Anderson Likelihood Ratio Test, the Infit mnsq and the Outfit mnsq measures. The six subscales were analyzed using the basic Rasch model for dichotomous items. The basic Rasch model considers a group of *n* candidates, who have an ability, $C_i$, ($i=1$,n), and who each takes a set of *m* tests, each of which has a difficulty $T_j$($j=1,m$). The probability of candidate *i* correctly answering test *j*, *Pij*,is estimated as:

$$logit\,(P_{ij}) = log(Pij/(1-P_{ij}))=C_i - T_j \quad ...$$

To assess the validity and unidimensionality of the subscale Categorization Skill (CS), where points are given for partially correct items, we used full-information factor analysis and examined the Infit MNSQ and the Outfit MNSQ measures. To help ensure that response categories were being used appropriately we followed the guidelines given by Linacre (1997, 2002). The CS subscale was analyzed using the Rasch Partial Credit Model given as.

$$logit\,(P_{ijk}) = log(P_{ijk}/(1-P_{ijk}))=C_i - T_j - M_k \; ...$$

where $P_{ijk}$ is the probability of candidate *i* on test *j* receiving a mark of *k*. The partial credit model allows the differences between the various points on a mark scale to be assessed.

To assess the subscales' reliability we calculated Rasch Person Separation Reliability measures (RR). In Rasch the reliability coefficient for person (person separation) is analogous to Cronbach's alpha. High person separation reliability means that apparent differences between people on the measure are less likely a result of errors in measurement and increases confidence that people who get different "scores" on the measure have separate abilities (Allen, 2007). Low values on any of these forms of reliability assessment indicate larger measurement errors and thus less confidence that comparisons can reveal true differences (Allen, 2007).

To examine cross-cultural equivalence across five different ethnic groups, (1) Asian, (2) Non-Hispanic Black, (3) Hispanic, (4) Non-Hispanic White, and (5) Other, we conducted differential item functioning (DIF). All data analyses were conducted using WINSTEPS and STATA (Linacre & Wright, 2000).

**Results**

The mean age of subjects was 23, with the oldest being 58 and the youngest 17. All subjects were enrolled in traditional daytime courses. Over 60 percent of the subjects were female (60.6%) and the ethnic group “White” (59.8%) constituted the majority of subjects who completed the LSP. Table 1 contains a distribution of subjects by ethnicity/race.

Unidimensionality and Validity

The fit of the subscales Analytic Skill, Spatial Skill, Discrimination Skill, Sequential Processing, Simultaneous Processing, and Memory Skill to the Rasch model were evaluated using the R1c statistic and the Anderson Likelihood Ratio Test. Here a P value > 0.05 indicates that the observed data have a satisfactory fit to the additive model. Item fit (validity) was analyzed using Rasch Outfit and Infit values. Outfit are outlier-sensitive fit statistics. This is based on the conventional chi-square statistic. This is more sensitive to unexpected observations by persons on items that are relatively very easy or very hard for them (Linacre, 2008a). Infit statistics are based on the chi-square statistic with each observation weighted by its statistical information (model variance). This is more sensitive to unexpected patterns of observations by persons on items that are roughly targeted on them (Linacre, 2008). The rule of thumb is to exclude in the analysis the persons and items that have Infit or Outfit statistics of less than -2 or greater than +2. Subscales were also examined for construct deficiency and/or construct saturation. Construct deficiency (CD) represents "gaps" on the construct continuum. These "gaps" represent the points at which the construct is poorly defined by the items (Schulz, 1995, Linacre, 2008a). Construct saturation (CS) is over-representation by similar items at a specific logit value. This is defined more fully as the point on the construct continuum where several items are measuring the same thing in almost the same way (Schulz, 1995, Linacre, 2008b).

