Analysis of Preschool Assessment Data

Desired Results Development Profile Preschool ©

DRDP – PS (2010)

Ifthika “Shine” Nissar

California State University, San Bernardino

Dr. Joseph Jesunathadas

EDUC 738 – Assessment and Evaluation

August 30, 2014

Table of Contents

 Description of the Context of the Assessment 3

Table of Specification 5

Oder of the Measures (from easiest to the hardest by reading and studying) 6

Classical Theory Calculations – Excel Sheet

Assessment Data of Desired Results Development Profile - Preschool (DRDP-PS) © 22

Classical Test Theory: Interpretation of Results 6

Modern Test Theory: Interpretation of Results 10

Winsteps Analysis

TABLE 1.0 Preschool DRDP Assessment Data Variable Table 10

 Analysis of the Variable Table 11

TABLE 3.1 Preschool DRDP Assessment Data Summary Statistics 13 Analysis of Summary Statistics 12

TABLE 6.1 Preschool DRDP Assessment Student (Person) Fit Order Table & Analysis 14

TABLE 6.4 Preschool DRDP Assessment Data Scalogram & Analysis 15

TABLE 10.1 Preschool DRDP Assessment Data Measures (Item) Fit Order Table 16

TABLE 10.3 Preschool DRDP Assessment Data MEASURES CATEGORY/OPTION/DISTRACTOR MISFIT ORDER & Analysis 17

TABLE 24.0.Question Dimensionality Table and Analysis 18

Conclusions & Professional judgments and 19 & 20

References 21

**Analysis of Preschool Assessment Data**

**Description of the Context of the Assessment**

Preschool students between the ages of Three to Five years attending Early Childhood Education (ECE) programs monitored by school districts are required to be assessed using a battery of assessment tools according to federal and state mandates. The Desired Results Developmental Profile © – Preschool (DRDP-PS (2010) is an assessment instrument developed by the California Department of Education, Child Development Division (CDE/CDD) to assess the growth and development of preschool children. The DRDP-PS© is designed for teachers to observe, document, and reflect on the learning, development, and progress of all children in an early care and education program. Parents are encouraged to provide input by sharing observations of their child’s growth and development. Results of the assessment are intended to be used by the teacher to plan and individualize the curriculum for individual children and groups of children and to guide continuous program improvement. Teachers develop student portfolios with work samples, photos and digital recordings to document student progress. ECE programs funded by federal funds such as the Head Start preschool program require this assessment completed three times a year to capture baseline, midyear and end of the year data. State funded preschool programs such as the California State Preschool Program mandates the completion of the assessment twice a year: initial assessment completed by 60 days of student enrollment and at the end of the year. In addition to curriculum planning and individualizing instruction, assessment data is discussed during parent teacher conferences twice a year. Families are an integral part of the students’ educational journey. The DR system was developed based on the following six Desired Results:

**Desired Results for Children**

DR 1: Children are personally and socially competent.

DR 2: Children are effective learners.

DR 3: Children show physical and motor competence.

DR 4: Children are safe and healthy.

**Desired Results for Families**

DR 5: Families support their child’s learning and development

DR 6: Families achieve their goals.

The DRDP PS (2010) has 43 Measures in 7 Domains: Self and Social Development (SSD). Language and Literacy Development (LLD), English Language Development (ELD), Cognitive Development (COG), Mathematical Development (MATH), Motor and Perceptual Development (MPD), Health and Safety (HLTH) as stated in Table of Specifications on page 5. (<http://www.cde.ca.gov/sp/cd/ci/documents/drdp2010preschooleng.pdf>). For the purpose of this assignment, student data was analyzed for the 12 Measures of the Self and Social Development (SSD) Domain. Each Measure has 4 Developmental levels. The developmental level of each measure represents a developmental continuum. Following are the Developmental Indicators and Ratings of the Self and Social Development (SSD) Domain:

|  |  |
| --- | --- |
| **Development Level** | **Rating** |
| Exploring | 0 |
| Developing | 1 |
| Building | 2 |
| Integrating | 3 |

Below is a sample from the DRDP assessment tool: Self and Social Development Domain: Measure 5

This Analysis of Preschool Assessment Data report utilizes data of 30 preschool students of a Head Start preschool classroom from a school district in the Coachella Valley. The assessment was completed during spring 2013 for the 2013-2014 school year. This assessment instrument measures construct and latent traits of self and social development. The focus of the construct of social development measure is to understand how children see themselves and how they relate to others.

