User Manual for IAPlus

# About IAPlus

IAPlus is an SPSS macro that has the functionality of processing test or questionnaire response data and compute a variety of classical test theory statistics. It should be used primarily to evaluate the quality of the response data. As such, SPSS is required to execute the instructions generated by IAPlus, and IAPlus is called from, and run from within an SPSS syntax window.

In addition to using SPSS, IAPlus has the option of calculating IRT related statistics. It calculates IRT statistics using the TAM package in R. When using TAM, IAPlus provides the option to use marginal maximum likelihood estimation for the Rasch, 2 parameter logistic model, or the generalized partial credit model.

Summary item statistics are calculated and presented overall, or by groups specified by the user.

# Requirements

IAPlus requires a working copy of SPSS installed in the computer where the analysis is conducted. If the optional IRT statistics are requested, the R programming language also needs to be installed in the computer. Within R, it requires the TAM package. IAPlus installs the TAM package when needed.

In addition to the macro, a working version of SPSS, and R (optional), IAPlus requires an SPSS system file with response data to a survey or a test. The responses must be stored in numeric format. Responses stored as SPSS system missing values (SYSMIS) will be treated as not administered and will be excluded from the analysis. Using numeric values, the responses can be coded as valid responses, or as one of several missing type response possibilities. As such, the macro allows for differential treatment of omitted, not reached, not administered or other missing type responses.

For the calculation of classical test theory statistics, values for the responses do not need to be sequential and can be one or more digits. But for the calculation of IRT based statistics, scored responses need to be sequential, and starting at 1 or 0. When responses start at 1, the program provides the option to down-code the response data. More details are presented later in this manual in the parameter section.

In addition to the response data, IAPlus can accept one or more ID variables for each response record, and one or more grouping variables used for the analysis. There are two types of grouping variables, classification variables (CLASSVARS) and group comparison variables (BYVARS). These are explained later in the syntax reference section, but when these are specified, the analysis is carried out for each of the groups defined by the grouping variables. Unlike the response data, which must be numeric, the grouping and ID variables can be string, numeric or a combination. There is no restriction of the number of digits, and the categories for the grouping variables do not need to be sequential.

# Using IAPlus

To use IAPlus to process response data, you will need to create an SPSS syntax file with an INCLUDE command, followed by a call to the macro. The syntax will look something like this:

include file = "C:\Temp\ItemAnalysis\IAPlus.ieasps".

!IAPlus indir = "C:\Temp\ItemAnalysis\Examples" /

infile = SampleDataPIRLS /

outdir = "C:\Temp\ItemAnalysis\Examples" /

outfile = PIRLSExample /

scale = Reading /

scalelbl = Reading Literacy /

items = R011C01C R011C02C R011C03C R011C04M R011C05M R011C06C

R011C07M R011C08C R011C09M R011C10C R011C11C R011C12M

R011C13M R011F01M R011F02M R011F03M R011F04M R011F05M

R011F06C R011F07C R011F08C R011F09C R011F10C R011F11M

R011F12C R011F13M /

keys = X X X 1 3 X 2 X 4 X X 2 3 2 3 4 1 4 X X X X X 1 X 2 /

maxscrs = 1 1 1 1 1 1 1 2 1 3 2 1 1 1 1 1 1 1 1 2 1 2 1 1 2 1 /

reverse = /

Omit = 9 99 /

OmitAs = 0 /

NotReach = 6 96 /

NotReachAs = sysmis /

NotAdmin = 8 88 /

classvars = /

atleast = 3 /

byvars = idcntry /

factor = n /

dodif = y /

idvars = IDStud /

critvar = asrrea01 /

wgtvar = totwgt /

NullByVar = n /

DoTables = Y /

qcscoring = n /

viewcod = n /

clean = Y /

DoIRT = y /

WhereIsR = "C:\Program Files\R\R-4.0.2\bin" .

Notice that the first line INCLUDEs the macro, and the following lines are the call to the macro. Some of the parameters are optional and can be left blank. You only need to include the macro once per session.

When calling the macro, you need to use the exclamation point (!) before IAPlus, as in !IAPlus. Following !IAPlus you will have a series of parameters with the corresponding value(s). Each parameter is separated by a slash, much in the same way you would separate subcommands when writing SPSS syntax. The last parameter ends with a period (“.”). Some commands are mandatory, like the name and location of the input and output file, but many of them are optional and/or have default values. The section on parameters in this User Manual will specify the parameters, their use, and accepted and default values.

The parameters and corresponding values are not case sensitive. The order of the parameters is also not important. Neither is the alignment, other than they should not start on the first column. You can also have multiple parameters in a single line, but always separated by a slash. If you repeat a parameter in the call, the last instance of the parameter will be the one used by the program.

The alignment and sequence of the parameters in the syntax above is done as such for looks and ease of reading.

# Output Files

Depending on the setup and configuration, IAPlus creates several output files. There is a set of files that are always generated, and there is a set of optional files. Files with summary item statistics are always generated, and these are in SPSS system file format and CSV format. Files with output (tables and graphic) are created in SPV format (SPSS native output format) and in HTML format that can be opened with any HTML file viewer.

All files generated by IAPlus are written to the directory specified with the parameter OUTDIR, and use the text specified with the parameter OUTFILE as the root for the file name. What follows is a list of the files generated by IAPlus.

Please note that if a file with the same name already exists, the existing file will be automatically replaced with no warning to the user.

## [OUTFILE]\_Output.SPV

This file contains the entire SPSS output file in native SPSS SPV format. Depending on the options requested, the contents of this file will vary. But in general, this file will mostly have output that should be used for quality control of the data and the corresponding analysis. There is a single SPV file for each call to the macro submitted.

