



Item Response Theory Guidelines -NEA 2021 Using Conquest



Bhutan Council for School Examinations and Assessment
Royal Government of Bhutan

Item Response Theory Guidelines -NEA 2021 Using Conquest



**Assessment and Monitoring Division
Bhutan Council for School Examinations and Assessment
Thimphu: Bhutan
2021**

Copyright

Assessment and Monitoring Division
Bhutan Council for School Examinations and Assessment, 2021
www.bcsea.bt

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of publisher.

This NEA Item Response Theory (IRT) Guidelines was developed by the Assessment and Monitoring Division (AMD) of Bhutan Council for School Examinations and Assessment (BCSEA) in collaboration with the Australian Council for Educational Research (ACER), India under NEA project.

The project is funded by Global Partnership for Education (GPE) under Education Sector Programme Implementation Grant (ESPIG) with support from United Nations International Children's Fund (UNICEF) and Save the Children International-Bhutan (SCI).

Acknowledgement

Bhutan Council for School Examinations and Assessment acknowledges the invaluable contributions made by the following officials in developing this guideline.

Authors:

1. Sonam Lhamo, Education Monitoring Officer, AMD, BCSEA
2. Karma Jigme Lepcha, Examination and Assessment Officer, SED, BCSEA
3. Kinley Dema, Principal Education Monitoring Officer, AMD, BCSEA
4. Sharda Rai, Examination and Assessment Officer, SED, BCSEA
5. Sapna Subba, Examination and Assessment Officer, SED, BCSEA
6. Arjun Kumar Gurung, Project Manager, NEA project, BCSEA

Logistic and Administrative

1. Sangay Tenzin, CPO, BCSEA
2. Bhim Kumar Pradhan, Project finance, BCSEA
3. Kezang Dema, Adim Assistant, BCSEA

Contents

1. Introduction	6
2. Data Analysis Processes	6
2.1 Creating folder	6
2.2 Data Cleaning	8
2.3 Check for overall data inconsistencies	10
2.4 Run1 (Form by Form Analysis)	13
2.4.1 Preparation of CQC	13
2.4.2 Run1B	17
2.5 Equating	18
2.6 Run 2 (Concurrent Analysis)	20
2.6.1 Update CQC file	22
2.7 Run 3 (Case Estimation Analysis)	24
2.8 Determining bands/cuts for the proficiency scale	27
2.9 Developing descriptors of the bands	29
2.10 Establishing baseline	30

1. Introduction

The National Education Assessment Item Response Theory (NEA IRT) Guidelines is a document that spells out the step by step processes required to be followed for the analysis of both cognitive items and reporting of the findings from the assessment. These processes are technically robust to guide the analysis of cognitive data for all NEA cycles. Further, the technical expertise rendered by ACER, India ensures the reliability and validity of the processes as well as the outputs. The data was run using ConQuest 5, SPSS version 22 and Microsoft Excel.

This document presents data analysis processes such as creating folder, data cleaning, checking for overall data inconsistencies, form by form analysis, Run1, equating, Run2 (concurrent analysis), update CQC file, Run3 (case estimation analysis), determining bands for the proficiency scale, developing descriptors of the bands and establishing baseline.

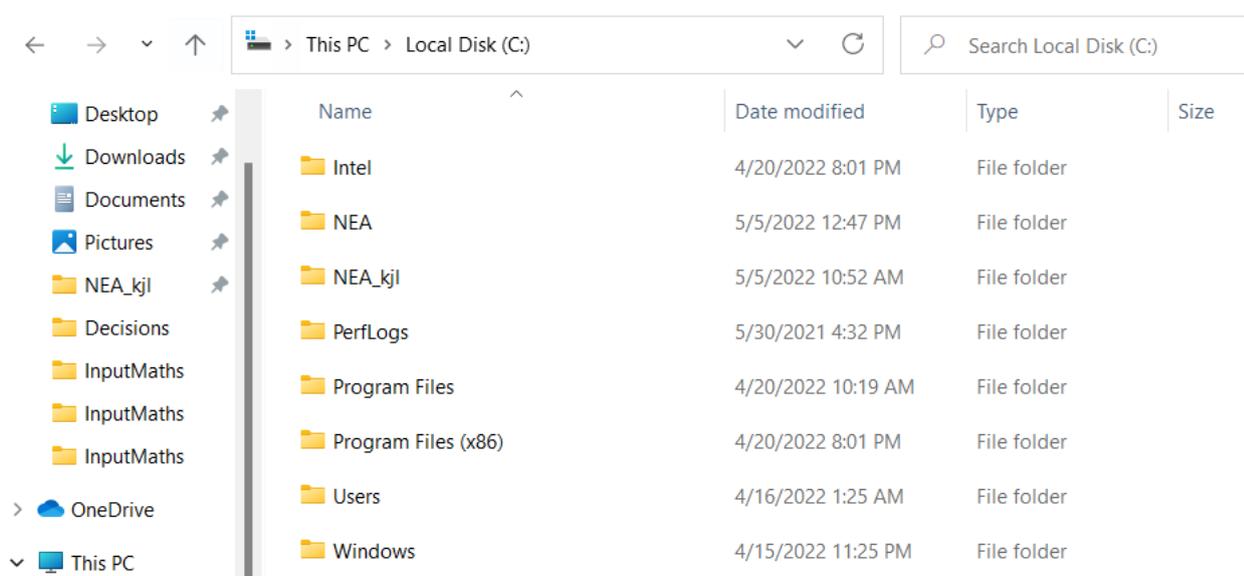
Objectives:

- To state explicitly the steps required for the analysis of NEA cognitive data;
- To determine bands and develop proficiency scale for each domain;
- To establish baseline for each domain;
- To ensure that the processes of NEA cognitive data analysis are technically robust;
- To generate accurate outputs of the academic performance for reporting; and
- To provide evidence based recommendations to the relevant stakeholders.

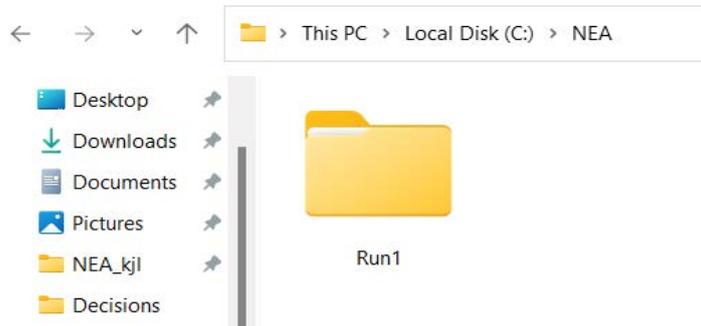
2. Data Analysis Processes

2.1 Creating folder

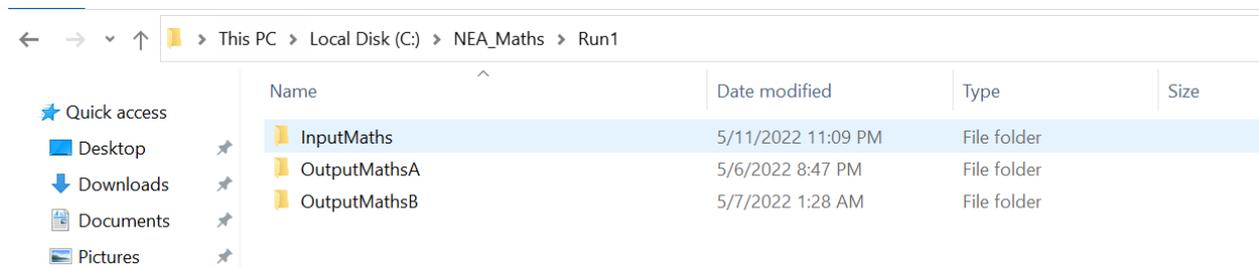
- Create a folder in C drive with folder name **NEA** so that the entire working directory can be transferred to different windows PC for analysis if not the file pathways would be lost.



