



2018 No. 085

Third-Party Checking of 2018 Scaling and Equating for the Kentucky Performance Rating for Educational Progress (K-PREP) Tests

Final Report

Prepared for: Kentucky Department of Education
Office of Assessment and Accountability
300 Sower Boulevard
Frankfort, KY 40601

Authors: Bethany H. Bynum
Tanya Longabach
Arthur A. Thacker

Prepared under: Human Resources Research Organization
700 North Hurstbourne Parkway, Suite 100
Louisville, KY 40222

Date: November 30, 2018

Third-Party Checking of 2018 Scaling and Equating for the Kentucky Performance Rating for Educational Progress (K-PREP) Tests

Table of Contents

Executive Summary	ii
Introduction	1
Sample Identification and File Construction.....	2
Calibration and Scaling Procedures.....	2
Equating Procedures	2
Raw-score-to-Scale-Score Procedures	3
Verification of 2018 Scoring Tables	3
Documentation.....	3
Conclusion	4
References.....	5
Appendix A – Control File (Math Grade 3).....	A-1
Appendix B – Winsteps Item Parameter Files (Math Grade 3).....	B-1
Appendix C – Winsteps Anchor File (Grade 3 Reading).....	C-1
Appendix D – Winsteps Score File (Grade 3 reading).....	D-1
Appendix E – Comparison of Files Output (Math Grade 3)	E-1

Executive Summary

Pearson and HumRRO independently calibrated, scaled and equated the 2018 Kentucky Performance Rating for Educational Progress (K-PREP) assessments and produced the raw-score-to-theta-score tables to be applied to students' test results. HumRRO further verified that scoring table were applied accurately by checking our scoring of the student sample against Pearson's. Results calculated by HumRRO were identical to those calculated by Pearson (M. Johnson, email communication, July 16-20, 2018). Given that HumRRO's results were identical to those of Pearson, we are assured that Pearson did not commit processing errors.

Third-Party Checking of 2018 Scaling and Equating for the Kentucky Performance Rating for Educational Progress (K-PREP) Tests

Introduction

In 2012, Kentucky transitioned from the Kentucky Core Content Test (KCCT) to the K-PREP system for spring testing. This transition represented a significant departure from the prior assessment system. The 3-parameter logistic Item-Response Theory (IRT) model was replaced with a Rasch model, a new item-type (i.e., short-constructed-response) was added to the assessments, a new scale-score reporting system was developed for sub-scores, and new cut scores were identified for the reading and mathematics assessments. The transition was also accompanied by a new primary testing contractor, Pearson. As a result, HumRRO's third-party checking process underwent significant changes to accommodate the transition¹.

Equating was added to the process in 2013 to permit comparison of the results across test years. The 2014 tests were equated to the 2013 tests using linking items. In this manner, comparable scores were produced for the 2014 and 2015 K-PREP. Forms for all subjects other than writing were repeated in 2016 from prior years, meaning that existing scoring tables could be used, and no equating was necessary. In 2017, new forms were constructed, and equating analyses was carried out for Reading and Mathematics in all grades. Beginning in 2015, scale scores were computed for the On-Demand Writing tests where simple number correct scores had been used in the past². Writing tests were equated in 2016 and 2017.

In 2018, new test forms were created for Reading and Mathematics. We calibrated the operational items on these test forms, performed equating via linking items, and derived scale scores for reporting that can be directly compared to previous scale scores. The Social Studies and Writing forms were repeated from 2016, as such no equating was necessary. However, HumRRO did verify the score tables used in 2018 were the same as those used in 2016. In 2018, the new Science test was first operational administration. For Science, two operational sets of items were administered across six forms. In other words, for each grade, two base forms were constructed along with two "scrambled" versions of each base form. For each grade, one cluster set of items was repeated across all six forms. The operation science items were calibrated and used to generate raw-to-theta conversion tables for use during standard setting, July 17-19, 2018.

This report describes how student test responses for the 2018 K-PREP assessments were used to create scale scores and place students in Novice, Apprentice, Proficient or Distinguished (NAPD) performance categories. The complex analyses to accomplish these tasks were conducted independently, but cooperatively, by both HumRRO and Pearson staff members. Several interim checks were conducted during the analyses and any discrepancies between the two companies was investigated and ultimately resolved. This process was conducted transparently among Pearson, HumRRO, KDE, and Kentucky's psychometric consultant (Dr. Bill Auty of EdMeasure) via frequent email communications and daily conference calls. The process was guided by a specifications document created by Pearson³ and regularly updated

¹ For additional details on how the assessment system and third-party checking procedures changed, see Bynum and Thacker (2013).