This file will have the following sections:

1. Two tables for each item with the crosstabulation of the original response value for the item and the scored value for the item, and a second one with the crosstabulation of the response value for the item and the proportion of the maximum score on the item. These tables help confirm proper assignment of the responses to the scored and p-plus version for each item. It also helps confirm the item reversal has been applied properly, when requested. This output can be suppressed with the option QCSCORING = N. The default is QCSCORING = Y.
2. Descriptive statistics of the scored responses. You should check here the minimum and maximum for each item and make sure it agrees with your expectations of the response data. This output cannot be suppressed.
3. Descriptive statistics for the items transformed to the p-plus metric. You should check here the minimum and the maximum values for each item. The maximum should not exceed 1.0. If it does, you need to check the parameter MAXSCRS for the items. The minimum score will generally be 0.0, or 1/MAXSCRS for the item. This will depend on the lowest value allowed for the item. This output cannot be suppressed.
4. A crosstabulation of the variables [SCALE]\_use with [SCALE]\_taken. Cases set to [SCALE]\_use = 0 will be excluded from the subsequent analysis and reports. You want you check to confirm the desired cases are excluded from the analysis. The minimum number of valid responses is set with the parameter ATLEAST. The default value is 1. This output cannot be suppressed.
5. Frequency distributions and histograms for the sum score and proportion correct score for the set of items. Use this to check that the sum scores and proportion correct scores match your expectations.
6. When requesting to conduct an IRT analysis (DOIRT = Y),
   1. A listing of each item included in the IRT analysis with the corresponding item parameters. You want to check that all intended items are included in this list. Note that items that have no valid responses, or that do not vary, are automatically excluded from the IRT analysis. You should check the corresponding \*.ROUT where the items eliminated from the IRT analysis are listed. If you specified IRTModel = 1PL, you should confirm that all values for alpha are = 1. If you specified IRTModel = 2PL, you should confirm that the values are all positive and not constant. If you requested to center the difficulties, it will display the shifting constant applied to the difficulty of the items, and the transformed item difficulty. When working with polytomous items it will also display the step parameter. This output cannot be suppressed.
   2. Descriptive statistics for the item parameter, before, and after centering, when difficulty centering is requested. When centering is requested, you will want to check the centering was accomplished by seeing a mean beta = 0.0. This output cannot be suppressed. The value of “beta” is the one used for scoring. The value of “oBeta” is the original item difficulty.
   3. Descriptive statistic for the WLE and EAP scores in the original metric, or in the transformed metric when centering of the difficulties is requested.
   4. A cross tabulation between indicators of the case being used in the IRT analysis (FromTAM) and found in the original data file (FromSource). You want to make sure no additional cases were introduced or lost during the processing. The cases FromTAM should match those displayed in the earlier table.
   5. A table with correlations between the WLE and EAP scores from TAM and the sum score and proportion correct scores. You want these correlations to be high and positive.
7. When the parameter CHKCRITERIA = Y, the default value, this file will include descriptive statistics and histogram for the criterion variables specified with CRITVAR.
8. When requesting to do a factor analysis (FACTOR = Y), it will include the following output:
   1. Table with the item commonalities.
   2. Table with the total variance explained by each of the items
   3. Table with the component matrix
   4. Table with the rotated matrix, if there is more than one factor extracted.
9. When requesting a DIF analysis using DODIF = Y (see description below), it will include the following output:
   1. Descriptive statistics for each of the criterion variables showing minimum and maximum values, and the number of cases within each of the level groupings created using the specified parameters. Note that the range of the criterion scores within each of the groupings do not overlap with the others.
   2. A crosstabulation for each of the criterion variables, by level grouping assignment, and within each combination of the classification and group comparison variables. The cells show the number of cases within each level grouping. You want to check here to make sure there are plenty of cases to report across the different groups and use fewer groups if necessary. The number of groups can be controlled with the NLEVELS parameter (default = 6), and for the IRT scores, with the parameter IRTCUTS. More details on how to use these parameters is presented later.
10. This file will contain the tables, bar charts and line graphs described in the next section where the HTML and XLSX files are described.
11. When the parameter VIEWCOD is set to Y, the SPV file will also include the entire expanded syntax executed by the program. By default, VIEWCOD = N.

## [OUTFILE]\_[BYVAR]\_CTables.html and .XLSX

When setting DOTABLES = Y, there will be one of these files for each group defined by the combination of classification (CLASSVARS) and group comparison variables (BYVARS) specified for the analysis. If no group comparison is specified, there will be a single file created. When setting the parameter NULLBYVAR = Y, it will create one of these files using each of the BYVARS, and one using none of them.

When there are no CLASSVARS and no BYVARS, this file will contain one table displaying for each item the following entries:

1. Item name and characteristics (sequence, name, key, reversal code)
2. Number of cases with valid responses to the items.
3. P-plus for the item.
4. Correlation with the scores computed by IAPlus, and each of the CRITVARS provided.
5. Proportion choosing each response option.
6. Biserial correlation for each of the response options.

When there are any CLASSVARS or BYVARS used for the analysis, there will be a separate table for each item, displaying the same entries listed above, and the group classification based on the corresponding CLASSVARS by BYVARS combination.

The tables will be followed by a set of bar charts displaying the proportion of respondents choosing each response option, and the biserial correlation between the response option and each of the scores and criterion variables processed by IAPlus. The will be one set of tables for each unique combination of CVARS. When there are BYVARS used, the bars will be grouped by the corresponding BYVARS value.

When DODIF = Y, and there are BYVARS provided, this output will also contain a set of line charts displaying, for each of the groups defined by the BYVARS, the p-plus on the item for each level on the scores and criterion variables processed by IAPlus. The number of levels for the scores can be set using the parameter NLEVELS (default = 6), and the parameter IRTCUTS.

## [OUTFILE]\_Scored.CSV and SAV

This file contains one entry for each response record read from the response file. It is only created as an SPSS system file. It contains the following variables:

CaseSeq Original sequence number for the record in the [INFILE]. Note that this sequence number corresponds to the sequence number prior to any selection done with the parameter SELCRIT. When the parameter SELCRIT is used, the values for CaseSeq might not be sequential.

[IDVARS] When specified, the ID variables for the response records.

[CLASSVARS] When specified, the classification variables for the response records.

[BYVARS] When specified, the group comparison variables for the response records.

[CRITVARS] When specified, any external criterion variable specified for the analysis.