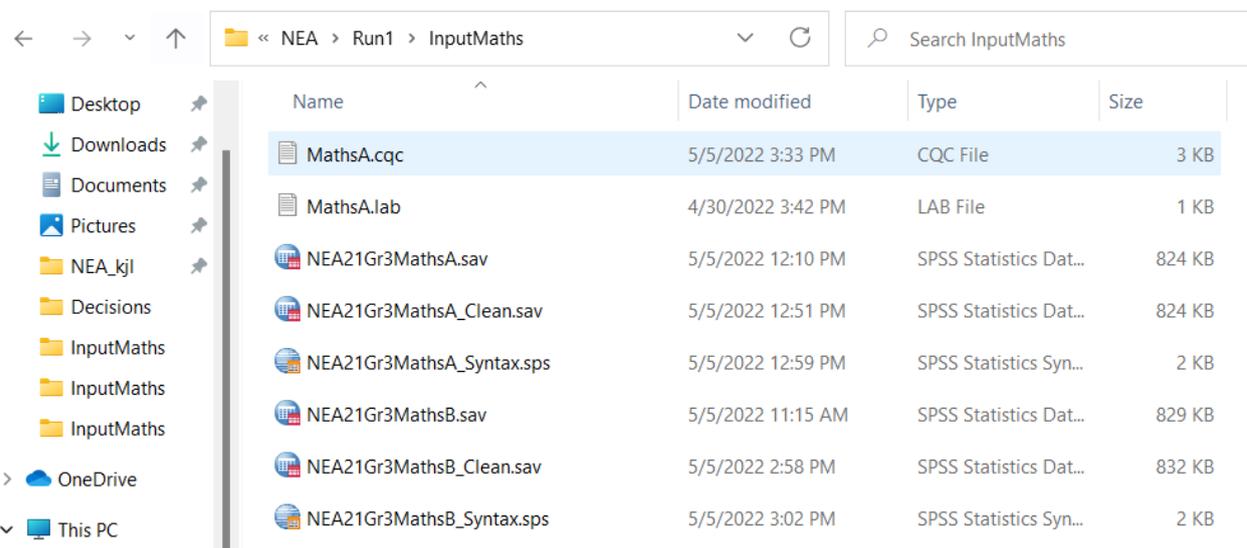
- Create a folder inside the NEA folder with a name **Run1** (to run data for all the domains).



- Create Input Folder (**InputMaths**) and 2 Output Folders (**OutputMathsA** and **OutputMathsB**) inside Run1 Folder.



- Insert the Maths A and B SPSS data sav files, and MathsA CQC and lab files in the Input Maths folder.



2.2 Data Cleaning

Data cleaning is the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, corrupted, inconsistent, irrelevant, duplicate, out of range data, missing or improperly formatted within a data set. This process enhances the quality of the data ensuring a reliable output.

- Open SPSS programme, click file, open, data and paste the [NEA21Gr3MathsA.sav](#) file from the InputMaths.
- Select the syntax and run.

GET

```
FILE='C:\NEA\Run1\InputMaths\NEA21Gr3MathsA.sav'.
```

```
DATASET NAME DataSet1 WINDOW=FRONT.
```

- Check duplicate data in the data view by clicking on the data and sort cases. Select only std ID and paste and run.

```
DATASET ACTIVATE DataSet1.
```

```
SORT CASES BY StudentID(A).
```

- In the data view of the dataset, go to data, click identify duplicate cases and select student ID and paste.
- Select the syntax of the duplicate cases and run.

```
* Identify Duplicate Cases.
```

```
SORT CASES BY StudentID(A).
```

```
MATCH FILES
```

```
  /FILE=*
```

```
  /BY StudentID
```

```
  /FIRST=PrimaryFirst
```

```
  /LAST=PrimaryLast3.
```

```
DO IF (PrimaryFirst).
```

```
  COMPUTE MatchSequence=1-PrimaryLast3.
```

```
ELSE.
```

```
  COMPUTE MatchSequence=MatchSequence+1.
```

```
END IF.
```

```
LEAVE MatchSequence.
```

```
FORMATS MatchSequence (f7).
```

```
COMPUTE InDupGrp=MatchSequence>0.
```

```
SORT CASES InDupGrp(D).
```

```
MATCH FILES
```

```
  /FILE=*
```

```
  /DROP=PrimaryFirst InDupGrp MatchSequence.
```

```
VARIABLE LABELS PrimaryLast3 'Indicator of each last matching case as Primary'.
```

```
VALUE LABELS PrimaryLast3 0 'Duplicate Case' 1 'Primary Case'.
```

```
VARIABLE LEVEL PrimaryLast3 (ORDINAL).
```

```
FREQUENCIES VARIABLES=PrimaryLast3.
```

```
EXECUTE.
```


2.3 Check for overall data inconsistencies

- In the data view of the dataset, click analyse, descriptive statistics, frequencies. Select only the items and paste.

```
FREQUENCIES VARIABLES=MA20G3NA01 MA20G3NA42 MA20G3NA06 MA20G3NA73 MA20G3GE45
MA20G3NA37 MA20G3NA50
```

```
MA20G3MA13 MA20G3NA29 MA20G3DP19 MA20G3MA12 MA20G3MA11 MA20G3NA36 MA20G-
3GE84 MA20G3NA33 MA20G3MA56
```

```
MA20G3GE66 MA20G3NA03 MA20G3NA65 MA20G3NA81 MA20G3GE44 MA20G3NA51 MA20G-
3NA04 MA20G3MA57 MA20G3NA38
```

```
MA20G3DP61 MA20G3GE83 MA20G3NA34 MA20G3NA10 MA20G3NA79
```

```
/ORDER=ANALYSIS.
```

- Select the frequency syntax and run and check the output. Scan the frequencies of the responses for each item and identify the errors (invalid codes/missing code (blanks and lower cases etc). The two images show the errors such as missing values and inconsistent entry.

MA20G3NA29

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	.0	.0	.0
7	3	.1	.1	.2
8	14	.6	.6	.8
9	136	5.9	5.9	6.7
A	1116	48.3	48.3	55.0
B	570	24.7	24.7	79.6
C	309	13.4	13.4	93.0
D	162	7.0	7.0	100.0
Total	2311	100.0	100.0	

MA20G3DP19

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 7	3	.1	.1	.1
8	9	.4	.4	.5
9	7	.3	.3	.8
A	335	14.5	14.5	15.3
B	158	6.8	6.8	22.2
C	998	43.2	43.2	65.3
d	1	.0	.0	65.4
D	800	34.6	34.6	100.0
Total	2311	100.0	100.0	

- In the syntax file, click transform, recode into the same variables, select items one by one with issues and recode the correct codes for the response codes with errors by clicking and entering in the old and new values. For ex. For shortcut, in the syntax file, type

```
RECODE MA20G3NA29 MA20G3NA10 ("9"="9").
```

```
EXECUTE.
```

```
RECODE MA20G3DP19 ("d"="D").
```

```
EXECUTE.
```

- Then select all the recoded response syntax and run, then in the data set, click analyse, descriptive statistics and frequencies. Select all the items and paste. In the syntax file, highlight the frequencies and run. Then, verify in the outputs to find if update is being done.
- Frequency before recode:

MA20G3DP19

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 7	3	.1	.1	.1
8	9	.4	.4	.5
9	7	.3	.3	.8
A	335	14.5	14.5	15.3
B	158	6.8	6.8	22.2
C	998	43.2	43.2	65.3
d	1	.0	.0	65.4
D	800	34.6	34.6	100.0
Total	2311	100.0	100.0	

- Frequency after recode:

MA20G3DP19

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 7	3	.1	.1	.1
8	9	.4	.4	.5
9	7	.3	.3	.8
A	335	14.5	14.5	15.3
B	158	6.8	6.8	22.1
C	999	43.2	43.2	65.4
D	801	34.6	34.6	100.0
Total	2312	100.0	100.0	

- Type the paths of the Input folder in C drive and the name of the clean data set of MathsA in the syntax file.

```
SAVE OUTFILE='C:\NEA\Run1\InputMaths\NEA21Gr3MathsA_Clean.sav'  
/COMPRESSED.
```

(Note: back slash and forward slashes must be used appropriately)

- Save the syntax file with proper name (**NEA21Gr3MathsA_Syntax.sps**) in the InputMaths folder.
- Save the dataset file as **NEA21Gr3MathsA_Clean.sav** in the MathsInput folder.
- With this, the dataset of Maths booklet A is cleaned and ready for item analysis.
- Repeat the same process for Maths Booklet B for which a new syntax file needs to be created. To create syntax for Booklet B, open the syntax file of booklet A, click file, new and then syntax).
- Copy the syntax file of Maths Booklet A and paste it in the new syntax file for Booklet B.