**Table of Specification**

|  |  |
| --- | --- |
| **Developmental Domains of the DRDP** | **Number of Measures** |
| Self and Social Development (SSD) | 12 (1-12) |
| Language and Literacy Development (LLD) | 10 (13-22) |
| English Language Development (ELD) | 04 (23-26) |
| Cognitive Development (COG) | 05 (27-31) |
| Mathematical Development (MATH) | 06 (32-37) |
| Motor and Perceptual Development (MPD) | 03 (38-40) |
| Health and Safety (HLTH) | 03 (41-43) |

**Measures of the Self and Social Development (SSD) Domain**

SSD1: Identity of self

SSD2: Recognition of own skills and accomplishments

SSD3: Expressions of empathy

SSD4: Impulse control

SSD5: Taking turns

SSD6: Awareness of diversity in self and others

SSD7: Relationships with adults

SSD8: Cooperative play with peers

SSD9: Socio-dramatic play

SSD10: Friendships with peers

SSD11: Conflict negotiation

SSD12: Shared use of space and materials

**Oder of the Measures of the Self and Social Development (SSD) Domain from easiest to the hardest by reading and studying the measures with my experience in the profession of ECE.**

|  |  |
| --- | --- |
| 1 | SSD7: Relationships with adults |
| 2 | SSD1: Identity of self |
| 3 | SSD2: Recognition of own skills and accomplishments |
| 4 | SSD9: Socio-dramatic play  |
| 5 | SSD10: Friendships with peers |
| 6 | SSD8: Cooperative play with peers  |
| 7 | SSD12: Shared use of space and materials |
| 8 | SSD5: Taking turns |
| 9 | SSD3: Expressions of empathy |
| 10 | SSD4: Impulse control |
| 11 | SSD6: Awareness of diversity in self and others |
| 12 | SSD11: Conflict negotiation |

**Classical Test Theory: Interpretation of Results**

Data from the Desired Results Development Profile © – Preschool (DRDP-PS (2010) gathered from 30 students of a Head Start preschool classroom from a school district in Coachella Valley was prepared for the Classical Test Theory calculations. For the purpose of this assignment only the Self and Social Development (SSD) Domain was analyzed. There are 12 measures (items) in the SSD. There are 4 ratings of the score in this assessment tool: 0 – Exploring, 1- Beginning, 2-Developing and 3 – Integrating. From the four ratings of 0,1,2,3, data was organized as 1,2, 3 as none of the 30 students rated 0, the exploring stage. For the purpose of this assignment persons are identified as students and items are identified as measures.

Please review excel sheet on page 22 (Assessment Data of Desired Results Development Profile - Preschool (DRDP-PS) © for the computation of the following test characteristics: Mean, median, and mode(s) of person scores; range and standard deviation; Item difficulty and item discrimination; and Reliability coefficient and Standard Error of Measurement (SEM).

Student total scores and the total of the measures (items) were calculated for each of the 12 measures (items). Scores were sorted (vertically) ranked from highest to lowest. Students earned a Total Score between 34 and 19 from a possible score of 36. The total of the measures (items) were sorted (horizontally) and reordered from easiest to the hardest. This it gave me the opportunity to endorse each item. Also reordering from easiest to the hardest allows me to compute the item difficulty.

Total possible score for the 12 measures from the 30 students was 90. Results of the scores ranged from 74 to 56. Measure 7 - SSD7: Relationships with adults scoring 74 was the easiest item and Measure SSD11: Conflict negotiation scoring 56 was the hardest item. This is typical in general with preschool students as relationship building with adults is very important to young students to feel safe and secure and young children thrive on adult praise and acceptance. Young children have a hard time with the concept of sharing. This is very common during the early months of attendance in preschool as students are gravitated to the novelty of instructional materials and toys in their classroom. When I ranked ordered the measures by studying each measure and based on my experience in the profession of ECE, as stated on page 6, I chose the easiest item to as SSD7: Relationships with adults hardest to be SSD 12 SSD11: Conflict negotiation.

The Mean (average of the total value) was 26. The Median (value in the middle) score was 27 and the Mode (frequently occurring value) was 30. The Range was 15 (the difference between the highest and lowest value).

 The Standard Deviation of the total student scores was 4.1 (Standard Deviation is the square root of the variance of the sample as indicated in figure 1)

|  |  |
| --- | --- |
| http://www.alghamdi-biostatistics.com/variance%20sample%20formula.gif | Figure http://lrieber.coe.uga.edu/edit6900/resources/SD_formula_large.gif |
| Formula to calculate the Variance. Variance is needed to compute the S.D | Formula to calculate the Standard Deviation |

Item Difficulties were calculated, generating a *p*-value for each item. Item difficulty is calculated by dividing the total score (item/measure total) earned by the total possible. Subtracting the *p*-value from 1 gave the *q*-value, or the proportion of students who did not answer the question correctly. *p*-value ranged from was 0.82 (SSD7: Relationships with adults – being the easiest Measure) to 0.62 (SSD11: Conflict negotiation – being the most difficult Measure) which indicates a narrow range in the item difficulties. *q*-value 0.18 (SSD7: Relationships with adults) to 0.38 (SSD11: Conflict negotiation). This indicates the spectrum of student performance on the assessment items as well as item difficulties.