² For additional information on how writing was calibrated and scaled, see K-PREP ODW Calibration and Scaling Specs v0.4.docx.

³ K-PREP Sp18 Calibration, Equating and Scaling Specs v2.0.docx.

based on decisions before and during calibration. This documentation is vital for ensuring consistency of processing across years and serves as a guiding document for subsequent years.

Sample Identification and File Construction

Kentucky selects most of its student population for use in the calibration sample for scaling and equating. However, some students are purposefully exempted. KDE established a set of invalidation codes for excluding students in the calibration file. Kentucky's exemption rules only apply to students who receive accommodations (e.g., Braille forms, audio, large print, etc.) and students with duplicate records (the same identification number and name). The accommodated students receive scores but are simply omitted from the calibration sample. Pearson and HumRRO verified n-counts after this step. For Reading and Math, grades 3-8 were equated.

The next step was to format all the grade/subject files to be read into the Winsteps IRT program and create Winsteps⁴ control files to read the student responses and estimate parameters. A sample control file is provided in Appendix A. HumRRO created specialized SAS programs to generate all input and control files automatically. The item documentation file was used to specify item types, location, keys, item use (e.g. field test vs. operational items), and other important information. HumRRO and Pearson did not share programming or methodology for creating the input and control files. Both companies did use the same student data files (containing all student responses). HumRRO followed the guidance provided by Pearson (with input from KDE) regarding the treatment of blank responses, condition codes, etc. in creating the input data files.

Calibration and Scaling Procedures

Once input and control files were prepared, Winsteps was used to calibrate items. Multiple-choice items were fit to the Rasch measurement model and constructed-response items (short constructed response and extended response items) were fit to the Partial Credit Model (PCM). Both types of items were simultaneously calibrated in Winsteps and item difficulty parameters (logits) were produced. "Step parameters" were also produced for constructed response items. Step parameters tell us how the various points possible on the item relate to the item's overall difficulty and are important for generating scoring tables. These parameters are produced on the theta scale (a commonly used scale with a mean of 0 and a standard deviation of 1). Appendix B contains an example of item parameters for one grade subject (logits and step parameters).

Equating Procedures

Two types of equating occurred for the K-PREP: (a) forms equating within a given test administration year and (b) equating across test administration years using common anchor items. The first of these, forms equating, is accomplished by calibrating all the items for a given grade/subject together. By calibrating all the items together (i.e., across all forms), this effectively equates the various forms for a given grade/subject such that test scores on form 2 and form 3, for example, are interchangeable in terms of difficulty. For all subjects other than Science and On-Demand Writing, Kentucky uses common forms for all operational items (those

⁴ HumRRO used Winsteps version 3-73-00 for this project.

that contribute to student scores), but the forms differ on field-test items. Field-test items are used to build future forms.

In addition to the need to equate the forms of a test within a given year, there is also the need for the current year's scores to be comparable to scores from prior years. Kentucky uses a common-item anchor design to equate K-PREP scores across years. The anchor items are "internal" in the sense that they are dispersed across forms rather than externally located in a separate anchor item form. Both multiple-choice and short answer items are designated as anchor items for equating for all grades and subjects.

Raw-score-to-Scale-Score Procedures

Once the final item parameters were estimated, they were used to create scoring tables. At this stage, the scoring tables produced by the final item calibration run in Winsteps are still on the theta metric. Each potential "number of total score points (multiple-choice items correct plus total points on constructed response items)" is associated with a theta estimate. This "person-level" theta would be their score on the theta scale. Output files were verified to match between HumRRO and Pearson at this stage.

Once theta scoring tables were obtained, they were linearly transformed to a reporting scale of 100-300 for all grade subjects. Performance levels (Novice, Apprentice, Proficient, and Distinguished) were also assigned to each score. Cut scores for the performance levels were determined following a standard setting workshop conducted in the summer of 2012 (see Pearson, 2013). The results of that workshop included cut scores on the theta metric that can be used to assign NAPD categories to students. Scale score cuts were used, as opposed to theta cuts, to assign performance levels to students' scale scores. Using these cuts allowed the scale scores associated with each performance level to be fixed across test administrations. HumRRO verified the raw-score-to-scale-score tables and the associated performance levels.

In addition to overall scores, Kentucky also reports cluster scores (subscores based on subsets of items within each test). The generation of cluster scores uses the previously estimated item parameters and is accomplished by generating scoring tables in Winsteps on the theta metric, based on the specific items identified for each scoring cluster. These theta scores are then transformed in exactly the same manner as the full test scores.