- Close the Maths Booklet A syntax file and start working on the new syntax file of Booklet B.
- In the beginning (top) of the syntax file, type **DATASET CLOSE ALL.** to close all the open data sets if there are any.

```
DATASET CLOSE ALL.
```

- Rename the file name (change from A to B).

```
GET FILE='C:\NEA\Run1\InputMaths\NEA21Gr3MathsB.sav'.
```

- Select the following and run.

```
DATASET CLOSE ALL.
```

```
GET FILE='C:\NEA\Run1\InputMaths\NEA21Gr3MathsB.sav'.
```

```
DATASET NAME DataSet1 WINDOW=FRONT.
```

- Check the duplicates by highlighting the syntax from 'dataset activate' to identify duplicate cases' till duplicate case execute and run.

```
DATASET ACTIVATE DataSet1.
```

```
SORT CASES BY StudentID(A).
```

```
* Identify Duplicate Cases.
```

```
SORT CASES BY StudentID(A).
```

```
MATCH FILES
```

```
  /FILE=*
```

```
  /BY StudentID
```

```
  /DROP = PrimaryLast3  /FIRST=PrimaryFirst
```

```
  /LAST=PrimaryLast3.
```

```
DO IF (PrimaryFirst).
```

```
  COMPUTE MatchSequencel=1-PrimaryLast3.
```

```
  ELSE.
```

```
  COMPUTE MatchSequencel=MatchSequencel+1.
```

```
END IF.
```

```
LEAVE MatchSequencel.
```

```
FORMATS MatchSequencel (f7).
```

```
COMPUTE InDupGrp=MatchSequencel>0.
```

```
SORT CASES InDupGrp(D).
```

```
MATCH FILES
```

```
  /FILE=*
```

```
  /DROP=PrimaryFirst InDupGrp MatchSequencel.
```

```
VARIABLE LABELS PrimaryLast3 'Indicator of each last matching case as Primary'.
```

```
VALUE LABELS PrimaryLast3 0 'Duplicate Case' 1 'Primary Case'.
```

```
VARIABLE LEVEL PrimaryLast3 (ORDINAL).
```

```
FREQUENCIES VARIABLES=PrimaryLast3.
```

```
EXECUTE.
```

- Select the item variables and delete them in order to insert the item variables of MathsB for which copy the item variable names from the variable view of the MathsB and paste them by keeping one space between each variable.
- Select the item variables in the syntax file and run and then in the dataset file click analyse, descriptive statistics and frequencies, then paste. Select the variables in the syntax file and then run.
- In the output file, check for errors in the data and do the recoding (as done for MathsA). Check the frequencies by rerunning to see if the codes are updated.
- Type the paths of the Input folder in C drive and the name of the MathsB clean dataset file in the syntax file.

```
SAVE OUTFILE='C:\NEA\Run1\InputMaths\NEA21Gr3MathsB_Clean.sav'  
/COMPRESSED.
```

(Note: back slash and forward slashes must be used appropriately)

- Save the dataset file in the MathsInput folder and name it NEA21Gr3MathsB_Clean.sav.
- Save the syntax file in the MathsInput folder and name it as

```
NEA21Gr3MathsB_Syntax.sps.
```

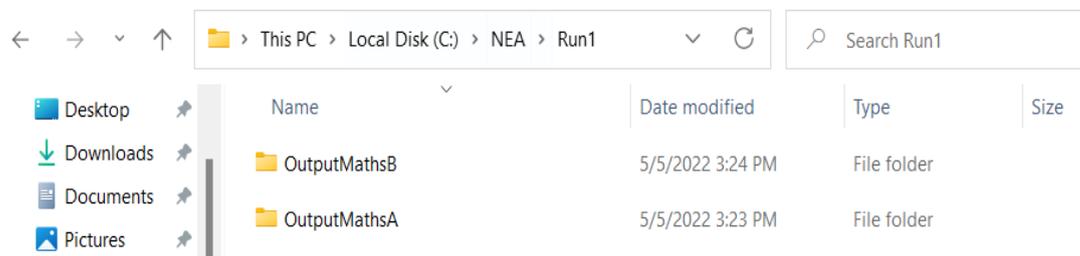
2.4 Run1 (Form by Form Analysis)

2.4.1 Preparation of CQC

- Check the titles in the CQC file and ensure that they are correct.
- Ensure the location of the files are correct (both input and output). To insert the input location, go to the folder where the input file is and select the path and paste it in the CQC file input.

```
let input=C:\NEA\Run1\InputMaths;  
let output=C:\NEA\Run1\OutputMathsA;
```

- Insert the output location by copying the path of the Output folder and pasting it in the CQC 'output'.



- In the 'let name' ensure the default name is correct for each booklet (just check MathsA or B).

```
let name=G3MathsA_Run1;
```

- Check and write the correct 'data file name' to track the path where the MathsA clean sav file is stored and also copy and paste the name of the MathsA clean save file.

```
(C:\NEA\Run1\InputMaths\NEA21Gr3MathsA_Clean.sav !)
```

- Check the name and sequence of the item variables as in the variable view in SPSS.

```
responses= MA20G3NA01 MA20G3NA42 MA20G3NA06 MA20G3NA73 MA20G3GE45 MA20G3NA37
MA20G3NA50 MA20G3MA13 MA20G3NA29 MA20G3DP19 MA20G3MA12 MA20G3MA11 MA20G3NA36
MA20G3GE84 MA20G3NA33 MA20G3MA56 MA20G3GE66 MA20G3NA03 MA20G3NA65 MA20G3NA81
MA20G3GE44 MA20G3NA51 MA20G3NA04 MA20G3MA57 MA20G3NA38 MA20G3DP61 MA20G3GE83
MA20G3NA34 MA20G3NA10 MA20G3NA79,
```

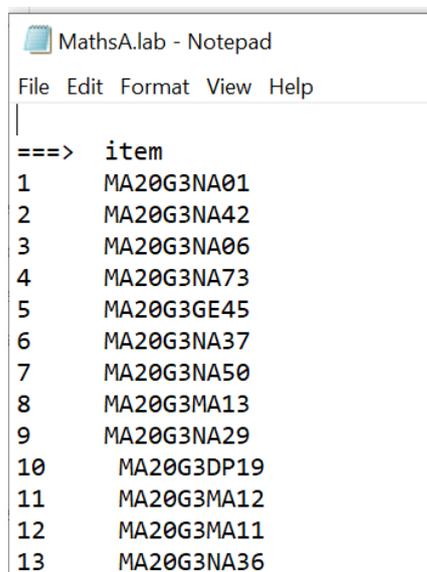
- PID (person ID of each student) need not be changed.

```
pid= StudentID;
```

- Labels- mention the location of the lab file (`MathsA.lab` to be retained)

```
labels << C:\NEA\Run1\InputMaths\MathsA.lab;
```

- Open the Maths lab A and delete all the items (retain first line ==> `item`)
- Open the SPSS Variable view, copy all the item IDs and paste in the Maths A lab file. Give numbering to all the item ID variables and keep uniform spacing between the item ID and the serial numbers and save it.



```
MathsA.lab - Notepad
File Edit Format View Help
===> item
1 MA20G3NA01
2 MA20G3NA42
3 MA20G3NA06
4 MA20G3NA73
5 MA20G3GE45
6 MA20G3NA37
7 MA20G3NA50
8 MA20G3MA13
9 MA20G3NA29
10 MA20G3DP19
11 MA20G3MA12
12 MA20G3MA11
13 MA20G3NA36
```

- Insert only the valid 'codes' A, B, C, D, 0, 1 (no partial credit for Maths) in MathsA CQC syntax. Codes 7, 8 and 9 are excluded.

```
codes A,B,C,D,0,1;
```

- Insert the 'key' for every item based on the codebook for MathA. Zero need not be mentioned as one of the keys because it will be automatically identified as the wrong answer by the system.

```
/*          10          20          30          40          50          60          70          80*
/* 1234567890123456789012345678901234567890123456789012345678901234567890*/
/* mmmmmmmmsmmmmmmmsmmmmmmmsmmmmmmmsmmmmmmmsmmmmmmmsmmmmmmmsmmmmmmmsmmmmmmms*/
KEY 1DBCBC1AACABAAB1CB1DBCBCACDBD1B!1;
```