[WgtVar] Weight variable used for the analysis. If none is specified, this takes on the value of 1 for all records.

[ITEMS] Original values used for the responses to the items. Not administered responses and system missing values are recoded as -999.

[ITEMS]\_s Scored responses to the items. These usually take values of 0 or 1 to the maximum score. Not administered items are coded as SYSMIS.

[ITEMS]\_p Proportion of the maximum score obtained on the item. This is calculated as the score on the item divided by the maximum score possible on the item. The maximum possible value is 1. Not administered items are coded as SYSMIS.

[SCALE]\_taken Number of valid responses in the response string. This is calculated as the number of items in the scale minus those not administered.

[SCALE]\_use Whether the record was used for the analysis (1) or not (0). When the parameter ATLEAST is used, records that do not have an ATLEAST number of valid responses are excluded from the analysis.

[SCALE]\_sumscr Number of points obtained on the item administered. This is calculated as the sum of [ITEMS]\_s.

[SCALE]\_PPlus Proportion of the maximum score possible on the items administered. This is calculated as the average [ITEMS]\_p across the items taken.

When running IRT (DOIRT = Y), you get these additional variables:

TAMcase Sequential case number as processed in TAM. This might differ from CaseSeq and cases might be selected using SELCRIT or excluded from the analysis by using only those that have a minimum number of responses.

TAMweight Survey weight as used in TAM. TAM normalizes the survey weights to add up to the sample size. This weight might differ in magnitude from the sampling weight specified with the parameter WGTVAR by a constant factor.

TAMsumscore Number of score points on the set of items with valid responses for the record. Note that the sumscore assumes each item has a minimum score of 0. While this will generally agree with the variable [SCALE]\_sumscr (see below), it will differ when the minimum score for the items is 1. This is often the case for attitudinal or other contextual scale items where responses are scores starting with 1. When running TAM, these responses need to be down coded (DOWNCODE = Y) so that the minimum score for each item is 0.

TAMmaxscore Maximum possible sum score on the set of items that were not coded as NA for the record.

[SCSALE]\_oEAP When selecting to do IRT (DoIRT = Y) and centering the difficulties at 0 (CtrDiff = Y), this are the EAP scores in the original metric, prior to centering the difficulties and applying the corresponding shift to the scores WLE estimates.

[SCSALE]\_EAP.se Standard error of the EAP estimates.

[SCSALE]\_oWLE When selecting to do IRT (DoIRT = Y) and centering the difficulties at 0 (CtrDiff = Y), this are the WLE scores in the original metric, prior to centering the difficulties and applying the corresponding shift to the scores WLE estimates.

[SCSALE]\_WLE.se Standard error of the WLE estimates.

Shift Shift applied to the scores when the difficulties are centered at 0 using CtrDIFF = Y.

[SCSALE]\_EAP EAP score based on the valid responses.

[SCSALE]\_WLE WLE score based on the valid responses.

## [OUTFILE]\_[BYVAR]\_SumStats.CSV and .SAV

There will be at least one of these files created with ITEM level summary statistics for each group defined by the classification (CLASSVARS) and group comparison variables (BYVARS) specified for the analysis. If no group comparison I specified, there will be a single file created. When setting the parameter NULLBYVAR = Y, it will create one of these files using each of the BYVARS, and one using none of them.

This file will have one record for each item for each CLASSVARS by BYVARS combination. However, IRT statistics are computed for the sample as a whole and therefore these will be repeated for each item, regardless of the group for which statistics are reported.

The contents of this file are the following:

ITEM Variable name for the item

Var\_Label A descriptive label for the item, as found in the data file from there the items were read.

Reverse An indicator whether the item was reversed coded for the analysis. Reversal is coded with a Y, non-reversal with N.

ISeq A sequential number for the item. This corresponds to the position of the item in the list of items entered with the ITEMS parameter.

Key A key for the item, if one is provided. If no key was provided, or the item is defined as a constructed response type item with no key provided, the files shows a –1 for that item.

[CLASSVARS] Classification variables used for the analysis.

[BYVARS] Group comparison variables used for the analysis.

Mean\_EAP Average EAP score for those in the group defined by the CLASSVARS and BYVARS. (when parameter DOIRT = Y)

Mean\_WLE Average WLE score for those in the group defined by the CLASSVARS and BYVARS. (when parameter DOIRT = Y)

CorrW\_[SCALE]\_sumscr The Pearson correlation between the scored responses to the item and the sum score. This would be the discrimination coefficient for the item under classical test theory.

CorrW\_[SCALE]\_pplus The Pearson correlation between the scored responses to the item and the pplus score.

CorrW\_[CRITVARS] The Pearson correlation between the scored responses to the item and each of the criterion variables specified with CRITVARS.

CorrW\_[SCALE]\_[SCORETYPE] The Pearson correlation between the scored responses to the item and the IRT score type requested with the parameter [SCORETYPE]. (when parameter DOIRT = Y)

SumWgts Sum of the sampling weights for the respondents in the group defined by the classification and group comparison variables.

ItemMean Average number of points on the item.

ItemStDev Standard deviation for the scored responses to the item.

ValidCases Number of valid responses, or cases with valid responses to the item.

ItemMin Minimum possible score for the item.

ItemMax Maximum possible score for the item.

NSteps Number of scored response categories for the item –1. (when parameter DOIRT = Y)

PPlus Average proportion correct on the item, calculated as the item mean divided by the maximum possible score on the item.

Alpha Item discrimination or slope parameter (when parameter DOIRT = Y)

Obeta Original item difficulty or threshold parameter (when parameter DOIRT = Y and CtrDIFF = Y)

tauCat# The step parameter for each of the categories of polytomous items (when the parameter DOIRT = Y and there are more than 2 categories for an item)

shift Shift applied to the item difficulties to make the average difficulty = 0.

beta Item difficulty or threshold parameter used for the calculation of the IRT based scores (when parameter DOIRT = Y).