Verification of 2018 Scoring Tables

After the final scoring tables were constructed, the scoring tables were applied to the 2018 student data. HumRRO checks the 2018 scored student data to verify that the scoring tables are being appropriately applied to the data and to check the distribution of students falling into each performance level. HumRRO verified Reading, Math, Social Studies, and Writing performance level distributions. HumRRO matched Pearson on the number and percent of students assigned to each performance level by subject and grade.

Documentation

As HumRRO and Pearson completed each step of the process described above, Winsteps control, item parameter, score, and output files were shared to check for inconsistencies. Winsteps output contained the number of cases in the calibration sample, item-level information

(e.g., p-values, parameters), and the theta scoring tables. A sample of the output files are appended to this document. They include:

1. Winsteps Control Files (Appendix A). These files contain the item parameter estimation specifications and important information for reading the student score files. It also specifies the output file names. The appendix includes an example control file for the initial item parameter estimation, equated item parameter estimation, and estimation of the cluster scores.
2. Winstep Item Parameter Files (Appendix B). These files contain the item parameters for the operational items. Each multiple-choice item has one parameter, a logit difficulty (named Measure in the Winstep files). Each constructed-response item has an overall difficulty parameter and a number of step parameters indicating how the points for the item are distributed along the theta scale. The file included in the appendix is an example of a final item parameter file. Initial item parameter files are in similar formats.
3. Winsteps Anchor File (Appendix C). The file includes the 2018 item parameter values for each anchor item with the equating shift estimate applied to the overall difficulty measure. The file is read by Winsteps and used to fix the item parameter values and estimate final score files.
4. Winsteps Score File (Appendix D). The file contains the raw score to theta estimation and includes the distribution of student scores.
5. Comparison of Files Output (Appendix E). This is a SAS output file from HumRRO's comparison program that checks scoring table results against Pearson's results. The files match if all comparison values are 0.

Conclusion

Pearson and HumRRO independently calculated the scaled/equated raw-score-to-scale-score tables for the 2018 K-PREP assessments. No differences were found between Pearson's and HumRRO's parameter estimation, Stocking-Lord transformation constants, or raw-score-to-scale-score tables. Given that HumRRO's and Pearson's scaling and equating results were identical, HumRRO is confident that Pearson did not commit processing errors.

References

- Bynum, B. H., & Thacker, A. A. (2011). *Third-party checking of calibration scaling and equating of the 2011 Kentucky core content test (FR-11-65)*. Alexandria, VA: Human Resources Research Organization.
- Bynum, B. H., & Thacker, A. A. (2013). *Third-party checking of 2012 scaling and equating of the for the Kentucky Performance Rating for Educational Progress (K-PREP) Tests (2013 No.11)*. Alexandria, VA: Human Resources Research Organization.
- Huynh, H. (2000, June). *Guidelines for Rasch Linking for PACT*. Memorandum to Paul Sandifer on June 18, 2000. Columbia, SC: Available from Author.
- Huynh, H., & Meyer, P. (2010). Use of robust z in detecting unstable items in item response theory models. *Practical Assessment, Research & Evaluation*, 15(2). Available online: <http://pareonline.net/getvn.asp?v=15&n=2>
- Huynh, H., & Rawls, A. (2009). A comparison between robust z and 0.3-logit difference procedures in assessing stability of linking items for the Rasch model. In Everett V. Smith Jr. & Greg E. Stone (Eds.) *Applications of Rasch Measurement in Criterion-Referenced Testing: Practice Analysis to Score Reporting*. Maple Grove, MN: JAM Press.
- Thacker, A. A., Dickinson, E. R., & Sinclair, A. L. (2013). *Policy capture for setting end-of-course and Kentucky Performance Rating for Education Progress (K-PREP) cut scores (2013 No. 007)*. Alexandria, VA: Human Resources Research Organization.
- Pearson (2012). *Kentucky performance rating for educational progress performance standards workshop: Performance level descriptor creation and standard setting, v1.1*. Pearson, Inc.