- In case of partial credit model, type + item*step in the 'model item' in MathsA CQC.

model item + item*step;

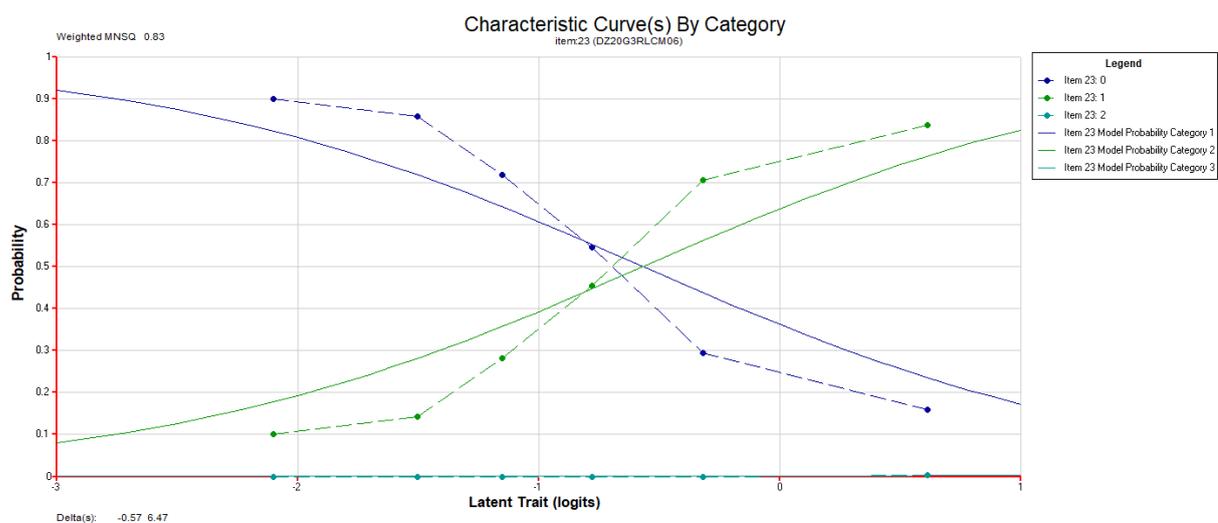
- Save all the changes made in the MathsA CQC file.
- Copy the CQC and lab file of MathsA and paste them and rename as MathsB in the same input folder and do the same for Maths B CQC and lab file as done for MathsA and save.
- Now save data, Syntax, CQC and lab files for both MathsA and MathsB are ready, open the conQuest for Run1.
- In the ConQuest, open MathsA CQC and run all. All the outputs of Maths A will get automatically saved in the mathsA Output Folder.
- In a similar manner, run for MathsB.
- To check the item behaviour and to ascertain which items have not functioned well, first check the summary file and itenals, then reliability, wright map, and finally GIN map.
- Open summary file of MathsA, Check the Item Rest Correlation and identify the items with less than 0.2 and for this select the Item Rest Correlation column and go to conditional formatting, highlight cell rules, less than, type 0.2, okay.
- Insert Remark column at the end of the summary excel sheet, and type the comment: Item-Rest Correlation less than 0.2. Review the Item Analysis report/item characteristics (itenal) for those flagged items in the itenals.

	A	B	C	D	E	F	G	H	I	J
4	Group All Students									
5	=====									
6	Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta	Remarks		
7	item:1 (MA	2225	44.26966292	0.398317944	0.472838296	0.980397221	-0.296819077			
8	item:2 (MA	2180	43.99082569	0.373233389	0.449043307	1.003155508	-0.281876835			
9	item:3 (MA	2282	33.08501315	0.37415902	0.447014255	0.998137164	0.256247052			
10	item:4 (MA	2251	19.76899156	0.112476875	0.184070619	1.198730264	1.097198591	Item rest correlation less than 0.2		
11	item:5 (MA	2296	83.49303136	0.271534978	0.334531365	0.984637828	-2.456508989			
12	item:6 (MA	2281	18.36913634	0.183130468	0.251509969	1.114105707	1.186525622	Item rest correlation less than 0.2		
13	item:7 (MA	2167	6.783571758	0.433909303	0.470243547	0.840488119	2.482381539			
14	item:8 (MA	2278	25.63652327	0.458505941	0.520566962	0.906077048	0.687655013			
15	item:9 (MA	2157	51.73852573	0.405444373	0.479411966	0.964272698	-0.668932302			
16	item:10 (M	2292	43.54275742	0.516750911	0.581823763	0.872477074	-0.277739636			
17	item:11 (M	2275	34.85714286	0.45732379	0.524775736	0.926084025	0.162141653			

- For. Ex. Item 4:

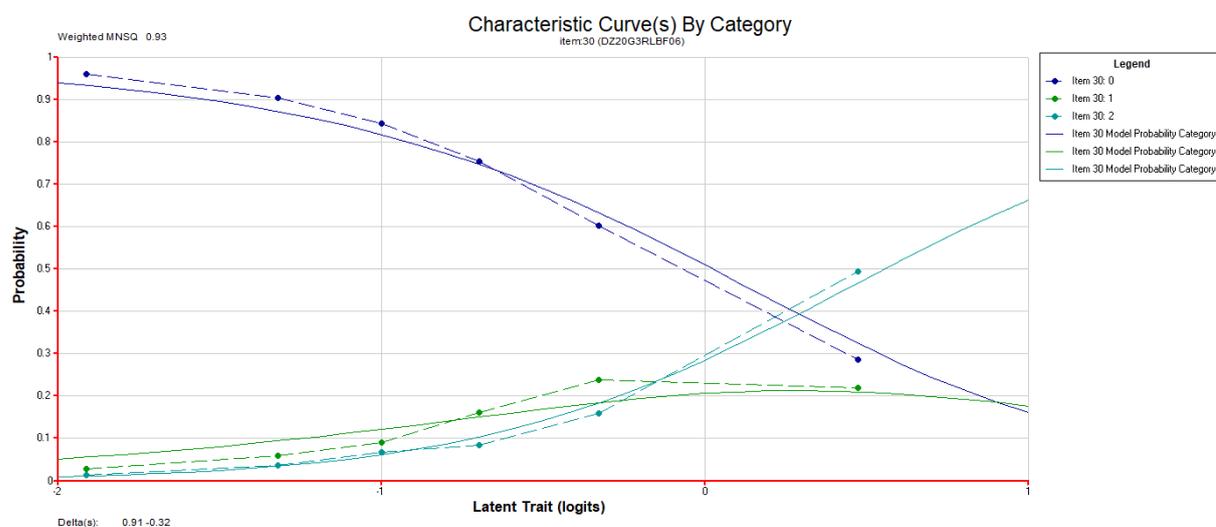
	A	B	C	D	E	F	G	H	I
37	Item 4								
38	item:4 (MA20G3NA73)								
39	Cases for this item: 2251 Item-Rest Cor. 0.11 Item-Total Cor. 0.18								
40	Item Threshold(s): 1.10 Weighted MNSQ 1.20								
41	Item Delta(s): 1.10								
42	Label	Score	Count	% of tot	Pt Bis	t	sig	PV1Avg:1	PV1 SD:1
43	A	0	1006	44.69	-0.04	-1.77	0.078	-0.634	0.906
44	B	0	548	24.34	0	0.09	0.925	-0.583	0.993
45	C	1	445	19.77	0.11	5.37	0	-0.252	1.042
46	D	0	252	11.2	-0.09	-4.1	0	-0.797	0.944