Variance ranged from 0.18 to 0.36. Variance values were calculated for each item by taking the product of the *p* and *q* values. The Sum of item Variances was 3.3. The Variances of students’ Total Scores had a value of 16.1. These two values were used to calculate the Kuder-Richardson 20 (KR20) reliability coefficient of 0.80 was calculated by:

 1 - VarE(variable of obtained score)

 VarX(variable of error Score) 1

 = 1 - 3.3

 16.1

The score of 0.80 is a high KR20 reliability coefficient value. This indicates that only about 20% of the total score variance was associated with error variance, while 80% was associated with true score variance. This coefficient value is a strong indication of internal consistency within the data. Since the reliability is high (0.80), the error rate is low (0.20) as indicated by this data. According to Frisbie (1998) the Reliability Coefficient, can range between 0.0 and + 1.0, usually the teacher-made tests has values around 0.50 for and for commercially prepared standardized tests around 0.90 The DRDP assessment tool is published by the Child Development Division of California Dept. of Education with input from researchers from the profession child development, special education, program administrators and many stakeholders such as experts from the field of assessment, program evaluation in the profession of early childhood education and practitioners from age groups: infants/toddlers and preschoolers (DRDP, 2010).

The Reliability Coefficient along with the other scores helps administrators of ECE programs to better interpret the data of the Self and Social Development (SSD Domain of the DRDP preschool assessment.

Item Discrimination values were calculated using the top and bottom 27% of the student sample population (8 students from each category). Values for item discrimination ranged between .17 and 0.38. Measures 9 and 10 generated values of 0.17 and measures 2 and 11 generated a value of 0.38.

The Standard Error of Measurement (SEM) value was 1.8. SEM is related to reliability. As the reliability increases, the SEM decreases. The SEM is an indication to estimate the dispersion of the measurement errors around the true scores. The SEM was calculated in the following way: S.D multiplied by the square root of 1- Reliability Coefficient.

**Modern Test Theory: Interpretation of Results**

The original data was prepared for the Modern Test Theory calculations through the use of the Rasch Model. The data of the 12 Measures were coded according to the specifications of the WINSTEPS software.

1. Variable Table - Table 1.0 on page 11

The figure represents the **student abilities and item calibration on a common Logit Scale**.

**On the Left is the student ability** and on **the Right is the Item calibration stated as measures**.

**The data are normally distributed** for students and measures.

**The spread of the distribution** of students were wider than the spread of the measure distribution.

**The Range** for the Students was 8.24 (5.29+(-3.05)) and for calibration is 4.24 (2.11+(-1.69)) which is almost double.

**Distribution of the variables**: students are distributed between -3.05 to 5.29 Logits and Measures were distributed between -1.69 to 2.11(Measure 1 and Measure 12) Logits.

**Item calibration -** Items were calibrated according to the ability of the measures. Item 12 (Measure 12 - SSD 12: Shared use of space and materials) is far above compared To Item 1 (Measure 1 - SSD1: Identity of self ) many students have confidence in getting this item correct. Measure 12 (Item 12 - SSD12: Shared use of space and materials) is calibrated much lower and addressed as an easier concept Therefore, students can endorse this item. Students whose measures are above item 12 have a greater chance of answering item 12 correctly compared to students’ measures are below item 12. Since there are no items above 12, it is apparent that students above 12 need items that are targeted to their ability level. This score is contradictory to the way measures are organized on the assessment (as being the hardest) well as the way I arranged (Midrange as #7) according to the difficulty for students.

**WINSTEPS Variable Table**

**TABLE 1.0 Preschool DRDP Assessment Data Variable Table**

**INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS**

----------------------------------------------------------------------------

 STUDENTS - MAP - MEASURESS

 <more>|<rare>

 6 T+

 |

 |

 |

 83 |

 5 +

 |

 23 53 |

 |

 |

 4 +

 03 12 S|

 |

 03 03 33 53 |

 |

 3 43 +

 |

 73 |

 |

 23 32 43 63 | Item 12

 2 +

 |T

 62 82 93 |

 M|

 |

 1 22 +S

 |

 | Item 11

 | Item 10 Item 7 Item 8 Item 9

 |

 0 42 72 72 +M

 | Item 6

 | Item 5

 | Item 4

 | Item 2 Item 3

 -1 12 S+S

 |

 |

 12 62 92 92 | Item 1

 |T

 -2 +

 82 |

 |

 32 |

 |

 -3 52 +

 <less>|<frequ>

1. Summary Statistics (Summary of Measured Students & Summary of Measured Measures/Items) – Table 3.1 – on page 13

This table showed general statistics and measures (items):