Appendix A – Control File (Math Grade 3)

```
;Winstep Control file f03MA_v0.CON
; HumRRO
&INST
Item1 = 25
NI = 47
TABLES = 00100000000001000001000000001
CODES = 012
CSV = N
FITP = 3.0
FITI = 3.0
XWIDE =1
HLINES = Y
data=f03MAmopv0.dat
IFILE= f03MAv0.ITM
ISFILE = f03MAv0.ISF
SFILE = f03MAv0.CSF
SCFILE = f03MAv0.RSS
PFILE = f03MAv0.PER
mprox=10
mucon=100
rconv=.50
lconv=.01
models=r
groups=0
stkeep=n
realse=n
stbias=n
target=n
extrsc=0.25
udecim=4
uimean=0
uscale=1
;upmean=0
;uanchor=y
ptbis=y
ILFILES = *
200322
200142
200244
202400
200532
200227
M3058
200053
M3020
202412
202405
200143
200095
```

200128
200224
M3009
M3022
200252
M3029
200126
202429
202413
200151
200021
M3025
200221
200109
202420
200115
200218
202419
200036
M3080
M3038
M3002
200017
M3006
200534
200040
200093
M3037
200323
200533
M3008
M3003
M3013
M3032
*
&END
END NAMES

Appendix B – Winsteps Item Parameter Files (Math Grade 3)