- Check the delta value, it has 1.10 which is quite a difficult item (normally an item should not have delta value more than +3 and less than -3 as in the wright map, if it is towards positive, the item is considered to be difficult). Though the answer is C, majority of the students have opted A and B. However, considering the ability of the students (PV1Avg:1), the PV1Avg is highest for the students who got the correct response (c) compared to others. Therefore, the item is functioning well and it could be retained for analysis.
- Item Total Correlation is the comparison of one item against the total score while Item Rest Correlation is the comparison of an item with the rest of the items (by removing the item from the total score).
- The reliability (EAP/PV reliability) of the entire test booklet should be 0.8-1.
- Item Rest Correlation should be between 0.2 to 0.6. If it is below 0.2, then the item is considered to be not reliable as the item does not discriminate between the high and low achievers and the same applies if it is higher than 0.6.
- Weighted MNSQ should be between 0.8 - 1.2. This shows how well the item fits the Rasch Model. Below 0.8 is overfit and 1.2 is underfit indicating that students of different abilities are not answering as expected.
- PV Average shows which ability of students have chosen the option.
- Pt.Bis is shows the discrimination between higher and lower ability test takers. It is more reliable than the PV AVG value and similar to Item Rest Correlation. It should be above 0.2 and the rest of the options should be negative ideally. If the Pt.Bis is negative, it negatively correlates/does not align with other items meaning higher ability students are not answering that item correctly. It could be either because of printing error, flaw in the question design etc. It gives a picture of whether the item is functioning well.
- Wright map is a map where the difficulty of items and the ability of students are put on the same scale. The objective is to see if the test is targeted well for all abilities of student. Ideally, the items should be spread across the difficulty level (bell curve) to be considered as a good fit. If the majority of the items are distributed at higher points of the map and few students are able to answer them correctly, the test is considered to be difficult and vice versa.
- Item threshold indicates the difficulty of items. In case of partial credit, the location where the response curves intersect is known as item threshold (in such case the probability of getting zero and correct is intersecting). As the threshold becomes positive, items are considered difficult. Item Threshold and Item Delta are same for MCQ but different for CRT.



- Above is an example of ICC curve for partial credit item. It is one of the best model fitting items. The threshold values (transition parameters) are the intersections of credit zero with partial credit1 and partial credit 1 with full credit 2. When the student is at $\theta = -1.14$, probability of getting zero and partial

credit is equal. Similarly, when the student is at $\theta = 0.76$, the probability of getting partial credit 1 and full credit is equal.



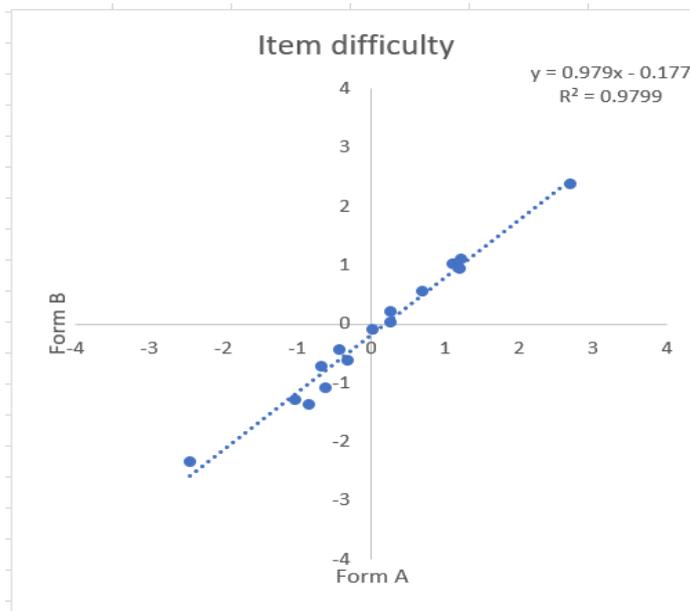
- The above ICC curve is an item with partial credit that did not behave as expected. The full credit 2 was scored by only one student. The threshold value ($\theta = 6.47$) for the full credit 2 shows that it is behaving beyond the acceptable range. In such cases, full credit item will be recoded (2 to 1) for further analysis.
- The above ICC curve is an item with partial credit where the partial credit 1 did not work as expected. The transition of getting full credit 2 is higher than partial credit 1 as the ability of the student increases as shown by the threshold value ($\theta = -0.32$).
- After examining the character of each item, decide whether to retain or drop the flagged items in consultation with the subject experts. Create a new column at the end in the MathsA summary file and write the decision (to keep/to drop/recode).
- Identify the item rest correlation in the summary file of MathB by using conditional formatting as done in MathsA and check all the items characteristics and identify which to retain and drop as done for MathA. And finally create a new column at the end in the MathsB summary file and write the decision (to keep/to drop/recode).

2.4.2 Run1B

- Run1B is carried out only if there is/are issue/s with items after Run1.
- If the item characteristics are not within the acceptable range, for example, poor or negative Item-Rest Correlation, partial credit codes not functioning and PV Average not as per the student's ability for MCQ items, the test items may have to be dropped from further analysis after Run1.
- For this, copy the entire folder of Run1 and paste it in the NEA folder and rename it as Run1B and also rename Run1 and Run1A. In case of English, since the item to be deleted was from BookletB, only the CQC file of booklet B needs to be updated and to obtain the outputs from Run1B, remove all the earlier outputs from the Output folder of Booklet B in the Run1B folder.
- Example, the item 22 was flagged with negative Item-Rest Correlation in Run1. This item has to be excluded from further analysis.

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta	Comments
item:22 (EN20G3RLLF02)	2286	24.71566	-0.014270002	0.056795139	1.355645209	1.206888154	To drop

- Filter out the linking items in both FormA and FormB sheets (remove all items except the linking items) by selecting only the linking item column and sort and filter. Check only 'Y'.
- Copy only the item IDs of MathsA and paste (value) in Equating Item Difficulty sheet and also copy the Avg Delta values of both FormA and B and paste (values) in Equating item Difficulty sheet. Now in the next comment column, insert the formula =B2-C2 (Avg Delta Value of FormA minus FormB). Study the general trend of the difference in delta value. Generally, the difference should be closer to zero to be along the regression line, if higher than zero highlight the items which deviate from the general trend and write in the comment column (large difference and delink item as it indicates that the item has behaved differently in FormA and FormB). An automatic item difficulty graph will be generated next to the comment column to see how the linking items have behaved in FormA and B. R squared Coefficient of determination-shows the strength of relationship) should be closer to 1 to show a perfect correlation of linking items in FormA and B.



- The linking items with large difference in delta average and that are not close to the regression line could be removed and made as unique items for future NEA cycles to have higher coefficient of determination (closer to 1). In Maths, item:16 (MA20G3MA56) and item:25 (MA20G3NA38) could be removed from the linking items and considered as unique items.

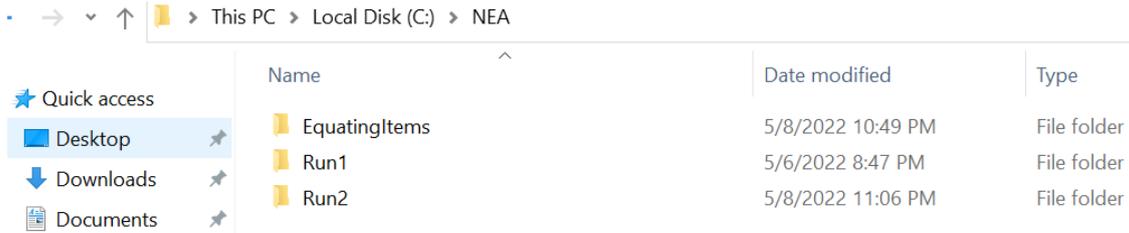
ItemID_Maths	Form A (X)	Form B (Y)	Difference(X-Y)	Comment
item:3 (MA20G3NA06)	0.256247	0.217158	0.04	
item:5 (MA20G3GE45)	-2.456509	-2.337198	-0.12	
item:6 (MA20G3NA37)	1.186526	0.953537	0.23	
item:8 (MA20G3MA13)	0.687655	0.565113	0.12	
item:9 (MA20G3NA29)	-0.668932	-0.715174	0.05	
item:12 (MA20G3MA11)	1.108572	1.036673	0.07	
item:16 (MA20G3MA56)	-0.846101	-1.368059	0.52	Large difference
item:17 (MA20G3GE66)	-0.32349	-0.608733	0.29	
item:18 (MA20G3NA03)	0.020642	-0.092354	0.11	
item:19 (MA20G3NA65)	2.685533	2.381102	0.30	
item:20 (MA20G3NA81)	1.219048	1.113477	0.11	
item:21 (MA20G3GE44)	0.25652	0.035263	0.22	
item:24 (MA20G3MA57)	-0.433249	-0.439024	0.01	
item:25 (MA20G3NA38)	-0.620155	-1.081778	0.46	Large difference
item:29 (MA20G3NA10)	-1.025629	-1.290398	0.26	

- Copy the item rest correlation of Form and Form B and paste (value) in Equating Item Discrimination sheet and find their difference (X-Y). The difference should be close to zero, as it indicated that the linking items have discriminated equally in both the Forms A and B. Check the item properties and

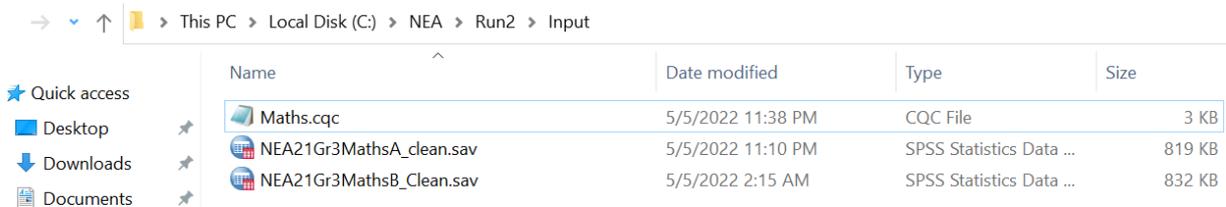
the item discrimination graph. Check if the R squared is closer to 1 as done before for the Equating Item Discrimination.