*Analytic Skill (AS)*

Results from the R1c statistic and the Anderson Likelihood Ratio Test indicate that the Analytic Subscale did not meet the Rasch model requirements. A review of the item and person Outfit and Infit statistics revealed that two questions (items 2 and 4), and 248 people possessed unacceptable measures (less than -2 or greater than +2) (Table 2). We removed questions two and four, deleted the misfitting 277 people and recalculated the Rasch model. The modified subscale fit the Rasch model assumption, R1c and Anderson test *p*= .06, and the item and person Outfit and Infit statistics were within the acceptable range (Table 3). However, a plot of the item and person scores shows us that the analytic subscale possessed both construct deficiency (CD) and construct saturation (CS) (Figure 1). Item one is very difficult while items three and five are easy and overlap.

*Sequential Processing Skills (SQPS)*

Using the reduced sample (N=460) the SQPS subscale, as measured by the R1c statistic and the Anderson Likelihood Ratio Test, did not fit the unidimensional requirement ($p$ = .044) (Table 2). A review of the Outfit and Infit statistics revealed that item three exceeded our criteria of acceptance (less than -2 or greater than +2). We removed question three and recalculated the R1c statistic and the Anderson Likelihood Ratio Test, Infit and Outfit statistics (Table 3). The results were encouraging, both unidimensional tests were was nonsignificant ($p$=.39 and $p$=.46) and Infit and Outfit statistics were within an “acceptable” range. We then plotted the item and person scores (Figure 1). It appears that the subscale holds adequate content coverage (non-CD) but some content saturation.

*Simultaneous Processing Skills (SMPS)*

The R1c statistic and the Anderson Likelihood Ratio Test were non-significant ($p$=.23 and p=.21) indicating that the subscale met the RASCH unidimensional requirement

(N=460). No items possessed unacceptable Infit and Outfit statistics (Table 2). Similar to the SQPS subscale, when we plotted the items and people logit scores the items provided adequate content coverage and a little content saturation (Figure 1).

*Spatial Skill (SS)*

The R1c statistic and the Anderson Likelihood Ratio Test were significant *p*=.000 thereby demonstrating that the Spatial Skill subscale did not meet the unidimensional requirement (N=460) (Table 2). Two items, questions two and three, possessed unacceptable Infit and Outfit statistics. We removed the two questions and reanalyzed the subscale (Table 3). Both the R1c statistic and the Anderson Likelihood Ratio Test remained significant *p*=.000. Infit and Outfit statistics were also unacceptable as all remaining items possessed Infit and Outfit statistics well-beyond the acceptable criteria. The item-person plot also demonstrates that the Spatial Skill subscale suffers from both CD and CS (Figure 1). Essentially, the analysis shows us that the Spatial Skill subscale, as written, is unacceptable for College students.

*Discriminant Skill (DS)*

The R1c statistic and the Anderson Likelihood Ratio Test were significant *p*=.01 thereby demonstrating that the Discriminant Skill subscale did not meet the unidimensional requirement (N=460). One item possessed unsatisfactory Infit and Outfit statistics--question five (Table 2). We removed the question and reanalyzed the subscale. The R1c statistic and the Anderson Likelihood Ratio analyses improved dramatically with both measures non-significant. Additionally, all remaining items were within satisfactory Infit and Outfit ranges (Table 3). The person and item logit plot shows fairly good coverage for the domain and some evidence of content saturation (Figure 1).

*Memory Skill (MS)*

The R1c statistic and the Anderson Likelihood Ratio Test were significant *p*=.000 (N=460). Three items (questions 10,11,16) held unacceptable Infit and Outfit statistics (Table 2). We reanalyzed the subscale without the six items. The R1c statistic and the Anderson Likelihood Ratio Test were not significant (*p*=.10, *p*=.06) and the remaining eight items held acceptable Infit and Outfit statistics (Table 3). The resulting item and person plot shows poor content coverage and some evidence of content saturation (Figure 1).

*Categorization Skill (CT)*

We analyzed the categorization subscale by means of a Rasch partial credit model (N=460). First, we conducted a Rasch factor analysis. The first factor or dimension accounted for 68 percent of the variance which is acceptable evidence of a unidimensional construct. Next we followed Linacre's suggested criteria for analysis of rating scales.