Item parameters - Math Grade 3 (f03MAv0.ITM)

```

; ITEM C:\K-PREP\K-PREP_2018\data_2018\WINSTEP\f03MAv0.con Jul 5 12:14 2018
;ENTRY MEASURE ST COUNT SCORE ERROR IN.MSQ IN.ZST OUT.MS OUT.ZS DISPL PTBISE WEIGHT OBSMA EXPMA DISCRM LOWER UPPER PVALU PBE-E RMSR G M R NAME
1 -1.1376 1 50776.0 41882.0 .0127 .95 -7.63 .85 -9.39 -.0006 .41 1.00 83.5 83.3 1.07 .10 1.00 .82 .35 .34 0 R . 200322
2 -1.7865 1 50776.0 45274.0 .0152 .96 -4.02 .80 -9.58 -.0007 .35 1.00 89.3 89.4 1.05 .26 1.00 .89 .29 .29 0 R . 200142
3 -1.5341 1 50776.0 44100.0 .0141 1.00 -.04 .94 -2.79 -.0009 .32 1.00 87.1 87.2 1.00 .23 1.00 .87 .32 .31 0 R . 200244
4 .1855 1 50776.0 31330.0 .0103 .93 -9.90 .90 -9.90 -.0004 .49 1.00 74.1 71.7 1.17 .00 1.00 .62 .43 .42 0 R . 202400
5 -1.0420 1 50776.0 41274.0 .0124 .96 -5.55 .93 -4.29 -.0006 .39 1.00 82.7 82.3 1.05 .00 1.00 .81 .36 .35 0 R . 200532
6 .2384 1 50776.0 30827.0 .0103 1.06 9.90 1.07 9.12 -.0002 .37 1.00 68.9 71.5 .86 .05 .97 .61 .43 .45 0 R . 200227
7 .2211 1 50776.0 30995.0 .0103 .99 -2.98 .97 -3.73 -.0002 .44 1.00 72.1 71.5 1.03 .01 1.00 .61 .43 .43 0 R . M3058
8 -.5272 1 50776.0 37555.0 .0112 1.01 1.90 1.03 2.13 -.0002 .38 1.00 76.7 77.0 .98 .00 1.00 .74 .39 .40 0 R . 200053
9 .5294 1 50776.0 28028.0 .0101 1.06 9.90 1.08 9.90 -.0002 .38 1.00 68.0 70.5 .84 .07 .98 .55 .43 .45 0 R . M3020
10 -.4164 1 50776.0 36666.0 .0110 .91 -9.90 .86 -9.90 -.0004 .48 1.00 78.6 76.0 1.15 .00 1.00 .72 .40 .38 0 R . 202412
11 -1.3275 1 50776.0 42999.0 .0133 .94 -8.12 .87 -7.69 -.0006 .39 1.00 85.7 85.2 1.07 .00 1.00 .85 .33 .32 0 R . 202405
12 1.7627 1 50776.0 16386.0 .0108 .93 -9.90 1.00 .44 .0002 .43 1.00 78.0 75.0 1.09 .00 1.00 .32 .41 .40 0 R . 200143
13 .0062 1 50776.0 32986.0 .0105 .90 -9.90 .83 -9.90 .0000 .51 1.00 75.9 72.7 1.22 .00 1.00 .65 .42 .40 0 R . 200095
14 -.8618 1 50776.0 40064.0 .0119 .94 -9.90 .89 -8.00 -.0008 .43 1.00 81.4 80.4 1.09 .00 1.00 .79 .37 .36 0 R . 200128
15 -1.0736 1 50776.0 41481.0 .0125 1.03 4.63 1.14 8.20 -.0010 .31 1.00 83.0 82.6 .94 .00 .99 .82 .35 .36 0 R . 200224
16 .9052 1 50776.0 24356.0 .0101 .95 -9.90 .95 -7.04 .0000 .46 1.00 73.1 70.5 1.11 .00 1.00 .48 .43 .43 0 R . M3009
17 .4703 1 50776.0 28604.0 .0102 1.09 9.90 1.12 9.90 -.0003 .35 1.00 66.9 70.6 .77 .08 .96 .56 .43 .46 0 R . M3022
18 .9982 1 50776.0 23453.0 .0102 1.06 9.90 1.09 9.90 .0003 .37 1.00 68.7 70.7 .85 .05 .99 .46 .43 .45 0 R . 200252
19 .7942 1 50776.0 25443.0 .0101 1.18 9.90 1.25 9.90 -.0001 .28 1.00 63.5 70.4 .54 .14 .93 .50 .43 .48 0 R . M3029
20 -.5056 1 50776.0 37390.0 .0112 .94 -9.90 .90 -8.73 -.0006 .45 1.00 78.4 76.8 1.11 .00 1.00 .74 .39 .38 0 R . 200126
21 -.7075 1 50776.0 38946.0 .0116 .96 -6.37 .90 -8.19 -.0005 .42 1.00 79.5 78.8 1.06 .00 1.00 .77 .38 .38 0 R . 202429
22 .4701 1 50776.0 28606.0 .0102 .99 -3.16 1.00 -.36 -.0002 .44 1.00 72.1 70.7 1.02 .00 1.00 .56 .43 .43 0 R . 202413
23 .2432 1 50776.0 30785.0 .0103 1.07 9.90 1.09 9.90 -.0002 .37 1.00 68.6 71.4 .83 .07 .97 .61 .43 .45 0 R . 200151
24 1.9461 1 50776.0 14848.0 .0111 1.17 9.90 1.48 9.90 .0005 .23 1.00 73.4 76.5 .66 .07 .94 .29 .40 .43 0 R . 200021
25 .9887 1 50776.0 23548.0 .0102 1.12 9.90 1.18 9.90 .0002 .33 1.00 66.3 70.7 .69 .07 .94 .46 .43 .46 0 R . M3025
26 .6318 1 50776.0 54640.0 .0063 1.04 6.97 1.09 6.76 -.0001 .58 1.00 55.4 56.0 .96 .00 1.96 1.08 .60 .72 0 R . 200221
27 1.5589 1 50776.0 42853.0 .0095 .92 -9.90 .91 -9.90 .0017 .52 1.00 73.2 71.4 1.09 .00 2.00 .84 .45 .45 0 R . 200109
28 .3497 1 50776.0 29764.0 .0102 1.20 9.90 1.44 9.90 -.0002 .26 1.00 64.8 71.0 .45 .08 .87 .59 .43 .48 0 R . 202420
29 -.9843 1 50776.0 40898.0 .0123 .90 -9.90 .81 -9.90 -.0008 .45 1.00 82.9 81.7 1.13 .00 1.00 .81 .36 .34 0 R . 200115
30 -1.3717 1 50776.0 43246.0 .0135 .90 -9.90 .80 -9.90 -.0007 .42 1.00 86.4 85.6 1.10 .00 1.00 .85 .33 .31 0 R . 200218
31 .1275 1 50776.0 31869.0 .0104 .86 -9.90 .80 -9.90 -.0002 .55 1.00 77.8 72.0 1.32 .00 1.00 .63 .42 .40 0 R . 202419
32 -1.1067 1 50776.0 41685.0 .0126 .94 -8.61 .89 -7.22 -.0001 .41 1.00 83.6 82.9 1.07 .00 1.00 .82 .35 .34 0 R . 200036
33 1.2104 1 50776.0 21411.0 .0103 1.13 9.90 1.21 9.90 .0001 .32 1.00 67.4 71.4 .69 .07 .94 .42 .43 .46 0 R . M3080
34 .1463 1 50776.0 31695.0 .0103 .89 -9.90 .84 -9.90 -.0001 .52 1.00 76.0 71.9 1.25 .00 1.00 .62 .43 .41 0 R . M3038
35 .3433 1 50776.0 29829.0 .0102 1.03 7.58 1.02 2.62 -.0002 .40 1.00 69.6 71.0 .93 .05 1.00 .59 .43 .44 0 R . M3002
36 -.8252 1 50776.0 39802.0 .0119 .95 -7.69 .88 -9.22 -.0003 .42 1.00 80.9 80.0 1.07 .00 1.00 .78 .37 .37 0 R . 200017
37 .4335 1 50776.0 28961.0 .0102 .98 -5.30 .95 -7.04 .0001 .45 1.00 71.7 70.8 1.06 .02 1.00 .57 .43 .43 0 R . M3006
38 1.6735 1 50776.0 17159.0 .0107 1.10 9.90 1.26 9.90 .0007 .31 1.00 72.7 74.3 .77 .05 .98 .34 .41 .44 0 R . 200534
39 -.6026 1 50776.0 38148.0 .0114 .93 -9.90 .87 -9.90 -.0008 .45 1.00 79.6 77.8 1.11 .00 1.00 .75 .39 .38 0 R . 200040
40 -.3173 1 50776.0 35838.0 .0109 1.03 5.68 .98 -2.20 -.0004 .38 1.00 73.6 75.1 .97 .07 1.00 .71 .41 .41 0 R . 200093
41 .2838 1 50776.0 30395.0 .0102 .92 -9.90 .87 -9.90 .0001 .50 1.00 74.1 71.3 1.20 .00 1.00 .60 .43 .42 0 R . M3037
42 -.0597 1 50776.0 33582.0 .0105 1.04 9.36 1.04 3.85 -.0006 .38 1.00 71.2 73.1 .92 .08 .99 .66 .42 .43 0 R . 200323
43 -.8732 1 50776.0 40139.0 .0120 .98 -3.75 .95 -3.63 -.0007 .39 1.00 81.1 80.5 1.03 .00 1.00 .79 .37 .37 0 R . 200533
44 1.1608 1 50776.0 21886.0 .0102 1.09 9.90 1.15 9.90 .0002 .35 1.00 68.6 71.2 .78 .06 .97 .43 .43 .45 0 R . M3008
45 -.3542 1 50776.0 36147.0 .0109 .95 -9.41 .91 -8.41 -.0006 .45 1.00 76.5 75.5 1.09 .00 1.00 .71 .40 .40 0 R . M3003
46 -.4364 1 50776.0 36829.0 .0111 .93 -9.90 .87 -9.90 -.0003 .47 1.00 78.3 76.2 1.13 .00 1.00 .73 .40 .39 0 R . M3013
47 .1724 1 50776.0 63617.0 .0071 .97 -4.51 .99 -.61 -.0004 .57 1.00 57.4 56.1 1.03 .00 2.00 1.25 .55 .62 0 R . M3032