2.6 Run 2 (Concurrent Analysis)

- Prepare Syntax file for Run 2 as two items were delinked in Maths.
- Create folder for Run2 in NEA folder. Inside the Run2 folder, create two new folders and name them InputMaths and OutputMaths.



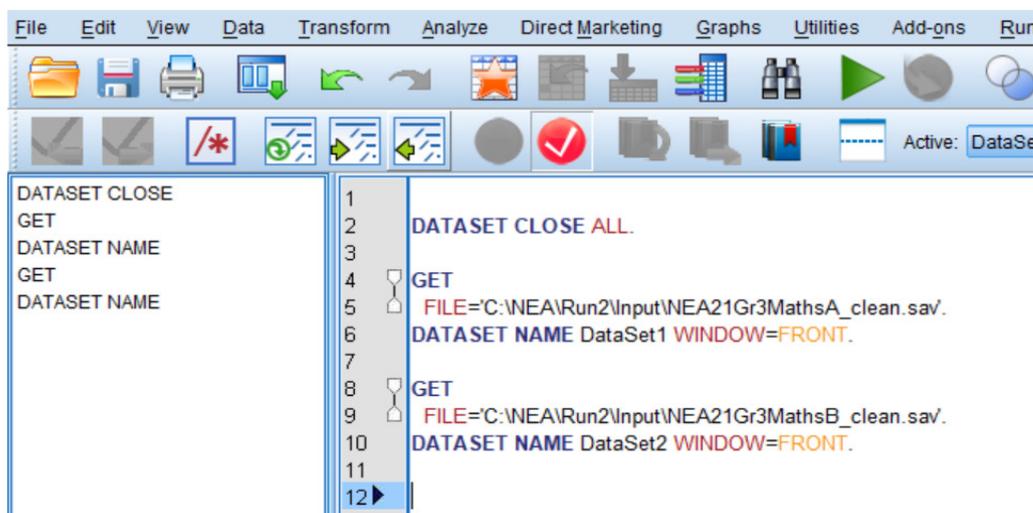
- Copy the MathsA CQC file, lab file and MathsA and B clean data save files from Run1 Input folders (In case of English, copy the EngB CQC file and lab file as one of the items was dropped in Run1B and also copy EngA and B clean data sav files from Run1B) and paste them in Run2 Input folder and re-name the MathsA CQC file as Maths lab file.



- Open SPSS, click open, data and then open the MathA clean data sav file from Run2 Input folder, then paste. Now open the syntax and write on the top DATASET CLOSE ALL.
- Copy
- GET

`FILE='C:\NEA\Run2\Input\NEA21Gr3MathsA_clean.sav'.`

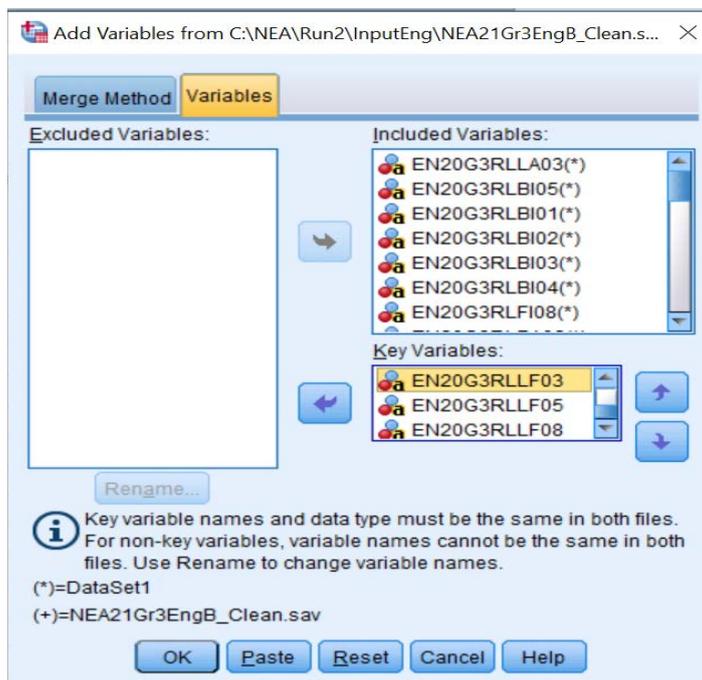
`DATASET NAME DataSet1 WINDOW=FRONT.` and paste below it in the syntax and in the pasted syntax change MathsA to MathsB and Set1 to Set2 as the data of formA and formB are getting merged. Select all and run. Two data sets of both MathsA and B will appear in separate SPSS data safe file.



- In the MathsB dataset clean, In the variable view, rename the delinked items (by adding underscore b at the end of the name) and save and close.

23	MA20G3NA20	String
24	MA20G3DP23	String
25	MA20G3NA41	String
26	MA20G3MA56_b	String
27	MA20G3GE66	String
28	MA20G3NA03	String
29	MA20G3NA65	String
30	MA20G3NA81	String
31	MA20G3GE44	String
32	MA20G3NA75	String
33	MA20G3GE87	String
34	MA20G3MA57	String
35	MA20G3NA38_b	String
36	MA20G3MA17	String

- In the SPSS dataset of MathsA, in the variable view, click data, merge files, add variables, select MathsB, continue. Click paste when the box below appears.



- Go to syntax file and select all the dataset activate and run, a new SPSS data save file with MathA and B merged will appear. In the variable view, go to file, save as and name it **NEA21Gr3Maths_Merged.sav** and paste.
- Go back to the syntax and select the

`SAVE OUTFILE='C:\NEA\Run2\Input\NEA21Gr3Maths_Merged.sav'
/COMPRESSED.` and run.

- In the MathsMerged data variable view, check where the items are merged where 'Primary last' would appear. Highlight and drag it to the end of the data so that all the merged item variables would be together in sequence and then save.

48	MA20G3DP23	String	1	0		None	None	1	Left	Nominal	Input
49	MA20G3NA41	String	1	0		None	None	1	Left	Nominal	Input
50	MA20G3MA...	String	1	0		None	None	1	Left	Nominal	Input
51	MA20G3NA75	String	1	0		None	None	1	Left	Nominal	Input
52	MA20G3GE...	String	1	0		None	None	1	Left	Nominal	Input
53	MA20G3NA...	String	1	0		None	None	1	Left	Nominal	Input
54	MA20G3MA...	String	1	0		None	None	1	Left	Nominal	Input
55	MA20G3NA80	String	1	0		None	None	1	Left	Nominal	Input
56	MA20G3NA82	String	1	0		None	None	1	Left	Nominal	Input
57	MA20G3NA39	String	1	0		None	None	1	Left	Nominal	Input
58	PrimaryLast1	Numeric	1	0	Indicator of eac...	{0, Duplicat...	None	14	Right	Ordinal	Input

2.6.1 Update CQC file

- Select all the item variables from the Maths Merged variable view and paste them in the responses in the CQC file and maintain same space between each item variable. Do the necessary updates inside the CQC file such as naming in the top (Grade 3 Maths Run2 Concurrent Analysis analysis). Update input and output location, let name.

```
*Maths.cqc - Notepad
File Edit Format View Help
/*-----*/
/* Bhutan NEA Main */
/* Grade 3 Maths Run2 Concurrent Analysis*/
/*-----*/
reset all;

let input=C:\NEA\Run2\Input;

let output=C:\NEA\Run2\OutputMaths;

let name=G3MathsRun2;
```