1. A minimum of ten observations per category.
2. Second the shape of the probability curves should be peaked for each category
3. Third, the average category measures should increase with the rating scale categories
4. Outfit mean-square statistics should be less than 2.0
5. Threshold calibrations should increase with the rating scale category
6. The category thresholds should be at least 1.4 logits apart and no more than five logits apart.

Table 4 provides a summary of category structure. All categories had observed counts over ten. The probability shapes for each curves are peaked at each category--Figure 2. Category measures increased with each scale and the Outfit and Infit statistics for each category and item are less than 2.0 (Table 2). Last threshold calibrations increased with each category and were generally 1.4 logits apart. A review of item Infit and Outfit statistics reveals that all items were within the acceptable parameters (Table 2). Although it appears that the CS subscale may contain some evidence of construct saturation, these results suggest that the Categorization subscale possesses acceptable psychometric properties.

Reliability

For the LSP, the Rasch analysis was able to provide a person-separation index (reliability) for only two subscales (N=460): (1) Categorization Skill = .68 and (2) Memory Skill = .57. All other subscale's person-separation indices were zero. The results provide evidence that for AS, SQPS, SMPS, DS, AND SS subscales any differences among ability estimates of persons, and therefore among any order that may be observed, are no greater than would be expected by chance relative to the error of measurement (Andrich, 2008). Apparently these five subscales lack enough questions or items to adequately separate test-takers.

Cross-Cultural Equivalence

Using the reduced sample and "adjusted" subscales we calculated 1-way ANOVAs to compare the mean scores for five different ethic groups (Black, Asian, White, Hispanic, and Other) on the seven subscales . Results show us that differences exist on four of the seven subscales: (1) analytic skill sequential processing, (2) simultaneous processing, and (4) spatial skill. Using Tukey's HSD test we find that generally Asian and White students tended to score the highest, while Blacks and Hispanics the lowest (Table 5).

To determine whether observed differences between Black, Asian, White, Hispanic, and Other students could be explained by non-equivalence of the items, we conducted differential item functioning analysis (DIF). DIF examines items to see if they display different properties for different groups after controlling for person ability (Angoff, 1993). For the six subscales measured with dichotomous variables (AS, SQPS, SMPS, DS, SS, and MS) we examined Differential Item Functioning (DIF) by likelihood ratio tests. The statistic of the test follows a chi-square distribution under the null assumption. For the Categorization subscale differences of item difficulties are tested for significance by dividing the difference by the joint standard errors of the items (Choi, Mericle, Harachi, 2006). The calculated *t-statistic* is used to detect items biased against subgroups. We consider items to display significant DIF if the difference in item difficulty between the baseline measure and each group creates a *t* statistic greater than +/- 1.96. Results show that none of subscales or questions displayed differential item functioning (Table 3).

## Discussion

This study examined the validity, reliability and cross-culture item equivalence of the Learning Style Profile's Cognitive Dimension using Rasch analysis. The Rasch model is arguably the most appropriate model to test dimensionality of dichotomously scored items (Vigneau, 2005). We were able to demonstrate unidimensionality for six of the seven subscales, after removing items and persons, but item difficulty and person separation varied across the dimensions

The Memory Skill and Categorization Skill subscales possessed the best psychometric properties.  Both possess an acceptable range of item difficulty, some degree of content saturation, and adequate reliability measures.  The subscales Discrimination Skill, Sequential Processing Skill, and Simultaneous Processing Skill hold an acceptable range of items, some content saturation, but unacceptable reliability. The Analytic Skill and Spatial Skill subscales hold unacceptable range of items, enough items, or reliable items.

In the presence of SMCDs and CMCDs, there are five steps recommended below as a possible solution (Linacre, 2008a; Linacre 2008b):

Step 1: Identification of any clinically or statistically meaningful gaps or redundancies in the continuum.

Step 2: Determine the number of items needed to fill each gap.

Step 3: Formulation of new items by a committee composed of experts.

Step 4: Testing of new and old items.

Step 5: Calibration of new items along the anchored continuum of the previous items.