POf# Probability of an average person within the group defined by the CLASSVARS and BYVARS combination to obtain a score of # on the item. At this point, if necessary, have been down coded to have a minimum score of 0. (when parameter DOIRT = Y)

EqMean Expected average score on the item of an average person within the group defined by the CLASSVARS and BYVARS combination. (when parameter DOIRT = Y)

EqPPlus Expected p-plus on the item of an average person within the group defined by the CLASSVARS and BYVARS combination. (when parameter DOIRT = Y)

Scale Name of the scale as specified with the parameter SCALE.

Scalelbl Label for the scale, as specified with the parameter SCALELBL

Weight Sampling weight used for the calculations. If none was provided, this is left blank.

Date Day when the analysis was run.

Time Time when the analysis was completed.

Indir Directory from where the input file with the item responses was read.

Infile Name of the file with the responses.

Selcrit Selection criteria for the cases, if any was applied.

## [OUTFILE]\_[BYVAR]\_ItemStats.CSV and .SAV

There will be at least one of these files created with ITEM RESPONSE level summary statistics for each group defined by the classification (CLASSVARS) and group comparison variables (BYVARS) specified for the analysis. If no group comparison I specified, there will be a single file created. When setting the parameter NULLBYVAR = Y, it will create one of these files using each of the BYVARS, and one using none of them.

This file will have one record for each possible item response, for each CLASSVARS by BYVARS combination. However, IRT statistics are computed for the entire sample as a whole and therefore these will be repeated for each item, regardless of the group for which statistics are reported.

The contents of this file are the following:

Item Variable name for the item.

Response Original response code for the item

Val\_Label Value label for the response

Var\_Label A descriptive label for the item, as found in the data file from there the items were read.

Reverse An indicator whether the item was reversed coded for the analysis. Reversal is coded with a Y, non-reversal with N.

ISeq A sequential number for the item. This corresponds to the position of the item in the list of items entered with the ITEMS parameter.

Key A key for the item, if one is provided. If no key was provided, or the item is defined as a constructed response type item with no key provided, the files shows a –1 for that item.

NCases Number of cases choosing the response option.

SumWgt Sum of the weights of those choosing the response option.

P Weighted proportion of those choosing the response option.

Q Weighted proportion of those NOT choosing the response option.

rpb\_[SCALE]\_[score] Point-biserial correlation between a variable indicating choosing the response option and the sum score and p-plus scores for the scale. If IRT is requested and/or criterion variables specified, the point-biserial correlation with these variables is also presented.

rbi\_[SCALE]\_[score] Biserial correlation between a variable indicating choosing the response option and the sum score and p-plus scores for the scale. If IRT is requested and/or criterion variables specified, the biserial correlation with these variables is also presented.

IsKey An indicator variable showing whether the response option is the keyed (correct) response option for the item. For items where no key is provided, this takes on the value of 1.

Criteria A list of the criterion variables specified with the parameter CRITVARS.

[CLASSVARS] Classification variables used for the analysis.

[BYVARS] Group comparison variables used for the analysis.

Mean\_EAP Average EAP score for those in the group defined by the CLASSVARS and BYVARS. (when parameter DOIRT = Y).

Mean\_WLE Average WLE score for those in the group defined by the CLASSVARS and BYVARS. (when parameter DOIRT = Y)

CorrW\_[SCALE]\_sumscr The Pearson correlation between the scored responses to the item and the sum score. This would be the discrimination coefficient for the item under classical test theory.

CorrW\_[SCALE]\_pplus The Pearson correlation between the scored responses to the item and the pplus score.

CorrW\_[CRITVARS] The Pearson correlation between the scored responses to the item and each of the criterion variables specified with CRITVARS.

CorrW\_[SCALE]\_[SCORETYPE] The Pearson correlation between the scored responses to the item and the IRT score type requested with the parameter [SCORETYPE]. (when parameter DOIRT = Y)

SumWgts Sum of the weights for those with valid responses to the item.

ItemMean Average score on the item.

ItemStDev Standard deviation of the scores on the item.

ValidCases Number of valid cases responding the item.

ItemMin Minimum possible score on the item.

ItemMax Maximum score on the item.

NSteps Number of scored response categories –1.

PPlus Average proportion of the maximum score on the item.

Scale Name of the scale as specified with the parameter SCALE.

Scalelbl Label for the scale, as specified with the parameter SCALELBL

Weight The sampling weight used for the calculations. If none was provided, this is left blank.

Date The day when the analysis was run.

Time The time when the analysis was completed.

Indir Directory from where the input file with the item responses was read.

Infile Name of the file with the responses.

Selcrit Selection criteria for the cases, if any was applied.

## [OUTFILE]\_[BYVAR]\_DIF.CSV and .SAV

There will be one of these files created for each of the BYVARS specified in the analysis, if any. If there are no BYVARS specified, this file will not be created.

This file will have one record for each level of the criterion specified by the parameter NLEVELS and for each CLASSVARS by BYVARS combination.

The contents of this file are the following:

Item The variable name for the item.

Var\_Label A descriptive label for the item, as found in the data file from there the items were read.

Reverse An indicator whether the item was reversed coded for the analysis. Reversal is coded with a Y, non-reversal with N.

ISeq A sequential number for the item. This corresponds to the position of the item in the list of items entered with the ITEMS parameter.

Key A key for the item, if one is provided. If no key was provided, or the item is defined as a constructed response type item with no key provided, the files shows a –1 for that item.

[CLASSVARS] classification variable specified with the parameter CLASSVARS.

Level Level assignment according to the criteria. Level –1 corresponds to those who did not have a value for the criterion variable.

[BYVARS] Comparison variable as specified with the parameter BYVARS.

NCases Number of cases in the level.

SumWgts Sum of the weights for the cases in the level.

PPlus Proportion correct on the item for persons in the level.

Criteria Criterion variable used for the assignment to levels.

## Additional Files

In addition to the files listed above, there is a series of files that are created and saved during the processing. These are the following:

[OUTFILE]\_[BYVAR]\_ItemStats4Graphs.CSV and .SAV

This file has the same contents of the [OUTIFLE]\_[BYVAR]\_ItemStats file but re-arranged and trimmed in a format to be used by the SPSS GRAPH command to create the plots included as part of the output of IAPlus.