- Open the Maths lab file and delete the item variables, copy the new item variables from the variable view of the Merged Maths data safe file and paste them in it and number them serially maintaining equal space between the serial numbers and the item variables and save.

```
*Maths.lab - Notepad
File Edit Format View Help

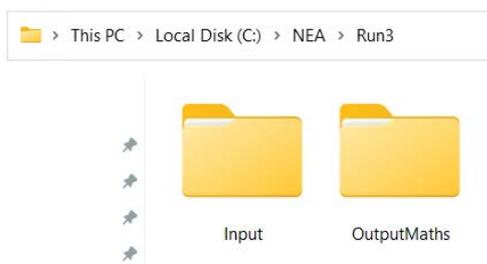
===> item
1 MA20G3GE44
2 MA20G3GE45
3 MA20G3GE66
4 MA20G3MA11
5 MA20G3MA13
6 MA20G3MA57
7 MA20G3NA03
8 MA20G3NA06
9 MA20G3NA10
10 MA20G3NA29
11 MA20G3NA37
12 MA20G3NA65
13 MA20G3NA81
14 MA20G3NA01
15 MA20G3NA42
16 MA20G3NA75
```


- Figure Par file

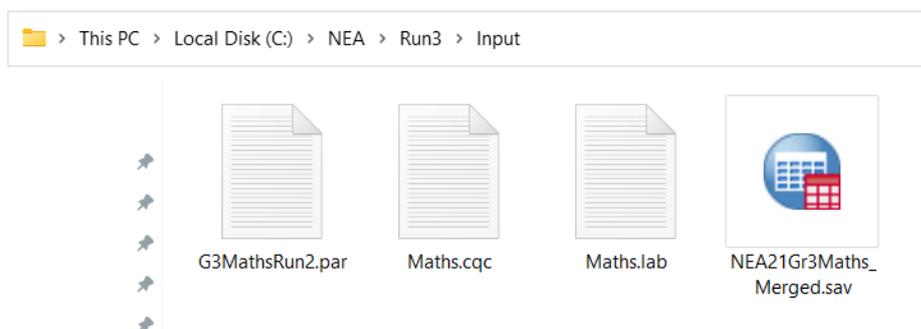
```
G3MathsRun2.par - Notepad
File Edit Format View Help
1 -0.31764 /* item MA20G3NA01 */
2 -0.30330 /* item MA20G3NA42 */
3 0.27226 /* item MA20G3NA06 */
4 1.07648 /* item MA20G3NA73 */
5 -2.35787 /* item MA20G3GE45 */
6 1.10332 /* item MA20G3NA37 */
7 2.46066 /* item MA20G3NA50 */
8 0.66131 /* item MA20G3MA13 */
9 -0.65764 /* item MA20G3NA29 */
10 -0.29868 /* item MA20G3DP19 */
...
40 -1.27616 /* item MA20G3MA56_b */
41 0.54467 /* item MA20G3NA75 */
42 -1.91472 /* item MA20G3GE87 */
43 -0.98960 /* item MA20G3NA38_b */
44 0.47025 /* item MA20G3MA17 */
45 1.52979 /* item MA20G3NA80 */
46 0.41280 /* item MA20G3NA82 */
```

2.7 Run 3 (Case Estimation Analysis)

- In the NEA folder, create Run3 folder and inside the Run3 folder, create 2 folders and name them Input and OutputMaths.



- Inside the Input folder of Run3, insert the CQC, lab file and merged dataset files from the Run 2 Input folder and also insert the Par file from the Run2 Output folder.



- In the outputs of CQC file, replace all the latent to wle (select one of the 'latent' words, then ctrl+H, and replace one by one with wle)

```
/* =====OUTPUTS=====*/
show !
    estimates= wle,
    table=1:2:3:4:5:7,
    filetype=excel,
    plotmin=-4, plotmax=4          >>%output%\show %name%.xls;

show !
    estimates= wle          >> %output%\%name%.shw;

itanal
itanal!          >> %output%\%name%.itn;
    estimates= wle,
    format=long,
    filetype=excel          >> %output%\%name%_itanal.xls;

itanal!
    estimates= wle,
    format=summary,
    filetype=excel          >> %output%\%name%_summary.xls;
show cases ! estimates= wle          >> %output%\%name%.wle;

show cases ! estimates= wle,
    filetype=spss          >> %output%\%name%_wle.sav;

export parameters >>          %output%\%name%.par;

export covariance >>          %output%\%name%.cov;
```

- Remove all the contents below the plots and save because in the Run3 only bands are required for developing the proficiency scale so, item analysis (ICC curves and graphs) are not required.

```
show cases ! estimates= wle,
    filetype=spss          >> %output%\%name%_wle.sav;

export parameters >>          %output%\%name%.par;

export covariance >>          %output%\%name%.cov;

/* =====PLOTS===== */

/*-----*/
```

- Update parameter file (For example, G3MathsRun2.par) generated in Run2. By default, last item parameter is not included in the parameter file. Maths has a total of 47 items and in parameter file only 46 items are reflected as shown below:

```
40 -1.27616 /* item MA20G3MA56_b */
41 0.54467 /* item MA20G3NA75 */
42 -1.91472 /* item MA20G3GE87 */
43 -0.98960 /* item MA20G3NA38_b */
44 0.47025 /* item MA20G3MA17 */
45 1.52979 /* item MA20G3NA80 */
46 0.41280 /* item MA20G3NA82 */
```

- Parameter for the 47th item can be obtained from Run2 summary file.

```

=====
ConQuest: Generalised Item Response Modelling Software  Mon May 09 14:56 2022
GENERALISED ITEM ANALYSIS
Group All Students
=====

```

Item	N	Facility	Item-Rest Cor	Item-Total Cor	Wghtd MNSQ	Avg Delta
item:43 (MA20G3NA38_b)	2315	58.31533477	0.408562177	0.482488777	0.943733151	-0.989599925
item:44 (MA20G3MA17)	2289	29.53254696	0.336077283	0.408867543	1.031938449	0.470253599
item:45 (MA20G3NA80)	2332	14.53687822	0.485695323	0.533944358	0.856837532	1.529792596
item:46 (MA20G3NA82)	2320	30.30172414	0.411565794	0.480436447	0.963728825	0.41280279
item:47 (MA20G3NA39)	2310	59.35064935	0.471979182	0.540523328	0.890440541	-1.039023706

- Updated parameter file is as follows:

```

43  -0.98960  /* item MA20G3NA38_b */
44   0.47025  /* item MA20G3MA17 */
45   1.52979  /* item MA20G3NA80 */
46   0.41280  /* item MA20G3NA82 */
47  -1.039023706 /* item MA20G3NA39 */

```

- Now all CQC file is ready for Run3. Open the ConQuest and open the CQC file from Run3 Input folder.

2.8 Determining bands/cuts for the proficiency scale

- Open the show file of the output and then open the summary sheet. Check if the deviance convergence criteria was reached, if so, it is successful which indicates how well the item response model has fit the data.

	A	B	C	D	E	F	G	H
16	Grouping Variables:							
17	The item model: item							
18	Slopes are fixed							
19	Cases in file: 4658 Cases in estimation: 4657							
20	Final Deviance: 145598.64835							
21	Akaike Information Criterion (AIC): 145604.64835							
22	Akaike Information Criterion Corrected (AICc): 145604.64577							
23	Bayesian Information Criterion (BIC): 145623.98673							
24	Total number of estimated parameters: 3							
25	The number of iterations: 7							
26	Termination criteria: Max iterations=2000, Parameter Change= 0.00001							
27	Deviance Change= 0.00010							
28	Iterations terminated because the deviance convergence criteria was reached							
29	Random number generation seed: 5.00000							

- Open the wright map sheet and start the selection of items for which it must begin from the lowest item (easiest item) to the most difficult item and for that open the summary file from the Run3 Output folder, and then select all the data including the headings of the column and custom sort by Avg Delta from largest to smallest order.