```

Step parameters 2018 - Math Grade 3 (f03MAv0.CSF)

; STRUCTURE MEASURE ANCHOR FILE FOR C:\K-PREP\K-PREP_2018\data_2018\WINSTEP\f03MAv0.con Jul 5
12:14 2018

; ITEM CATEGORY Rasch-Andrich threshold MEASURE

1	0	.0000
1	1	.0000
2	0	.0000
2	1	.0000
3	0	.0000
3	1	.0000
4	0	.0000
4	1	.0000
5	0	.0000
5	1	.0000
6	0	.0000
6	1	.0000
7	0	.0000
7	1	.0000
8	0	.0000
8	1	.0000
9	0	.0000
9	1	.0000
10	0	.0000
10	1	.0000
11	0	.0000
11	1	.0000
12	0	.0000
12	1	.0000
13	0	.0000
13	1	.0000
14	0	.0000
14	1	.0000
15	0	.0000
15	1	.0000
16	0	.0000
16	1	.0000
17	0	.0000
17	1	.0000
18	0	.0000
18	1	.0000
19	0	.0000
19	1	.0000
20	0	.0000
20	1	.0000
21	0	.0000
21	1	.0000
22	0	.0000
22	1	.0000
23	0	.0000
23	1	.0000
24	0	.0000
24	1	.0000
25	0	.0000
25	1	.0000
26	0	.0000

26	1	1.0234
26	2	-1.0234
27	0	.0000
27	1	-2.2045
27	2	2.2045
28	0	.0000
28	1	.0000
29	0	.0000
29	1	.0000
30	0	.0000
30	1	.0000
31	0	.0000
31	1	.0000
32	0	.0000
32	1	.0000
33	0	.0000
33	1	.0000
34	0	.0000
34	1	.0000
35	0	.0000
35	1	.0000
36	0	.0000
36	1	.0000
37	0	.0000
37	1	.0000
38	0	.0000
38	1	.0000
39	0	.0000
39	1	.0000
40	0	.0000
40	1	.0000
41	0	.0000
41	1	.0000
42	0	.0000
42	1	.0000
43	0	.0000
43	1	.0000
44	0	.0000
44	1	.0000
45	0	.0000
45	1	.0000
46	0	.0000
46	1	.0000
47	0	.0000
47	1	-.4215
47	2	.4215