[OUTFILE]\_PersonScores.CSV and .SAV

This file has the response records processed with TAM, with their corresponding score variables. The information included in this file is incorporated in the [OUTFILE]\_Scored.SAV file described earlier.

[OUTFILE]\_IRTStats.CSV and .SAV

This file has one record per item processed with TAM. It contains the IRT item parameters calculated within TAM. The information included in this file is incorporated in the [OUTFILE]\_SumStats.CSV and .SAV files.

[OUTFILE]\_Responses.CSV

The response file that was used as the input file for TAM. This response information is all contained in the [OUTFILE]\_PersonScores.CSV and .SAV files.

# Parameters

What follows is a list of the parameter that can be used when calling IAPlus. They are in the logical order in which you would want to consider them, and not in alphabetical order. But the order of the parameters in the call can be any order. Some parameters are optional, some are mandatory, and some have default values. This is all indicated below.

Depending on the parameter, the values assigned can be one or many. They are not case sensitive. The parameters are assigned a value or values using the equal sign. When more than one value is specified for a parameter, these need to be separated with spaces, and the values from one parameter to the next are always separated by a forward slash (‘/’).

## Data Input and Output

INDIR (not optional; no default)

Directory with response file. The directory location must be fully spelled out and enclosed in quotes.

Example:

INDIR = "c:\temp\itemanalysis\example" /

INFILE (not Optional; no default)

File with item responses. It does not need the SAV extension for the file name. Do not use quotes.

Example:

infile = SampleDataPIRLS /

OUTDIR (not optional; no default)

Directory where to write the results. The directory location must exist and be accessible from your computer, be fully spelled out and enclosed in quotes. This directory is also used to save many temporary files created by the program, therefore we recommend using a local directory so that the program will run with optimal performance.

Example:

outdir = "c:\Temp\ItemAnalysis\Examples" /

OUTFILE (not optional; no default)

Name used for the output files. Presented as [OUTFILE] in the descriptions above. This value cannot start with a number, cannot have special characters or spaces, and must comply with the SPSS variable naming convention.

Example:

outfile = PIRLSExample /

SELVAR (optional; no default)

A variable or list of variables used to subset the data read from the INFILE. The variable(s) used to subset the data can be one of the ITEMS, the CLASSVARS, the BYVARS, IDVARS or CRITVARS. Therefore, you only need to specify the SELVAR when this is not specified anywhere else.

Example:

selvar = var1 var2 /

SELCRIT (optional; no default)

The selection criteria to apply when reading the INFILE. Only records that meet the specific selection criteria are read and accounted for in the processing, therefore the sequence number of the records in the INFILE and the working file might differ when this parameter is used.

The selection criteria must be specified using SPSS IF/THEN logic, but no IF is needed. Any, and all operations must fit in a single line ending with a forward slash (“/”).

The selection criteria is applied whenever SELCRIT is not blank, regardless of whether the SELVAR parameter if left blank or not.

Example:

Selcrit = (var1 = 1 and var2 > 4) /

In the example above, only cases with VAR1 = 1 and VAR2 > 4 will be read from the INFILE. Parenthesis are used for readability but are not necessary.

## Scale

SCALE (optional; no default)

Name for the test or scale that defines the ITEMS. This value is used as the root of the different score variables created by IAPlus. Only one name is allowed, and the name must conform to SPSS variable naming convention.

When it is left blank, it will assume the ITEMS should NOT be treated as forming a scale with the rest of the items and therefore no scores will be computed, no factor analysis or IRT calibration will be done. In this case a CRITVAR is expected in the call, and item statistics will be computed using the CRITVAR as a reference.

Example:

scale = reading /

SCALELBL (optional; default: “<no label provided>”)

This specifies the variable label to be used for the different score variables created by IAPlus. It can be up to 96 characters long and should not contain special characters.

Example:

scalelbl = PIRLS Reading Score /

ITEMS (not optional; no default)

The list variable names for the items in the scale. These need to be listed one after another, with spaces in between. You can use multiple lines, but none should begin in the first column. The last variable name should be followed by a forward slash. The items will be read, processed, and assigned a sequence number (ISEQ) based on the order in which they appear on this list, regardless of the order of the items in the INFILE. Items can only be numeric.

Example:

Items = R011C01C R011C02C R011C03C R011C04M R011C05M R011C06C

R011C07M R011C08C R011C09M R011C10C R011C11C R011C12M

R011C13M R011F01M R011F02M R011F03M R011F04M R011F05M

R011F06C R011F07C R011F08C R011F09C R011F10C R011F11M

R011F12C R011F13M /

## Response Processing and Coding

KEYS (optional; no default)

A sequence of numeric values, usually 1 thru 5, separated by a space indicating the value of the correct response for the item. When used, there should be one entry for every item listed in ITEMS parameter. When left blank, or entered as an X, it assumes the item responses have already been scored and there is no need to apply a key to the item. A scale could be composed of items that require a KEY and items that do not. When applying the value of the KEY, the program recodes the responses = KEY to 1, and the rest to 0 or not administered depending on the scoring scheme specified with the scoring related parameters.

Example:

Keys = X X X 1 3 X 2 X 4 X X 2 3 2 3 4 1 4 X X X X X 1 X 2 /

This example shows the scoring keys for a set of items that include already scored items, and items that need a key to be applied. When working with all open-ended items, already scored items, or items that do not require a key, such as is usually the case of opinion or attitudinal scales, you can use a series of Xs, or leave blank.

Whenever applying a KEY, we recommend using QCSCORING = Y to ensure the keys were entered and applied correctly.

IsZERO...IsFIVE (optional; default = 0, 1, 2, 3, 4 and 5, respectively)

When there is no KEY provided, or the item is an open-ended item, these represent the values that are to be recoded to 0, 1, 2, 3, 4 or 5, respectively. These need to be entered using the logic used within the SPSS RECODE command to specify the “old-value” list. The recoding is applied uniformly to all items. This command should not be used to reverse item scoring. To reverse item scoring use the REVERSE parameter.