**Mean -**Mean of the student in Self and Social Development was 1.36 Logits higher than the mean calibration of item difficulties (0 Logits), I would recommend to include more difficult items to the assessment. There was a far greater distance between the mean of the students and the mean of the Measure. This is too wide. Because the mean of the student from the item is > 1 this can yield to an inaccurate measurement. Items that are aligned to abilities are measured more accurately. As Measure 6, 5, 4, 3 and 2 are situated closer to the mean of the Measures there could be no difference in skill mastery.

**S.D :** S.D was 2.37 Logits and for measures was .92 Logits. S.D for students were about 2 ½ times higher than for students item calibration.

**Mean of the Raw score:**  Mean of the students was 26.4 and Mean of the Measures was 66.1

**Standard Deviation of the Raw Score:** S.D of students was 4.0 and S.D for the Measures was 4.4

**RMSE (Root Mean Square Error):** RSME for students was .81 and .49 for items

**Student Reliability** was .88 and **Measure Reliability** was .72

STUDENT RAW SCORE RELIABILITY = .87. This compared with the Rasch Model are about the same .80. This reflects that the data is very reliable to the sample of 12 measures of the DRDP.

**WINSTEPS Summary Statistics**

TABLE 3.1 Preschool DRDP Assessment Data Summary Statistics

INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS

--------------------------------------------------------------------------------

 **SUMMARY OF 30 MEASURED STUDENTS**

-------------------------------------------------------------------------------

| RAW MODEL INFIT OUTFIT |

| SCORE COUNT MEASURE ERROR MNSQ ZSTD MNSQ ZSTD |

|-----------------------------------------------------------------------------|

| MEAN 26.4 12.0 1.36 .75 .94 .0 .98 .0 |

| S.D. 4.0 .0 2.37 .12 .46 1.2 .62 1.1 |

| MAX. 34.0 12.0 5.29 1.01 2.05 3.4 2.61 2.2 |

| MIN. 19.0 12.0 -3.05 .62 .08 -1.7 .05 -1.6 |

|-----------------------------------------------------------------------------|

| REAL RMSE .81 ADJ.SD 2.22 SEPARATION 2.73 STUDEN RELIABILITY .88 |

|MODEL RMSE .75 ADJ.SD 2.24 SEPARATION 2.97 STUDEN RELIABILITY .90 |

| S.E. OF STUDENT MEAN = .44 |

-------------------------------------------------------------------------------

STUDENT RAW SCORE-TO-MEASURE CORRELATION = .99

CRONBACH ALPHA (KR-20) STUDENT RAW SCORE RELIABILITY = .87

 **SUMMARY OF 12 MEASURED MEASURESS**

-------------------------------------------------------------------------------

| RAW MODEL INFIT OUTFIT |

| SCORE COUNT MEASURE ERROR MNSQ ZSTD MNSQ ZSTD |

|-----------------------------------------------------------------------------|

| MEAN 66.1 30.0 .00 .46 .98 -.2 .98 -.2 |

| S.D. 4.4 .0 .92 .01 .38 1.5 .51 1.3 |

| MAX. 74.0 30.0 2.11 .49 1.81 2.9 2.10 2.4 |

| MIN. 56.0 30.0 -1.69 .45 .57 -2.1 .44 -1.9 |

|-----------------------------------------------------------------------------|

| REAL RMSE .49 ADJ.SD .78 SEPARATION 1.59 MEASUR RELIABILITY .72 |

|MODEL RMSE .46 ADJ.SD .80 SEPARATION 1.75 MEASUR RELIABILITY .75 |

| S.E. OF MEASURES MEAN = .28 |

-------------------------------------------------------------------------------

UMEAN=.000 USCALE=1.000

MEASURES RAW SCORE-TO-MEASURE CORRELATION = -1.00

360 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 357.79 with 318 d.f. p=.0615

iii Student (Person) Fit Order Table – Table 6.1 This table gives us the **INTFIT and OUTFIT** information and tells how well the student’s responses fit the model. Looking at the outfit and using a criteria of 2.0 for standardized Z (ZSTD), there are 6 students who do not fit the model. Their patterns of responses reflect that they don’t adhere to the Rasch model. The pattern of Responses has too many inconsistencies. These students’ response patterns underfits the Rasch Model. I will delete students 16, 21, 10, 17, 6 and 8 and run the data again. **Point Measure:** This table also shows a Point Measure correlations which shows if the Measures are corresponding 1:1 with the score.