It conclusion it appears that the Learning Style Profile's Cognitive Dimension is not psychometrically sound.  Although the LSP was written as a first-level diagnostic instrument, in most cases the number of acceptable items is not enough to cover each subscale's content domain.  Where we do have enough items, the subscales suffer from content saturation. While he have evidence for unidimensionality and no differential item functioning we conclude that the Learning Style Profile should not be used with undergraduate college students. More items should be added that possess adequate psychometric properties and which are relevant to college-aged students.

## References

Allen, D.D. (2007). Validity and reliability of the movement ability measure: A self-report instrument proposed for assessing movement across diagnoses and ability levels. Physical Therapy, 87(7), 899-916.

Angoff W.H. (1993) Perspectives on differential item functioning methodology. In: Differential Item Functioning. (Holland PW, Wainer H, eds) Hillsdale, NJ: Lawrence Erlbaum.

Andrich, D. (2008). An Index of Person Separation in Latent Trait Theory, the Traditional KR-20 Index, and the Guttman Scale Response Pattern. www.rasch.org/erp7.htm

Boyd, B.L. & Murphrey, T.P (2004). Evaluating the scope of learning style instruments used in studies published in the Journal of Agricultural Education. Journal of Southern Agricultural Education Research, 54(1), 124-133.

Bruner, J.S. & Rodrigues, J.S. (1953). Some determinants of apparent size. The Journal of Abnormal and Social Psychology, 48(1), 17–24.

Choi, Y., Mericle, A., & Harachi, T. W. (2006). Using a Rasch Model to Test the Cross-Cultural Item Equivalence ofthe Harvard Trauma Questionnaire and the Hopkins Symptom Checklist Across Vietnamese and Cambodian Immigrant Mothers. Journal of Applied Measurement, 7(1), 16-38.

Cross, P. (1983). Helping adults learn to learn. In R. Smith (Ed.), New Directions for continuing education. (Vol. 19, pp. 41–50). San Francisco: Jossey-Bass

Das, J.P., Kirby, J.R., & Jarman, R.F. (1979). Simultaneous and successive process. New York: Academic press.

Dunn, R, Dunn, K. (1986). Teaching students through their individual learning styles: A practical approach. Reston, VA: Reston Publishing Co.

Gardner, R.W. (1953). Cognitive styles in categorizing behavior. Journal of Personality, 22, 214-333.

Hardigan, P.C. (2000). Factor Validity Study of the Learning Style Profile. (ERIC Document Reproduction Service No. ED 449748).

Holzman, P.S. (1954). The relationship of assimilation tendencies in visual, auditory, and kinesthetic time–error to cognitive attitudes of leveling and sharpening. Journal of Personality, 22, 375–394.

Jame, W.B., & Maher, P.A. (2007). Are you Teaching with “Style”? Resource Packet Appendix 24/7 November 12-26, 2007 www.starlinktraining.org.

Keefe, J.W. & Monk, J.S. (1988). Learning style examiner's manual. Reston, VA: National Association of Secondary School Principals.

Keeley, M. (2008). Special Populations Office, Bucks County Community College. http://www.bucks.edu/~specpop/Lrnprfil.htm. Accessed March 2008.

Linacre, J.M. & Wright, B.D. (2000). WINSTEPS computer program for Rasch measurement. Chicago:MESA Press.

Linacre, J.M. (1999). Investigating rating scale category utility. Journal of Outcome Measurement, 3, 103-122.

Linacre, J.M. (2002). Optomizing rating scale category effectiveness. Journal of Applied Measurement, 3(1), 85-106.

Linacre, J.M. (2008a). Assessing Statistically and Clinically Meaningful Construct Defiency/Saturation: Recommended Criteria for Content Coverage and Item Writing. http://www.rasch.org/rmt/rmt174d.htm.

Linacre, J.M. (2008b). Misfit diagnosis: Infit Outfit mean-square standardized. Winsteps Help for Rasch Analysis. http://www.winsteps.com/winman/index.htm?diagnosingmisfit.htm.