Appendix C – Winsteps Anchor File (Grade 3 Reading)

Item Anchor File (G03RDv0anchors.txt IAF)

```
30 -1.2469
23 -.6402
28 -.8052
24 -.3949
26 -.8520
 1 -.5881
10 .4559
32 .6463
17 .9736
18 -.1466
19 .2362
```

Step Parameter Anchor File (G03RDv0anchors.SAF)

```
; ITEM CATEGORY Rasch-Andrich threshold MEASURE
 1 0 0.0000
 1 1 0.0000
10 0 0.0000
10 1 -0.7101
10 2 0.7100
17 0 0.0000
17 1 0.0000
18 0 0.0000
18 1 0.0000
19 0 0.0000
19 1 0.0000
23 0 0.0000
23 1 0.0000
24 0 0.0000
24 1 0.0000
26 0 0.0000
26 1 0.0000
28 0 0.0000
28 1 0.0000
30 0 0.0000
30 1 0.0000
32 0 0.0000
32 1 0.0000
```


Appendix D – Winsteps Score File (Grade 3 reading)

PERSON SCORE FILE FOR C:\K-PREP\K-PREP_2018\data_2018\WINSTEP\g03RDv0.con Jul 12 11:50 2018
USCALE=1.00

SCORE	MEASURE	S.E.	INFO	NORMED	S.E.	FREQUENCY	%	CUM.FREQ.	%	PERCENTILE
0	-5.3568	2.0099	.25	-39	175	163	.3	163	.3	1
1	-3.9406	1.0199	.96	84	89	4	.0	167	.3	1
2	-3.2070	.7355	1.85	148	64	6	.0	173	.3	1
3	-2.7602	.6124	2.67	187	53	13	.0	186	.4	1
4	-2.4304	.5408	3.42	216	47	44	.1	230	.5	1
5	-2.1644	.4933	4.11	239	43	74	.1	304	.6	1
6	-1.9383	.4593	4.74	259	40	126	.2	430	.8	1
7	-1.7394	.4337	5.32	276	38	205	.4	635	1.3	1
8	-1.5601	.4138	5.84	292	36	324	.6	959	1.9	2
9	-1.3955	.3981	6.31	306	35	416	.8	1375	2.7	2
10	-1.2422	.3854	6.73	320	34	521	1.0	1896	3.7	3
11	-1.0978	.3750	7.11	332	33	610	1.2	2506	4.9	4
12	-.9604	.3666	7.44	344	32	706	1.4	3212	6.3	6
13	-.8286	.3597	7.73	356	31	791	1.6	4003	7.9	7
14	-.7013	.3542	7.97	367	31	899	1.8	4902	9.7	9
15	-.5775	.3497	8.18	378	31	948	1.9	5850	11.5	11
16	-.4564	.3463	8.34	388	30	1033	2.0	6883	13.6	13
17	-.3374	.3438	8.46	399	30	1106	2.2	7989	15.7	15
18	-.2198	.3421	8.55	409	30	1292	2.5	9281	18.3	17
19	-.1032	.3412	8.59	419	30	1483	2.9	10764	21.2	20
20	.0131	.3410	8.60	429	30	1546	3.0	12310	24.2	23
21	.1296	.3416	8.57	440	30	1719	3.4	14029	27.6	26
22	.2468	.3430	8.50	450	30	1742	3.4	15771	31.1	29
23	.3651	.3451	8.39	460	30	1820	3.6	17591	34.6	33
24	.4852	.3481	8.25	471	30	2041	4.0	19632	38.7	37
25	.6077	.3520	8.07	481	31	2282	4.5	21914	43.2	41
26	.7333	.3569	7.85	492	31	2319	4.6	24233	47.7	45
27	.8627	.3629	7.59	504	32	2532	5.0	26765	52.7	50
28	.9970	.3702	7.30	515	32	2552	5.0	29317	57.7	55
29	1.1373	.3791	6.96	528	33	2601	5.1	31918	62.9	60
30	1.2850	.3899	6.58	540	34	2615	5.2	34533	68.0	65
31	1.4420	.4030	6.16	554	35	2709	5.3	37242	73.3	71
32	1.6108	.4192	5.69	569	37	2613	5.1	39855	78.5	76
33	1.7949	.4395	5.18	585	38	2627	5.2	42482	83.7	81
34	1.9991	.4654	4.62	603	41	2490	4.9	44972	88.6	86
35	2.2312	.4996	4.01	623	44	2072	4.1	47044	92.7	91
36	2.5039	.5473	3.34	647	48	1597	3.1	48641	95.8	94
37	2.8410	.6187	2.61	676	54	1131	2.2	49772	98.0	97
38	3.2961	.7414	1.82	716	65	668	1.3	50440	99.3	99
39	4.0389	1.0246	.95	781	89	267	.5	50707	99.9	99
40	5.4627	2.0125	.25	905	176	69	.1	50776	100.0	99