Example:

IsZero = 70 thru 79 /

IsOne = 10 thru 19 /

IsTwo = 20 thru 29 /

In the example above, values of 70 to 79 will be recoded to 0, values of 10 thru 19 recoded to 1, and values of 20 to 29 will be recoded to 2.

MAXSCRS (optional; default = 1 for all ITEMS)

The maximum number of points achievable for each item. When omitted, it assumes 1 for all items. As with the parameter ITEMS, KEYS, REVERSE, there should be one entry per ITEM, or none. The value of MAXSCRS is used to calculate the p-plus for the response to the item.

Example:

Maxscrs = 1 1 1 1 1 1 1 2 1 3 2 1 1 1 1 1 1 1 1 2 1 2 1 1 2 1 /

In the example above, there is a combination of 1-, 2- and 3- point items. To verify that the maximum score has been entered and applied correctly, you can check the minimum and maximum values for the ITEMS\_p variables shown in the output file (\*.SPV). When properly entered and applied, the maximum value should not exceed 1. The minimum value will depend on the lowest possible score for the responses to the item. For example, a Likert scale type item, scored 1 thru 4, will have a minimum p-plus value of 0.25 (or 1/4), and a maximum p-plus of 1.0 (or 4/4), whereas a test item, scored 0 through 2, will have a minimum p-plus value of 0.0 and a maximum of 2.

DOWNCODE (optional; default = N; options = Y or N)

This option reduces the value of the scored item responses by 1. It is necessary when running IRT with TAM and working with items that have a minimum scored value of 1. When DOWNCODE is set equal to Y, ALL items are down coded. The down code is applied only to the responses in the file that is used by TAM. All other instances of the scored responses in the files and the output use the original score values.

Example:

downcode = Y /

REVERSE (optional; no default; options = Y or N)

This is a series of N or Y indicating whether the scoring of the item should be reversed. When omitted it assumes none of the items need to be reversed.

Example:

reverse = Y Y Y Y Y N /

In the example above, there are 6 items. The first 5 are to be reverse coded. The last one is not. The reversal of the scored responses is done according to this formula: MAXSCRS[ITEM\_S] + MINRSCORE - ITEM\_S

MINRSCORE (optional; default = 1)

The minimum score for the items. This parameter is used only when an item needs to be reverse coded. It is assumed to be the same for all items.

Example:

minrscore = 1 /

In the example above, the minimum score on the items is set to 1, and therefore responses to an item with original scores of 1, 2, 3 and 4 will be reverse coded to 4, 3, 2 and 1, respectively.

ATLEAST (optional; default = 1)

Indicates the minimum number of valid responses required for a case to be used in the analysis. There always needs to be at least one valid response for a case to be processed. Valid responses are any possible responses, except those coded as not-administered.

Example:

atleast = 3 /

In the example above, only records with at least 3 valid responses will be used for the analysis. Records excluded from the analysis for not having the minimum number of valid responses are coded with the variable [SCALE]\_use = 0 in the scored file. All other records used for the analysis have [SCALE]\_use = 1. The output file shows a crosstabulation of cases using the variables [SCALE]\_taken, the number of valid responses to the items, with the case use indicator [SCALE}\_use.

Note that the selection of cases based on the ATLEAST parameter is done AFTER all recoding of the responses is done to the items using the parameters described in this section.

OMIT (optional; no default; options = numeric response values)  
OTHERMIS (optional; no default; options = numeric response values)   
NOTREACH (optional; no default; options = numeric response values)

The value(s) in the response variables that are used to indicate the response to an item that was administered, yet not responded, not reached, or missing for some other reason.

Example:

omit = 9 99 /

othermis = 7 97 /

notreach = 6 96 /

In the example, values 9 and 99 represent omitted responses, 7 and 97 other missing responses (could be, for example, double responses to multiple choice questions), and 6 and 96 responses to items that were not reached.

OMITAS (optional; no default; options = single numeric response value, usually 0 or sysmis)  
NOTREACHAS (optional; no default; options = single numeric response value, usually 0 or sysmis)  
OTHERMISAS (optional; no default; options = single numeric response value, usually 0 or sysmis)

The value to assign responses that were coded as omitted (OMIT), not reached (NOTREACH) or missing for some other reason (OTHERMIS). In general, these responses are usually coded as 0 or system missing. But any other numeric value is allowed.

Example:

omitas = 0 /

othermisas = 0 /

notreachas = sysmis /

In the example above, omitted and other missing responses are recoded to 0, and not reached responses are coded as system missing, and therefore excluded from the analysis.

NOTADMIN (optional; no default)

The value in the response data that is used to identify responses to items that were not administered. By default, responses coded as SYSMIS are treated as not administered and excluded from the analysis.

Example:

Notadmin = 8 98 /

In the example above, value 8 and 98 are used to represent responses to the items that were not administered.

## Grouping and Reporting

USESUMSCR (optional; default = N; options = Y or N)

Whether to use the sum score ([SCALE]\_sumscr) as one of the criterion variables in the reporting statistics.

Example:

usesumscr = N /

CLASSVARS (optional; no default; options = any variable(s) in the INFILE)

A single variable, or a list, that will be used to report item statistics calculated with IAPlus. When multiple CLASSVARS are specified, the item statistics are computed for each of the possible groups available in the data crossing the values of the CLASSVARS, in combination with each of the BYVARS, if any are provided.

Example:

classvars = idcntry itlang /

In the example above the analysis will be done for each combination of the values in the variable IDCNTRY and ITLANG.

BYVARS (optional; no default; options = any variable(s) in the INFILE)

A single variable, or a list, that will be used to report item statistics calculated with IAPlus. When multiple BYVARS are specified, each of them is treated separately from each other and results are not crossed between them. The BYVARS will also be used to create graphics where statistics between groups can be compared, and therefore are also referred to in this documentation as comparison variables.