Melear, C.T. (1989). Cognitive processes in the Curry Learning Style Framework as Measured by the Learning Style Profile and the Myers-Briggs Type Indicator among non-Majors in College Biology. Ph.D., Dissertation, The Ohio State University. (ERIC Document Reproduction Service No. ED 313 240).

Schlesinger H.J. (1954). Cognitive attitudes in relation to susceptibility to interference, Journal of Personality, 22, 354–374.

Schulz E. M. (1995) Construct deficiency? RMT 9(3), p. 447.

Slater, P. (1940). Some group tests of spatial judgments on practical ability. New York: Praeger Press.

Smith, I.M. (1964). Spatial ability. San Diego, CA: Knapp

Smith, R.M. (1983). Helping adults learn to learn. In R. Smith (Ed.) New Directions for continuing education. (Vol. 19, pp. 88–94). San Francisco: Jossey-Bass.

Vignaeau, F. (2005). Items in Context: Assessing the Dimensionality of Raven's Advanced Progressive Matrices. Educational and Psychological Measurement, 65(1) 109-123.

Table 1: Distribution of Subjects by Race

| Ethnicity | Count | Relative Frequency | Cumulative Frequency |
|---|---|---|---|
| Asian | 65 | 9% | 9% |
| Black | 70 | 10% | 19% |
| Hispanic | 123 | 17% | 36% |
| White | 442 | 60% | 96% |
| Other | 37 | 4% | 100% |
| Total | 737 | 100% | 100% |

Table 2: Item Statistics Before Dropping Items (N=764)

| Subscale | Item | Difficulty | Error | Outfit | Infit | p-value |
|---|---|---|---|---|---|---|
| Analytic Skill | 1 | 1.42 | 0.09 | 1.03 | -1.24 | 0.00 |
| R1c = 118.43 p =.000 | 2 | 0.71 | 0.08 | 7.32 | 7.44 | 0.00 |
| Anderson LR Test Z=103.22 p=.000 | 3 | -0.85 | 0.09 | -2.15 | -1.68 | 0.07 |
| | 4 | -0.44 | 0.08 | -4.54 | -4.36 | 0.01 |
| | 5 | -0.84 | 0.09 | -1.61 | -1.90 | 0.16 |
| Sequential Processing | 1 | -1.07 | 0.19 | -1.80 | -0.86 | -0.29 |
| R1c = 31.85 p =.044 | 2 | -0.40 | 0.16 | -1.19 | -1.33 | 0.57 |
| Anderson LR Test Z=31.78, p=.045 | 3 | 0.48 | 0.13 | 1.99 | 2.00 | 0.03 |
| | 4 | -0.22 | 0.15 | -1.08 | -0.95 | 0.82 |
| | 5 | 0.89 | 0.12 | 1.57 | 1.52 | 0.06 |
| | 6 | 0.33 | 0.13 | 0.60 | 0.47 | 0.53 |
| Simultaneous Processing | 1 | 0.15 | 0.14 | 1.73 | 1.52 | 0.17 |
| R1c = 15.11 p = .23 | 2 | 0.77 | 0.13 | -0.57 | -0.81 | 0.91 |
| Anderson LR Test Z=15.56 p=.21 | 3 | 0.22 | 0.14 | -1.22 | -1.21 | 0.67 |
| | 4 | 0.03 | 0.15 | 1.50 | 1.59 | 0.29 |
| | 5 | -1.18 | 0.19 | -1.15 | -1.15 | 0.43 |
| Spatial Skill | 1 | -0.36 | 0.08 | 1.01 | 1.34 | 0.48 |
| R1c = 32.76 p = .001 | 2 | 0.35 | 0.07 | -2.34 | -2.68 | 0.01 |
| Anderson LR Test Z=33.55 p=.001 | 3 | 1.10 | 0.07 | 2.61 | 3.30 | 0.00 |
| | 4 | -0.97 | 0.09 | -0.32 | -0.59 | 0.43 |
| | 5 | -0.11 | 0.07 | -1.90 | -1.87 | 0.54 |