**WINSTEPS Student (Person) Fit Order Table**

TABLE 6.1 Preschool DRDP Assessment Student (Person) Fit Order Table

INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS

--------------------------------------------------------------------------------

STUDENT: REAL SEP.: 2.73 REL.: .88 ... MEASURES: REAL SEP.: 1.59 REL.: .72

 STUDENT STATISTICS: MISFIT ORDER

---------------------------------------------------------------------------------------------

|ENTRY TOTAL MODEL| INFIT | OUTFIT |PT-MEASURE |EXACT MATCH| |

|NUMBER SCORE COUNT MEASURE S.E. |MNSQ ZSTD|MNSQ ZSTD|CORR. EXP.| OBS% EXP%| STUDENT|

|------------------------------------+----------+----------+-----------+-----------+--------|

| 16 26 12 1.64 .77|1.93 1.7|2.61 2.0|A .50 .30| 66.7 81.6| 82 |

| 21 24 12 .03 1.01|2.05 1.2|2.34 1.3|B .09 .26| 83.3 91.2| 42 |

| 10 29 12 3.02 .63|1.97 3.4|1.91 2.2|C .14 .37| 41.7 66.7| 43 |

| 17 26 12 1.64 .77|1.31 .7|1.74 1.2|D-.25 .30| 83.3 81.6| 62 |

| 6 30 12 3.41 .62|1.41 1.8|1.68 2.0|E-.09 .38| 50.0 66.2| 33 |

| 8 30 12 3.41 .62|1.41 1.8|1.68 2.0|F-.09 .38| 50.0 66.2| 03 |

| 4 31 12 3.81 .63|1.25 1.1|1.67 1.8|G .04 .38| 75.0 65.9| 12 |

| 30 19 12 -3.05 .63|1.32 1.3|1.30 1.0|H .07 .39| 41.7 65.8| 52 |

| 27 22 12 -1.63 .79| .81 -.2|1.13 .4|I .32 .35| 91.7 83.0| 92 |

| 13 27 12 2.17 .69|1.10 .4|1.02 .2|J .20 .34| 66.7 75.9| 32 |

| 18 26 12 1.64 .77| .85 -.2|1.06 .3|K .26 .30| 83.3 81.6| 93 |

| 29 20 12 -2.64 .66|1.02 .2| .94 .0|L .36 .38| 58.3 71.1| 32 |

| 15 27 12 2.17 .69| .97 .0| .96 .1|M .30 .34| 83.3 75.9| 63 |

| 3 33 12 4.70 .72| .89 -.2| .93 .1|N .45 .37| 83.3 78.2| 53 |

| 11 28 12 2.62 .65| .91 -.2| .84 -.3|O .43 .36| 75.0 69.6| 73 |

| 25 22 12 -1.63 .79| .77 -.3| .87 .0|o .42 .35| 91.7 83.0| 62 |

| 23 23 12 -.91 .92| .82 .0| .62 -.2|n .20 .31| 91.7 88.2| 12 |

| 7 30 12 3.41 .62| .80 -.9| .74 -.8|m .58 .38| 83.3 66.2| 53 |

| 2 33 12 4.70 .72| .72 -.7| .57 -.7|l .66 .37| 83.3 78.2| 23 |

| 19 25 12 .96 .90| .71 -.2| .52 -.4|k .26 .27| 91.7 88.1| 22 |

| 9 30 12 3.41 .62| .69 -1.6| .63 -1.3|j .69 .38| 83.3 66.2| 03 |

| 28 21 12 -2.18 .70| .68 -.8| .55 -.8|i .66 .37| 83.3 77.6| 82 |

| 1 34 12 5.29 .83| .67 -.5| .47 -.6|h .66 .34| 91.7 84.3| 83 |

| 14 27 12 2.17 .69| .66 -.9| .53 -.9|g .64 .34| 83.3 75.9| 43 |

| 5 31 12 3.81 .63| .65 -1.6| .58 -1.4|f .73 .38| 91.7 65.9| 03 |

| 26 22 12 -1.63 .79| .64 -.6| .50 -.7|e .61 .35| 91.7 83.0| 92 |

| 12 27 12 2.17 .69| .62 -1.1| .49 -1.1|d .69 .34| 83.3 75.9| 23 |

| 24 22 12 -1.63 .79| .59 -.7| .44 -.8|c .66 .35| 91.7 83.0| 12 |

| 20 24 12 .03 1.01| .08 -1.7| .05 -1.6|b .00 .26|100.0 91.2| 72 |

| 22 24 12 .03 1.01| .08 -1.7| .05 -1.6|a .00 .26|100.0 91.2| 72 |

|------------------------------------+----------+----------+-----------+-----------+--------|

| MEAN 26.4 12.0 1.36 .75| .94 .0| .98 .0| | 79.2 77.3| |

| S.D. 4.0 .0 2.37 .12| .46 1.2| .62 1.1| | 15.9 8.5| |

---------------------------------------------------------------------------------------------

iv. Scalogram – Table 6.4

The **Scalogram** gives us the most misfitting response strings and what the misfitting string was. The data reveals that:

Student number 16 did not fit the responses for measures (items) 11 (SSD11: Conflict negotiation), and 12 (SSD12: Shared use of space and materials) 16 student’s responses did not fit.

**WINSTEPS Scalogram**

TABLE 6.4 Preschool DRDP Assessment Data Scalogram

INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS

--------------------------------------------------------------------------------

**MOST MISFITTING RESPONSE STRINGS**

STUDENT OUTMNSQ |MEASURES

 | 1 11

 |132450712

 high---------

 16 82 2.61 A|.......31

 21 42 2.34 B|.3..1....

 10 43 1.91 C|......1..

 17 62 1.74 D|......33.

 6 33 1.68 E|........3

 8 03 1.68 F|........3

 4 12 1.67 G|2.......3

 27 92 1.13 I|.1.......

 13 32 1.02 J|......3..

 18 93 1.06 K|.......3.

 15 63 .96 M|.....3...

 3 53 .93 N|...2.....

 25 62 .87 o|....1....

 23 12 .62 n|.......1.

 19 22 .52 k|..3......

 |-----low-

 |132451711

 | 0 12

1. Measures (Item) Fit Order Table – Table 10.1

This table gives us the INFIT and the OUTFIT information tells how well the well the measures (items) fit the model. Looking at the outfit and using a criteria of 2.0 for standardized Z (ZSTD), measure 11 does not fit the model. I will delete items 11 (as it has a value of 2.4 ZSTD and run the data again. The pattern of the Measures above 2.0 does not fit the Rasch Model.

**WINSTEPS Measures (Item) Fit Order Table**

TABLE 10.1 Preschool DRDP Assessment Data Measures (Item) Fit Order Table

INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS

--------------------------------------------------------------------------------

STUDENT: REAL SEP.: 2.73 REL.: .88 ... MEASURES: REAL SEP.: 1.59 REL.: .72

 **MEASURES STATISTICS: MISFIT ORDER**

----------------------------------------------------------------------------------------------

|ENTRY TOTAL MODEL| INFIT | OUTFIT |PT-MEASURE |EXACT MATCH| |

|NUMBER SCORE COUNT MEASURE S.E. |MNSQ ZSTD|MNSQ ZSTD|CORR. EXP.| OBS% EXP%| MEASURES|

|------------------------------------+----------+----------+-----------+-----------+---------|

| 11 63 30 .62 .45|1.81 2.9|2.10 2.4|A .51 .60| 50.0 76.4| Item 11 |

| 12 56 30 2.11 .48|1.28 1.0|1.54 1.2|B .66 .65| 83.3 80.4| Item 12 |

| 7 64 30 .42 .45|1.48 1.9|1.38 1.1|C .60 .61| 66.7 75.9| Item 7 |

| 3 70 30 -.79 .46|1.24 1.0|1.41 1.1|D .69 .65| 73.3 78.0| Item 3 |

| 5 68 30 -.38 .45| .84 -.6|1.16 .5|E .72 .63| 90.0 77.0| Item 5 |

| 2 70 30 -.79 .46|1.05 .3| .85 -.3|F .67 .65| 73.3 78.0| Item 2 |

| 4 69 30 -.58 .46| .87 -.4| .75 -.6|f .68 .64| 76.7 77.6| Item 4 |

| 10 64 30 .42 .45| .70 -1.3| .60 -1.2|e .56 .61| 86.7 75.9| Item 10 |

| 9 64 30 .42 .45| .67 -1.5| .54 -1.4|d .58 .61| 86.7 75.9| Item 9 |

| 1 74 30 -1.69 .49| .67 -1.2| .57 -.9|c .76 .70| 90.0 80.4| Item 1 |

| 6 67 30 -.18 .45| .58 -2.0| .45 -1.8|b .61 .62| 86.7 76.3| Item 6 |

| 8 64 30 .42 .45| .57 -2.1| .44 -1.9|a .62 .61| 86.7 75.9| Item 8 |

|------------------------------------+----------+----------+-----------+-----------+---------|

| MEAN 66.1 30.0 .00 .46| .98 -.2| .98 -.2| | 79.2 77.3| |

| S.D. 4.4 .0 .92 .01| .38 1.5| .51 1.3| | 11.4 1.6| |

----------------------------------------------------------------------------------------------

1. Measures (Item) Fit Order Table – Table 10.3

This table gives us the frequencies of the items and how each student endorsed the item. Measures (Items) 1 and 6 are not working well.