Appendix E – Comparison of Files Output (Math Grade 3)

Item Parameter Differences – Math Grade 3

obs	NAME	MEASURE_diff	Rasch1_diff	Rasch2_diff	Rasch3_diff	Rasch4_diff	Rasch5_diff
1	200017	0	0	0	.	.	.
2	200021	0	0	0	.	.	.
3	200036	0	0	0	.	.	.
4	200040	0	0	0	.	.	.
5	200053	0	0	0	.	.	.
6	200093	0	0	0	.	.	.
7	200095	0	0	0	.	.	.
8	200109	0	0	0	0	.	.
9	200115	0	0	0	.	.	.
10	200126	0	0	0	.	.	.
11	200128	0	0	0	.	.	.
12	200142	0	0	0	.	.	.
13	200143	0	0	0	.	.	.
14	200151	0	0	0	.	.	.
15	200218	0	0	0	.	.	.
16	200221	0	0	0	0	.	.
17	200224	0	0	0	.	.	.
18	200227	0	0	0	.	.	.
19	200244	0	0	0	.	.	.
20	200252	0	0	0	.	.	.
21	200322	0	0	0	.	.	.
22	200323	0	0	0	.	.	.
23	200532	0	0	0	.	.	.
24	200533	0	0	0	.	.	.
25	200534	0	0	0	.	.	.
26	202400	0	0	0	.	.	.
27	202405	0	0	0	.	.	.
28	202412	0	0	0	.	.	.
29	202413	0	0	0	.	.	.
30	202419	0	0	0	.	.	.
31	202420	0	0	0	.	.	.
32	202429	0	0	0	.	.	.
33	M3002	0	0	0	.	.	.
34	M3003	0	0	0	.	.	.
35	M3006	0	0	0	.	.	.
36	M3008	0	0	0	.	.	.
37	M3009	0	0	0	.	.	.
38	M3013	0	0	0	.	.	.
39	M3020	0	0	0	.	.	.
40	M3022	0	0	0	.	.	.
41	M3025	0	0	0	.	.	.
42	M3029	0	0	0	.	.	.
43	M3032	0	0	0	0	.	.
44	M3037	0	0	0	.	.	.
45	M3038	0	0	0	.	.	.
46	M3058	0	0	0	.	.	.
47	M3080	0	0	0	.	.	.

All RSSS Differences – Math Grade 3

Obs	rs	theta_ diff	SS_ diff	PL_ diff
1	0	0	0	0
2	1	0	0	0
3	2	0	0	0
4	3	0	0	0
5	4	0	0	0
6	5	0	0	0
7	6	0	0	0
8	7	0	0	0
9	8	0	0	0
10	9	0	0	0
11	10	0	0	0
12	11	0	0	0
13	12	0	0	0
14	13	0	0	0
15	14	0	0	0
16	15	0	0	0
17	16	0	0	0
18	17	0	0	0
19	18	0	0	0
20	19	0	0	0
21	20	0	0	0
22	21	0	0	0
23	22	0	0	0
24	23	0	0	0
25	24	0	0	0
26	25	0	0	0
27	26	0	0	0
28	27	0	0	0
29	28	0	0	0
30	29	0	0	0
31	30	0	0	0
32	31	0	0	0
33	32	0	0	0
34	33	0	0	0
35	34	0	0	0
36	35	0	0	0
37	36	0	0	0
38	37	0	0	0
39	38	0	0	0
40	39	0	0	0
41	40	0	0	0
42	41	0	0	0
43	42	0	0	0
44	43	0	0	0
45	44	0	0	0
46	45	0	0	0
47	46	0	0	0
48	47	0	0	0
49	48	0	0	0
50	49	0	0	0
51	50	0	0	0