| Subscale | Item | Difficulty | Error | Outfit | Infit | p-value |
|---|---|---|---|---|---|---|
| Discriminant Skill | 1 | 0.82 | 0.07 | -1.59 | -1.56 | 0.46 |
| R1c = 32.39 p = .001 | 2 | -1.11 | 0.09 | 1.18 | 0.87 | 0.15 |
| Anderson LR Test Z=33.55 p=.000 | 3 | -0.30 | 0.08 | 1.97 | 1.75 | 0.40 |
| | 4 | 0.62 | 0.07 | 0.77 | 1.05 | 0.38 |
| | 5 | 0.00 | 0.07 | -3.22 | -3.32 | 0.00 |
| Memory | 1 | 0.58 | 0.07 | 3.76 | 4.33 | 0.00 |
| R1c = 160.24 p = .001 | 2 | 0.69 | 0.07 | 1.00 | 0.87 | 0.35 |
| Anderson LR Test Z=163.25 p=.000 | 4 | 0.14 | 0.07 | -2.09 | -2.11 | 0.72 |
| | 6 | -0.45 | 0.07 | 0.47 | 0.64 | 0.97 |
| | 8 | 0.72 | 0.07 | 0.60 | 0.81 | 0.57 |
| | 10 | 0.00 | 0.07 | -3.45 | -3.22 | 0.12 |
| | 11 | -0.59 | 0.07 | 3.85 | 3.57 | 0.00 |
| | 12 | -0.12 | 0.07 | 0.49 | 0.20 | 0.15 |
| | 13 | 0.54 | 0.07 | -2.12 | -1.81 | 0.56 |
| | 15 | 0.52 | 0.07 | -0.50 | -0.34 | 0.98 |
| | 16 | -1.08 | 0.08 | -2.26 | -2.12 | 0.04 |
| | 18 | -0.93 | 0.08 | 0.42 | -1.08 | 0.63 |
| Categorization | 1 | 0.87 | 0.06 | 0.91 | 0.87 | |
| Variability accounted = 68% | 2 | -0.55 | 0.10 | 1.14 | 1.18 | |
| | 3 | 0.31 | 0.10 | 1.10 | 1.19 | |
| | 4 | -0.30 | 0.09 | 1.08 | 1.12 | |
| | 5 | 0.13 | 0.09 | 0.96 | 1.00 | |
| | 6 | 0.21 | 0.10 | 0.97 | 0.92 | |

| Subscale | Item | Difficulty | Error | Outfit | Infit | p-value |
|---|---|---|---|---|---|---|
| | 7 | 0.38 | 0.10 | 0.82 | 0.89 | |
| | 8 | 0.47 | 0.10 | 0.87 | 0.96 | |

*The Memory subscale contains 18 items but only 12 items are used to calculate a student's score

Table 3: Item Statistics After Dropping Items (N=460)