Example:

classvars = idcntry itlang /

byvars = itsex itbook /

In the example above the item statistics will be compute broken down by the variable ITSEX, and the variable ITBOOK. If this is used in combination with CLASSVARS = IDCNTRY ITLANG, two sets of results will be created: one for each category combination created by the variables IDCNTRY \* ITLANG \* ITSEX and a second one for each category combination created by the variables IDCNTRY \* ITLANG \* ITBOOK. Notice that ITSEX and ITBOOK are no crossed.

NULLByVar (optional; default = N; options = Y or N)

When specifying BYVARS, whether to also create reports for the overall group, or the groups defined by the CLASSVARS, or not.

Example:

classvars = idcntry itlang /

byvars = itsex itbook /

Nullbyvar = Y /

In the example above, in addition to reporting the results for each combination of the CLASSVARS with the BYVARS, it reports the results for the CLASSVARS alone.

DOTABLES (optional; default = Y; options = Y or N)

This parameter generates a set of tables and graphs presenting the statistics computed. In addition to calculating a series of item statistics, IAPlus also provide the option to create bar and line charts with the statistics. The bar and line charts are saved to an HTML and an XLSX file using the parameters DOHTML and DOXLSX, respectively. These files are described in the “Output Files” section of this manual.

Example:

dotables = Y /

The tables created using this command are created using CTABLES. This is sometimes an optional module in SPSS. If CTABLES is not installed, the program will run and calculate the statistics correctly, but will omit these tables and issue a warning.

## Supplementary Variables

IDVARS (optional; no default; options = any variable(s) available in the INFILE)

Identification variable(s) that you would want to have included in the [OUTFILE]\_Scored file. These variables are simply copied over and are not processed. These are useful to later match the scores calculated within IAPlus with additional variables in the INFILE.

Example:

idvars = idschool idstud /

CRITVARS (optional; no default; options = any variable(s) available in the database)

In addition to scores generated by IAPlus (sum score and the proportion correct on the set of items), IAPlus will also calculate correlation coefficients between the item responses and each of the CRITVARS, as well as biserial and point-biserial correlations between each of the response options and each of the variables specified as CRITVARS.

Example:

Critvars = Variable1 Variable2 /

When SCALE is left blank, indicating the items do not conform a scale and should not be combined as such, the program will require one or more CRITVARS.

WGTVAR (optional; no default; options = a single variable with the survey or sampling weight)

The name of the variable with the survey or sampling weight. When no WGTVAR is specified, IAPlus assigns all cases a weight of 1, therefore all response records contribute equally to the calculations. When a WGTVAR is specified, this is used to weight the response records. Cases with zero, missing and negative sampling weights are excluded from the analysis from the outset.

Example:

wgtvar = TOTWGT /

NORMWGT (optional; default = N; options = Y or N)

Using this option normalizes weights so they add to 100 within each group defined by the crossing of the classification variables. If no classification variables are specified, then no normalization takes place. Normalization is done even when no WGTVAR is specified. Since all item statistics are calculated within the groups defined by the CLASSVARS, the normalization affects only TAM analysis when the entire set of records is processed to calculate IRT based item parameters and related statistics.

Example:

normwgt = Y /

## Quality Control and Output Control

DOHTML and DOXLSL (optional; default = Y; options = Y or N)

Writes the output from the DOTABLES parameter to HTML or XLSX formatted files.

Example:

dohtml = Y /

doxlsx = Y /

QCSCORING (optional; default = Y; options = Y or N)

Using this option generates writes to the output tables crossing the original response code with the scoring of the items and assignment of proportion correct. You will want to use this option to verify that the items are being scored correctly, omitted and other missing responses are being assigned the correct code, and items are reversed accordingly, if reversal was requested.

Example:

qcscoring = Y /

CHKCRITERIA (optional; default = Y; options = Y or N)

This parameter triggers the program to display descriptive statistics for the criterion variables created by IAPlus, and those provided using the parameter CRITVARS. It also creates frequency histograms and correlation coefficients for these. It is very useful to help with the quality control of the data.

Example:

chkcritria = Y /

BISERIALN (optional; default = 5)

The minimum number of cases choosing a response option to calculate the biserial and point-biserial correlation for that response option, within the group defined by the CLASSVARS and BYVARS. In addition to using this criteria, biserial and point-biserial correlations are not calculated for extreme cases such as p<0.01 and p>0.99.

Example:

biserialn = 5 /

VIEWCOD (optional; default = N; options = Y or N)

Expands and displays in the output file the code generated by the macro. Useful for troubleshooting the processing and identifying where exactly problems or warnings occur, if any.

Example:

viewcod = N /

GRAPHN (optional; default = 10)

The minimum number of cases for a point to be shown in the graphics created and reported with IAPlus. This value refers to the actual number of cases, and not to the sum of the weights of the cases in a group.

Example:

graphn = 5 /

CLEAN (optional; default = Y; options = Y or N)

Instructs the program to delete (CLEAN = Y) the temporary files created during processing. During processing, several temporary files are created and stored in the OUTDIR. At times, preserving and viewing these files could be useful for troubleshooting the code or identify problems with the data.

Example:

clean = Y /

## Factor Analysis Definition

FACTOR (optional; default = N; options = Y or N)

Instructs the program to do a factor analysis with the response data for the ITEMS. It uses the variance covariance matrix as input, and therefore it can handle an incomplete data matrix as long as the variance covariance matrix is complete. Unless otherwise specified with the parameters FEXTRACT and FROTATE, is uses a PC extraction and conducts a varimax rotation, when more that one factor is extracted.

Example:

factor = Y /

This factor analysis is done across all groups and is only intended to give a general overview of the factor structure captured by the response matrix.

FEXTRACT (optional; default = PC; options = any extraction option available with SPSS)

Specifies the factor extraction technique. It only takes effect when FACTOR = Y.

Example:

fextract = PC /

FROTATE (optional; default = VARIMAX; options = any rotation option available with SPSS)

Specifies the factor rotation method. It only takes effect when FACTOR = Y.

Example:

frotate = varimax /

## Differential Item Function

DODIF (optional; default = Y; options = Y or N)

When using BYVARS, this option will create and display line charts contrasting the proportion correct for each of the items, by levels in the scores and criterion used by IAPlus, for each of the groups defined by the corresponding BYVARS. The number of levels can be set using the parameter NLEVELS, or the parameter IRTCUTS.