**WINSTEPS**

**MEASURES CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER TABLE**

TABLE 10.3 Preschool DRDP Assessment Data MEASURES CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER TABLE

INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS

--------------------------------------------------------------------------------

 MEASURES CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER

-----------------------------------------------------------------------

|ENTRY DATA SCORE | DATA | AVERAGE S.E. OUTF PTMEA| |

|NUMBER CODE VALUE | COUNT % | MEASURE MEAN MNSQ CORR.| MEASURE|

|--------------------+------------+--------------------------+--------|

| 11 A 1 1 | 4 13 | -1.84 .37 1.0 -.53 |Item 11 | 1 Developing

| 2 2 | 19 63 | 1.58 .56 1.8 .12 | | 2 Building

| 3 3 | 7 23 | 2.60 .34 2.3 .29 | | 3 Integrating

| | | | |

| 12 B 1 1 | 7 23 | -1.45 .55 1.4 -.65 |Item 12 | 1 Developing

| 2 2 | 20 67 | 2.02 .45 1.1 .39 | | 2 Building

| 3 3 | 3 10 | 3.54 .13 1.3 .31 | | 3 Integrating

| | | | |

| 7 C 1 1 | 4 13 | -.96 1.36 1.0 -.39 |Item 7 | 1 Developing

| 2 2 | 18 60 | .94 .50 .9 -.22 | | 2 Building

| 3 3 | 8 27 | 3.48 .42 1.2 .54 | | 3 Integrating

| | | | |

| 3 D 1 1 | 2 7 | -2.34 .71 1.0 -.42 |Item 3 | 1 Developing

| 2 2 | 16 53 | .47 .51 .8 -.41 | | 2 Building

| 3 3 | 12 40 | 3.18 .43 1.5 .63 | | 3 Integrating

| | | | |

| 5 E 1 1 | 2 7 | -.80 .83 2.1 -.24 |Item 5 | 1 Developing

| 2 2 | 18 60 | .22 .46 .5 -.59 | | 2 Building

| 3 3 | 10 33 | 3.86 .25 .4 .74 | | 3 Integrating

| | | | |

| 2 F 1 1 | 1 3 | -3.05 .4 -.35 |Item 2 | 1 Developing

| 2 2 | 18 60 | .44 .49 .8 -.48 | | 2 Building

| 3 3 | 11 37 | 3.28 .42 1.1 .61 | | 3 Integrating

| | | | |

| 4 f 1 1 | 1 3 | -3.05 .4 -.35 |Item 4 | 1 Developing

| 2 2 | 19 63 | .48 .48 1.1 -.49 | | 2 Building

| 3 3 | 10 33 | 3.48 .32 .6 .63 | | 3 Integrating

| | | | |

| 10 e 1 1 | 1 3 | -2.64 .4 -.31 |Item 10 | 1 Developing

| 2 2 | 24 80 | .99 .44 .9 -.32 | | 2 Building

| 3 3 | 5 17 | 3.97 .59 .8 .49 | | 3 Integrating

| | | | |

| 9 d 1 1 | 1 3 | -2.64 .4 -.31 |Item 9 | 1 Developing

| 2 2 | 24 80 | .97 .44 .9 -.33 | | 2 Building

| 3 3 | 5 17 | 4.06 .52 .6 .51 | | 3 Integrating

| | | | |

| 1 c 2 2 | 16 53 | -.31 .49 .8 -.76 |Item 1 | 2 Building

| 3 3 | 14 47 | 3.28 .29 .4 .76 | | 3 Integrating

| | | | |

| 6 b 2 2 | 23 77 | .57 .44 .7 -.61 |Item 6 | 2 Building

| 3 3 | 7 23 | 3.98 .39 .5 .61 | | 3 Integrating

| | | | |

| 8 a 1 1 | 1 3 | -2.64 .4 -.31 |Item 8 | 1 Developing

| 2 2 | 24 80 | .92 .43 .8 -.38 | | 2 Building

| 3 3 | 5 17 | 4.30 .40 .4 .55 | | 3 Integrating

-----------------------------------------------------------------------

**WINSTEPS Question Dimensionality Table**

This table gives us the STANDARDIZED RESIDUAL variance and shows the level of Raw Variance that is explained within the data set.