| SubScale | Item | Difficulty | Error | OUTFIT | INFIT | P Value |
|---|---|---|---|---|---|---|
| Analytic Skill | 1 | 1.68 | 0.11 | 1.26 | 0.18 | 0.03 |
| R1c =5.49 p = .064 | 3 | -0.84 | 0.09 | -0.85 | -0.83 | 0.51 |
| Anderson LR Test Z=4.65 p=.097 | 5 | -0.83 | 0.09 | -0.62 | -0.60 | 0.68 |
| DIF LR = 8.12, *p* = .423 | | | | | | |
| Sequential Processing | 1 | -1.01 | 0.19 | -0.55 | -0.32 | 0.96 |
| R1c = 12.69 p =.391 | 2 | -0.33 | 0.16 | -0.78 | -0.92 | 0.76 |
| Anderson LR Test Z=11.79 p=.462 | 4 | -0.13 | 0.16 | 0.01 | -0.22 | 0.53 |
| DIF LR = 16.00, *p* = .453 | 5 | 1.03 | 0.13 | 1.45 | 1.59 | 0.05 |
| | 6 | 0.44 | 0.14 | 0.30 | 0.10 | 0.58 |
| Simultaneous Processing | 1 | 0.15 | 0.15 | 1.73 | 1.52 | 0.17 |
| R1c = 15.11 p = .235 | 2 | 0.78 | 0.14 | -0.57 | -0.81 | 0.91 |
| Anderson LR Test Z=15.56 p=.212 | 3 | 0.22 | 0.15 | -1.22 | -1.21 | 0.67 |
| DIF LR = 25.25, *p* = .065 | 4 | 0.04 | 0.15 | 1.50 | 1.59 | 0.29 |
| | 5 | -1.19 | 0.20 | -1.77 | -1.15 | 0.43 |
| Spatial Skill | 1 | 0.13 | 0.08 | 5.45 | 5.40 | 0.00 |
| R1c = 37.49 p = .000 | 4 | -0.51 | 0.08 | -2.12 | -2.06 | 0.00 |
| Anderson LR Test Z=35.24 p=.000 | 5 | 0.38 | 0.08 | -2.96 | -3.00 | 0.02 |
| DIF LR = 8.73, *p* = .362 | | | | | | |
| Discriminant Skill | 1 | 0.78 | 0.07 | -1.12 | -0.94 | 0.53 |
| R1c = 2.25 p = .894 | 2 | -1.08 | 0.09 | 0.07 | -0.01 | 0.87 |
| Anderson LR Test Z=2.34 p=.885 | 3 | -0.29 | 0.08 | -0.13 | -0.07 | 0.96 |
| DIF LR = 7.72, *p* = .805 | 4 | 0.59 | 0.07 | 1.04 | 0.99 | 0.68 |
| Memory Skill | 2 | 0.56 | 0.08 | 0.54 | 0.57 | 0.72 |
| R1c = 53.84 p = .104 | 4 | 0.01 | 0.07 | -1.40 | -1.22 | 0.15 |
| Anderson LR Test Z=56.71 p=.064 | 6 | -0.59 | 0.08 | 1.07 | 0.86 | 0.48 |
| DIF LR = 34.14, *p* = .196 | 8 | 0.59 | 0.08 | -0.41 | -0.37 | 0.20 |
| | 12 | -0.26 | 0.08 | 0.21 | 1.05 | 0.23 |

| SubScale | Item | Difficulty | Error | OUTFIT | INFIT | P Value |
|---|---|---|---|---|---|---|
| | 13 | 0.41 | 0.08 | -1.52 | -1.70 | 0.73 |
| | 15 | 0.39 | 0.08 | 0.19 | -0.16 | 0.95 |
| | 18 | -1.10 | 0.08 | 0.71 | 0.68 | 0.61 |
| Categorization | 1 | 0.87 | 0.06 | 0.91 | 0.87 | $p > .05$ |
| Variability accounted = 68% | 2 | -0.55 | 0.10 | 1.14 | 1.18 | $p > .05$ |
| | 3 | 0.31 | 0.10 | 1.10 | 1.19 | $p > .05$ |
| | 4 | -0.30 | 0.09 | 1.08 | 1.12 | $p > .05$ |
| | 5 | 0.13 | 0.09 | 0.96 | 1.00 | $p > .05$ |
| | 6 | 0.21 | 0.10 | 0.97 | 0.92 | $p > .05$ |
| | 7 | 0.38 | 0.10 | 0.82 | 0.89 | $p > .05$ |
| | 8 | 0.47 | 0.10 | 0.87 | 0.96 | $p > .05$ |

Table 4. Category Structure (N=460)

| Item | Observed Count | Structure Calibration | Category Measure | Outfit | Infit |
|---|---|---|---|---|---|
| 0 | 1208 | ---- | -2.24 | 1.05 | 1.04 |
| 1 | 1108 | -0.91 | -0.63 | 0.90 | 0.93 |
| 2 | 781 | 0.03 | 0.63 | 0.76 | 0.80 |
| 3 | 407 | 0.88 | 2.23 | 1.16 | 1.16 |