Example:

dodif = Y /

In addition to the line charts, it created a file with the corresponding statistics.

NLEVELS (optional; default = 6)

The number of levels to use for the scores and criterion variables processed by IAPlus when creating the line charts generated by the option DODIF = Y.

Note that when there are too few cases in a level, the point is suppressed for that level, therefore you will want to choose enough levels to show the trend, yet not too many levels that results in very few cases by level for any one group. Between 4 and 6 levels is usually a useful number of levels.

The levels are created by categorizing the distribution of scores or the criterion variable in equally wide segments represented by the observed scores of the respondents in the file. For example, if a distribution for a score ranges from exactly -3.0 to exactly +3.0, and 6 levels are requested, the levels will be (1) from -3 to -2, (2) from -2 to -1, (3) from -1 to 0, and so on.

Example:

nlevels = 6 /

IRTCUTS (optional; no default)

At times, when using the option DODIF = Y in combination with DOIRT = Y, you might want to specify the cut-points used for the line displays, rather than use equally wide segments of the distribution. In this case you can make explicit the cut-points you want to use. The number of levels will be the number of cut-points provided plus 1.

These cut-points apply ONLY to the IRT score calculated with TAM (either EAP or WLE), and only affect the number of levels for this score, and not the other scores. When using this parameter, you can use positive or negative cut-points. The negative numbers need to be enclosed in quotes, otherwise they will not be process correctly. The cut-points also need to be provided in ascending order. Assuming the abilities and item difficulties are well aligned, and the abilities are somewhat normally distributed, scores will be on a metric of mean = 0 and standard deviation = 1, and therefore using IRTCUTS = “-2” “-1.0” 0.0 1.0 2.0 / will results in creating levels that represent about 1 standard deviation of scores of the distribution.

Example:

irtcuts = "-2.0" "-1.0" 0.0 1.0 2.0 /

## Item Response Theory

DOIRT (optional; default = N; options = Y or N)

IAPlus has the capability of using the TAM package in R to conduct an IRT type calibration with the ITEMS provided. This is done by setting the parameter DOIRT=Y. This IRT calibration is intended for quality control purposes and to evaluate the set of items analyzed. Therefore, IAPlus is only intended to be used as an IRT program to the extend that it can conduct a calibration with the set of items and cases provided. For more information about using TAM and its many options, please refer to the TAM manual.

When using TAM, it calculates item parameters for the items in the scale, and WLE and EAP scores for the respondents. It uses the default parameters. It does a single IRT calibration across the entire sample, using sampling weights, if any are provided.

A sample call looks like this:

results <- TAM::tam.mml(responses[,UseTheseItems],

irtmodel="PCM",

pweights = responses$pwgt,

pid=responses$CaseSeq,

control=list(nodes=seq(-4,4,len=41), convD=.001, conv=.001,

convM=.001), constraint="cases")

When using TAM, you will need to indicate with one of the parameters the location of the R executable (WHEREISR). You can also choose whether to calibrate using a 1PL or a 2PL model (IRTMODEL), and whether to use EAP or WLE scores for the subsequent analysis (SCORETYPE).

IAPlus create an R syntax file saved to the OUTDIR, and called “[SCALE]\_TAMcode.R”. The processing log is stored in a file called “[SCALE]\_TAMcode.Rout”. If there are errors in the processing of the data with TAM, you will find these listed in the \*.Rout file.

If the response data start with 1, you need to use the DOWNCODE option. For polytomous items, the categories used must be sequential. For example, 0, 1, 2. A item with response categories 0, 2, 4 will cause the program to crash.

The R code will exclude from the calibration any item that does not vary (all respondents use only one response category) and any item that was coded as not-administered across the entire sample. While these items will be processed by IAPlus and shown in the output, they are excluded from all IRT related calculations.

Example:

Doirt = Y /

SCORETYPE (optional; default = WLE; options = WLE or EAP)

The choice of IRT score calculated with TAM to use as criteria for biserial, point-biserial correlations, and discrimination coefficient.

Example:

scoretype = WLE /

IRTMODEL (optional; default = 1PL; options = 1PL or 2PL)

When using TAM, the choice of IRT model to use for the calibration and scoring. When choosing 1PL it will use irtmodel=”PCM”, and when choosing 2PL it will use irtmodel=”GPCM”.

Example:

Irtmodel = 1PL /

CTRDIFF (optional; default = Y; options = Y or N)

When doing IRT with TAM, the calibration is generally run with no constraints on the distribution of abilities or item difficulties. This option can be used to center the item difficulty estimates around 0, and to shift the ability estimates accordingly. As a result, you might be able to compare the results from the calibration of the same set of items performed with different samples.

Example:

ctrdiff = Y /

WhereIsR (optional; no Default)

This is the directory where RScript.exe file is located or accessible from your computer. This parameter is necessary if you are using the option DOIRT = Y; unless this directory is already in your PATH environment variable. The directory location should be enclosed in quotes.

Example:

WhereIsR = "C:\Program Files\R\R-4.0.2\bin" /

# Troubleshooting and Bug Report

To report bugs, or troubleshoot the program, please send details in an email to [egonzalez@ets.org](mailto:egonzalez@ets.org) or [eugene.gonzalez@iea-hamburg.de](mailto:eugene.gonzalez@iea-hamburg.de). When troubleshooting, please set VIEWCOD = Y / and submit process the syntax. Accompany your email with the corresponding \*.sps and \*.sps files, as well as any additional information you might consider relevant.

# Examples

Two examples accompany this User Manual. In the first one, the syntax file is called “PIRLSBook1and2.sps”, and the corresponding output files are all called “PIRLSExample\_\*”. In this we analyze 2 blocks of items from the PIRLS 2016 reading assessment.

In the second example the syntax file is called “ICCSImmig.sps” and the corresponding output files are all called “ICCImmig\_\*”. In this second example we analyze the attitude towards immigrants from the ICCS 2013 survey.