	A	B	C	D	E	F	G
1	=====						
2	ConQuest: Generalised Item Response Modelling Software Mon May 09 23:41 2022						
3	GENERALISED ITEM ANALYSIS						
4	Group All Students						
5	=====						
6	Item	N	Facility	em-Rest	Cem-Total	C/ghtd MNS	Avg Delta
7	item:5 (MA20G3GE45)	4623	82.04629	0.280582	0.345077	0.986291	-2.3579
8	item:42 (MA20G3GE87)	2263	75.82855	0.200463	0.274697	1.067006	-1.91475
9	item:27 (MA20G3GE83)	2282	70.59597	0.386019	0.455958	0.939171	-1.63679
10	item:35 (MA20G3DP62)	2320	70.60345	0.337867	0.410929	0.972996	-1.6197
11	item:26 (MA20G3DP61)	2269	65.27104	0.425065	0.495035	0.929836	-1.34777
12	item:14 (MA20G3GE84)	2289	64.04543	0.4085	0.480639	0.935898	-1.28957
13	item:40 (MA20G3MA56_b)	2319	64.07934	0.303709	0.382715	1.018909	-1.27619
14	item:29 (MA20G3NA10)	3722	61.31112	0.380503	0.45347	0.960974	-1.12065
15	item:47 (MA20G3NA39)	2310	59.35065	0.471957	0.540498	0.890707	-1.03873
16	item:22 (MA20G3NA51)	2271	57.94804	0.389613	0.464898	0.96497	-0.99
17	item:43 (MA20G3NA38_b)	2315	58.31533	0.40854	0.482464	0.943697	-0.98963
18	item:16 (MA20G3MA56)	2280	55.4386	0.322338	0.404644	1.019844	-0.86656
19	item:30 (MA20G3NA79)	2259	53.91766	0.357359	0.435491	0.998166	-0.79084
20	item:28 (MA20G3NA34)	2259	51.52722	0.34499	0.423916	1.017299	-0.6761
21	item:9 (MA20G3NA29)	4233	51.35837	0.405172	0.478584	0.964348	-0.65766

- In the international assessments, generally 0.8 logit difference is maintained between the categories of student abilities in the proficiency scale. Good pool/number of items must be included in each category/band/cuts.
- Copy the Delta value of the easiest item (which will be the base point) from the sorted summary file and paste it in the wright map next to the item number. Then add 0.8 to the easiest item's ave delta and get the sum, that becomes the highest ave delta value for the particular band. For ex. The ave delta value of the easiest item in maths is -2.36 and the highest is -1.56 (which is the sum of delta value of easiest item and 0.8). The cut for this first band must begin from -2.36 to the delta values closer to -1.56, which comes to around -1.62.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
7																
3																
9																
0																
1	1															
2																
3																
4																
5																
6																
7																
3	0															
3																
9																
0																
1																
2																
3																
4																
5																
6																
7																
3	-1															
4																
5																
6																
7																
3																
9																
0																
1																
2																
3																
4																
5																
6																
7																
3	-2															
4																
5																
6																
7																

- Similarly, for the next band, to the highest delta value of the previous band, add 0.8 which becomes the highest delta avg (-1.62+0.8= -0.82) for this second band. So for the second band, the lowest avg delta value is -1.29 and the highest delta avg closest to -0.82 is -0.79. In this similar manner determine the remaining bands.
- In case of less number of items in each band, seek expert opinions and suggestions to make necessary adjustments on the expansion of some of the bands to have adequate number items to develop descriptors for the proficiency scale.
- Depending upon the spread of the items and the items in each band, the total number of bands could be decided. The number of bands could differ from domain to domain. For. ex, Maths has 5 bands while Eng and Dzo have 4 bands.

2.9 Developing descriptors of the bands

- First identify the items and the item descriptors/characteristics of the items from the code book of each band.
- Populate all the item descriptors for each band.
- Refer the learning outcomes and then build on the item descriptors.
- Refer the cognitive processes and the complexity of each item in the bands and finalize the band descriptors.
- Ensure there is vertical alignment of the learning outcomes along the band; the descriptors have to be progressive as the band level goes up.

	A	B	J	K	L	M
1	Item	Variable	Item Descriptor	Band descriptor_Draft 1 (Based on learning Outcome)	Band descriptor_Draft 1 (Based on cognitive processes)	Band descriptor_Draft 1 (Based on complexity)
13	item:15 (MA20G3NA33)	MA20G3NA33	Converts fraction into decimal	performs basic multiplication and division	performs basic multiplication and division	
14	item:8 (MA20G3MA13)	MA20G3MA13	Converts minutes to hours	Identifies fraction as part of a set in context and demonstrates understanding of relationship between fractions and decimals	recognises the whole can be redivided into different fractional parts for different purposes; connects and converts between decimals and fractions	
15	item:31 (MA20G3NA70)	MA20G3NA70	Multiplies a 2-digit number by a 1-digit number in real life situation	Identifies right angle in the context and conversion of minutes to hour	identifies angles as greater than, less than or equal to a right angle; conversion of minutes to hour	
16	item:32 (MA20G3MA16)	MA20G3MA16	Identifies right angle in the context	Uses data in bar graph to solve a problem	uses data in bar graph to solve a problem	
17	item:41 (MA20G3NA75)	MA20G3NA75	Identifies even number in			

Proficiency Scale Descriptor for Grade 3 Mathematics	
5	Applies the concept of place values to solve simple problems in familiar situations; relates repeated addition and multiplication; multiplies 2-digit numbers by 1-digit numbers and divides 2-digit numbers by 1-digit numbers; recognises growing patterns with shapes; identifies a single operation rule in numerical patterns and finds the missing term; interprets different representations of time on an analogue clock to solve simple problems in familiar situations; aligns the corresponding faces of an object and its net; uses data in bar graphs to solve simple problems in familiar situations
4	Recognises odd and even numbers in familiar situations; relates multiplication and division; selects and applies multiple strategies for solving problems involving addition and subtraction up to 3-digit numbers; performs basic multiplication and division (2-digit by 1-digit) to solve simple problems in familiar situations; connects and converts between decimals (up to tenths) and fractions (unit fractions); converts minutes to hour; identifies angles as greater than, less than or equal to a right angle; uses data in pictographs to solve simple problems in familiar situations
3	Adds and subtracts up to 3-digit numbers to solve simple problems in familiar situations; recognises that different wholes can be divided to show the same fractional parts; recognises repeating patterns with shapes; measures, compares and estimates length and mass using formal units; converts between formal units of measurement; calculates elapsed time; classifies simple geometrical shapes based on their attributes; identifies basic transformation (turns and flips); uses data from a tally chart to solve simple problems in familiar situations; uses a calendar to solve simple problems in familiar situations
2	Compares up to 5-digit numbers; subtracts up to 2-digit numbers by regrouping; recognises unit fraction and decimals up to tenth digit; represents familiar situations using number sentence; recognises patterns involving skip counting; reads time to the hour, half-hour and quarter-hour on analogue clocks; recognises parallel, perpendicular lines and line of symmetry; reads data from a tally chart to solve simple problems in familiar situations
1	Recognises up to 4-digit numbers; performs addition and subtraction on up to 2-digit numbers without regrouping; uses multiple non-standard units to measure length, mass or capacity; identifies and classifies 2-D and 3-D shapes; retrieves information from a tally chart

A student at the top of a band is likely to have demonstrated all of the skills in that band, and all of the skills in the bands below. A student in the middle of a band is likely to have demonstrated half of the skills in that band, and all of the skills in the bands below.

2.10 Establishing baseline

Study the WrightMap and check the percentage of population in each band. The band in which maximum population is located is generally considered as the baseline. However, the ultimate decision must be made by the domain experts, curriculum experts and teachers by referring to the learning outcomes and the items in each band so that the minimum learning standard required at grade III is not compromised. The band that describes the minimum level for a grade should be dependent on the content of the items and the syllabus. The Learning Outcomes need to be carefully examined and once the minimum acceptable ones have been selected, the band description that most closely matches the LOs should be selected.

**“An educated and enlightened society
guided by GNH principles”**

Vision - Ministry of Education, Bhutan



**“An internationally recognized centre of excellence in
educational assessment that provides quality services to
build the integrity and profile of the education system leading
to an improvement in the quality of learning in Bhutan”**

**Vision - Bhutan Council for School
Examinations and Assessment, Bhutan**