**Table 5. Descriptive Statistics for Mean Comparisons (N=460)**

| Level | | N | Mean | Lower 95% | Upper 95% | Subscale |
|---|---|---|---|---|---|---|
| Asian | A | 41 | 3.50 | 3.15 | 3.85 | Analytic |
| White | A | 276 | 3.43 | 3.29 | 3.57 | Skill |
| Other | A B C | 18 | 2.98 | 2.53 | 3.42 | |
| Hispanic | B C | 78 | 2.97 | 2.71 | 3.23 | |
| Black | C | 46 | 2.65 | 2.30 | 3.01 | |
| White | A | 276 | 5.61 | 5.51 | 5.71 | Sequential |
| Asian | A B | 41 | 5.49 | 5.23 | 5.74 | Processing |
| Hispanic | B | 78 | 5.26 | 5.07 | 5.45 | |
| Other | B | 18 | 5.23 | 4.90 | 5.55 | |
| Black | B | 46 | 5.01 | 4.75 | 5.28 | |
| Asian | A | 41 | 4.76 | 4.53 | 4.98 | Simultaneous |
| White | A | 276 | 4.64 | 4.55 | 4.73 | Processing |
| Hispanic | A B | 78 | 4.58 | 4.42 | 4.74 | |
| Black | B | 46 | 4.25 | 4.02 | 4.47 | |
| Other | B | 18 | 4.18 | 3.90 | 4.46 | |
| White | A | 276 | 3.53 | 3.41 | 3.65 | Spatial |
| Hispanic | A | 78 | 2.90 | 2.67 | 3.13 | Skill |
| Asian | A B | 41 | 3.47 | 3.16 | 3.78 | |
| Black | B C | 46 | 2.49 | 2.18 | 2.81 | |
| Other | C | 18 | 3.27 | 2.88 | 3.67 | |
| Black | A | 46 | 2.93 | 2.55 | 3.30 | Discriminant |
| Hispanic | A | 78 | 2.74 | 2.47 | 3.01 | Skill |
| Asian | A | 41 | 2.70 | 2.33 | 3.07 | |
| Other | A | 18 | 2.57 | 2.10 | 3.04 | |
| White | A | 276 | 2.41 | 2.27 | 2.56 | |
| Other | A | 18 | 7.33 | 4.27 | 10.40 | Memory |
| Hispanic | A | 78 | 6.43 | 4.42 | 8.44 | |
| White | A | 276 | 6.18 | 5.43 | 6.94 | |
| Asian | A | 41 | 5.71 | 4.42 | 6.99 | |
| Black | A | 46 | 5.20 | 2.83 | 7.57 | |
| Asian | A | 41 | 9.77 | 8.72 | 10.82 | Categorization |
| Black | A | 46 | 8.81 | 7.75 | 9.87 | |
| White | A | 276 | 8.73 | 8.31 | 9.14 | |
| Hispanic | A | 78 | 8.57 | 7.80 | 9.34 | |
| Other | A | 18 | 8.50 | 7.18 | 9.83 | |

Levels not connected by same letter are significantly different

Figure 1. Person and Item Measures

Figure 2. Category Probabilities-- Structure measures at intersections

```
P       ++---------+---------+---------+---------+---------+---------++
R  1.0 +                                                             +
O       |                                                             |
B       |0                                                          33|
A       | 0000                                                  333  |
B   .8 +     000                                            333     +
I       |        00                                          333        |
L       |          00                                    33           |
I       |            00                                33             |
T   .6 +              00                             3               +
Y       |                00                        33                |
    .5 +                  0                       3                  +
O       |                   00                   33                   |
F   .4 +                  111*11111111  2222222*22                   +
        |               111    00     2**     33   222                |
R       |            111         00 22   11133        222             |
E       |         111             2*       311           222          |
S   .2 +      111             222  00   33   11            2222      +
P       |  1111             222       0*3       111             2222  |
O       |11              222        333 000        111              22|
N       |         2222222     333333       000000     1111111         |
S   .0 +*********333333333333                   000000000000*********+
E       ++---------+---------+---------+---------+---------+---------++
       -3        -2        -1         0         1         2         3
```