In this table 46.1% of Raw variance explained by measures and 35.3% Raw variance explained by persons. Raw variance of 53.9% is unexplained

TABLE 24.0 Preschool DRDP Assessment **Question Dimensionality Table**

INPUT: 30 STUDENTS 12 MEASURESS MEASURED: 30 STUDENTS 12 MEASURESS 3 CATS

--------------------------------------------------------------------------------

 Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

 -- Empirical -- Modeled

Total raw variance in observations = 55.6 100.0% 100.0%

 Raw variance explained by measures = 25.6 46.1% 44.4%

 Raw variance explained by persons = 19.6 35.3% 34.0%

 Raw Variance explained by items = 6.0 10.8% 10.4%

 Raw unexplained variance (total) = 30.0 53.9% 100.0% 55.6%

 Unexplned variance in 1st contrast = 7.7 13.8% 25.5%

 Unexplned variance in 2nd contrast = 5.8 10.5% 19.4%

 Unexplned variance in 3rd contrast = 4.4 7.9% 14.7%

 Unexplned variance in 4th contrast = 3.1 5.5% 10.2%

 Unexplned variance in 5th contrast = 2.6 4.6% 8.5%

.

**Conclusions and Professional Judgments**

12 Measures of the Self and Social Development (SSD) Domain of the DRDP assessment was analyzed using the Classical Test Theory calculations and Modern Test Theory calculations.

Data analysis of the Classical Test Theory yielded a measure of Kuder-Richardson 20 (KR20) of reliability coefficient of 0.80. This is a high reliability which is closer to 1 (as the Reliability Coefficient range should be between 0.0 and + 1.0. that is considered most reliable). According to Frisbie (1998) for commercially prepared standardized tests the Reliability Coefficient is around 0.90. Reliability was calculated with the Modern Test Theory too. Cronbasch Alpha Reliability for Polytomous data for student Reliability was .88 and item (measure) Reliability was .72. The Standard Deviation (S.D.) of Classical Test Theory was 4.08. The S.D of Modern Test Theory for Raw Scores was 4.0 for students and 4.4 for Measures. The S.D for students (2.37 Logits) was much higher than the S.D for the item calibration (.92 Logits) about 2 ½ times higher ( According to the DRDP (2010) “The reliability of an assessment constitutes further evidence for validity. An assessment instrument is considered reliable if different observers rate the same child at the same developmental level for each item or measure and arrive at the same results (p.2).” Item Discrimination values were calculated using the top and bottom 27% of the student sample population (8 students from each category). 9 tables were generated to analyze the data according to the Modern Test Theory using WINSTEPS. The Classical Test Theory and the Modern Test Theory yielded a mean of 24. Based on the data analysis, the DRDP assessment tool can be considered of having a high construct validity.

Validity of this assessment is not an easy judgment. I questioned some of the measures and the interpretation of the data when I reviewed the Variable Table. Measure (Item number) 12 – SSD -Shared use of space and materials. This is a hard skill for preschool students to learn. In the assessment tool, this is item #12 and I rated this as item #7. The assessment results reflected this as a very easy item in the variable table. According to Messik (1995) validity, reliability and comparability are not just measurements of principles. They also have social values. Validity is an overall judgment within the concept it’s employed. “The content aspect of construct validity includes evidence of content relevance, representativeness and technical quality” (p.745). In terms of substantive validity, the scores of the assessment tool was not consistent most of the time. I would like to share that the structural aspect of this assessment was also poor as there were inconsistencies with the scoring structure as reflected by the Scalogram. Messick (1995) discusses that the structural aspect appraises the fidelity of the scoring structure. There were parts of the Self and Social Development (SSD) Domain of this assessment that measured what it was supposed to as reflected by the data.

The data analysis reflected this assessment tool to have a high reliability (0.80). Just because a test is reliable it is not valid. I now understand that reliability and validly are two important measures which have to be paid close attention. A test result could benefit or harm students. Tests should be administered by trained professionals. Formative assessments are very important to help students along and get them prepared for summative assessments. I have truly enjoyed this course as I learned a lot and pushed myself beyond my comfort zone.

References

Desired Results Developmental Profile Preschool © DRDP – PS (2010)

<http://www.cde.ca.gov/sp/cd/ci/DRDPforms.asp>

Frisbie, David A. (1988). Reliability of scores from teacher-made tests*. Educational Measurement, 7*, 25-35.

Frisbie, D. A. (2005). Measurement 101: some fundamentals revisited. *Educational Measurement,*

24*(3)*, 21-8.

Messick, Samuel. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*. 50(9), 741-749.

**Assessment Data of Desired Results Development Profile - Preschool (DRDP-PS) © - excel sheet**

**Please review the excel document**