Mplus syntax builder
for testing forced-choice data with the Thurstonian IRT model

User guide

Introduction

Brown and Maydeu-Olivares (2011) have recently introduced an item response theory (IRT) model capable of modeling responses to any forced-choice questionnaire. The Thurstonian IRT model is a multidimensional item response model with some special features that can be straightforwardly estimated using the general modeling software M*plus* (Muthén & Muthén, 1998-2010). However, programming these models in M*plus* is tedious and error-prone except for very small models, as one need to impose parameter constraints that reflect the within-block patterned relationships among items. However, the model is conceptually so simple that the M*plus* programming can be easily automated.

This is a user guide to a very simple Excel macro that writes the M*plus* syntax necessary to fit the IRT model to any forced-choice questionnaire. Furthermore, a detailed tutorial on how to model different types of forced-choice questionnaires and how to score respondents on the measured attributes is provided by Brown and Maydeu-Olivares (2012).

In this user guide, we describe how to code responses to forced-choice questionnaires and how to build Mplus syntax files for different forced-choice designs. We cover different block sizes (items presented in pairs, triplets, quads) and their common and specific features. We cover both full ranking and partial ranking designs.

Preparing the data file
Binary coding of pairwise preferences

M*plus* syntax for the Thurstonian IRT model requires the forced-choice responses to be coded using binary outcomes (dummy variables).

The outcome of any pairwise comparison {A, B} is coded

1, if item A was preferred to item B

0, if item B was preferred to item A

In other words, it is coded 1 if the first item in the pair was preferred to the second, and it is coded 0 otherwise.

How to code full ranking blocks

Any rank ordering of n items can be equivalently coded as n(n-1)/2 pairwise comparisons using the binary coding above. For instance, to code a rank ordering of 4 items A, B, C and D, we need to consider outcomes of 6 pairwise comparisons:

{A, B}   {A, C}   {A, D}   {B, C}   {B, D}   {C, D}

Then, the ordering {B, A, D, C} can be equivalently coded using binary outcome variables as follows:

{A,B}=0   {A,C}=1   {A,D}=1   {B,C}=1   {B,D}=1   {C,D}=0

Or simply

0 1 1 1 1 0

How to code partial ranking blocks ('most'-'least' response format or similar)

Sometimes respondents are only asked to report one item that best describes them and one that least describes them. Such partial ranking formats result in missing binary outcomes whenever the block size is 4 items or more.

For instance, if out of 4 items A, B, C and D the item B was selected as 'most' and item C was selected as 'least', the ordering of items A and D is not known:

{A, B}=0   {A, C}=1   {A, D}=*   {B, C}=1   {B, D}=1   {C, D}=0

Or simply

0 1 * 1 1 0

The respondents might be only asked to report one item that best describes them. Such partial ranking formats result in missing binary outcomes whenever the block size is 3 items or more.

For instance, if out of 3 items A, B, and C, the item B was selected as 'most', the ordering of items A and C is not known:

{A, B}=0   {A, C}=*   {B, C}=1

Or simply

0 * 1

Converting the rank orders or 'ipsative points' to binary outcomes

If the forced-choice data have been recorded using rank orders of items within each block, or reversed rank orders as is often the case with already "ipsative scored" items, the responses should be recoded as binary outcomes of pairwise comparisons before submitting them to the M*plus* analysis. This recoding can be easily performed using standard statistical software prior to modeling with M*plus*. For instance, in SPSS conditional statements are used:

```
IF (i1>i2) i1i2=1.
IF (i1<i2) i1i2=0.
RECODE  i1i2  (SYSMIS=99).
EXECUTE .
```

Alternatively, DEFINE commands can be used to recode the data within M*plus*.

For rank-orderings, binary outcomes of all pairwise combinations of *n* items within each block are computed as differences between the ranks of the second and the first items in the pair, as follows:

```
DEFINE:
i1i2 = item2-item1;
i1i3 = item3-item1;
etc.
```

For 'ipsative item scores', which are inverted rank orders, we use

```
DEFINE:
i1i2 = item1-item2;
i1i3 = item1-item3;
etc.
```

Then, all outcomes are cut at 0, creating binary variables with value 1 if the difference was greater than 0, and value 0 otherwise.

```
CUT i1i2 i1i3 … (0);
```

For incomplete rankings, outcomes of comparisons where preferences between items are not known should be coded as missing data, using conditional statements, for example:

```
IF (i2 GT i1) THEN i1i2=1;
IF (i2 LT i1) THEN i1i2=0;
IF (i2 EQ i1) THEN i1i2=_MISSING;
```

In addition, when missing data is present, the missing responses have to be imputed prior to model estimation. This is described in Example 2 below.

STEP-BY-STEP TUTORIAL FOR CREATING M*PLUS* SYNTAX

STEP 1. ENTERING FILE NAMES AND BASIC INFORMATION ABOUT THE QUESTIONNAIRE DESIGN

When the Excel application is first opened, the 'Step 1' page is presented.

This page requires as input the name of the data file containing the binary outcomes (the data file may contain additional variables), the name of a file to save the respondents scores (this is optional), the number of forced-choice blocks in the questionnaire and the block size. The page looks as follows:

All fields that require information are marked with an asterisk. If no file name is supplied for saving the people's scores, no corresponding line of Mplus syntax will be created.

After the user presses the 'Next' button, the 'Step 2' page is presented.

STEP 2. ENTERING THE NUMBER OF MEASURED ATTRIBUTES AND THE QUESTIONNAIRE KEY

At Step 2, the user is required to enter the number of attributes measured by the questionnaire.

Next, the user specifies whether the data contains full or partial rankings by selecting either 'full' or 'partial' from the dropdown list. This question only appears if the block size is 3 or more.
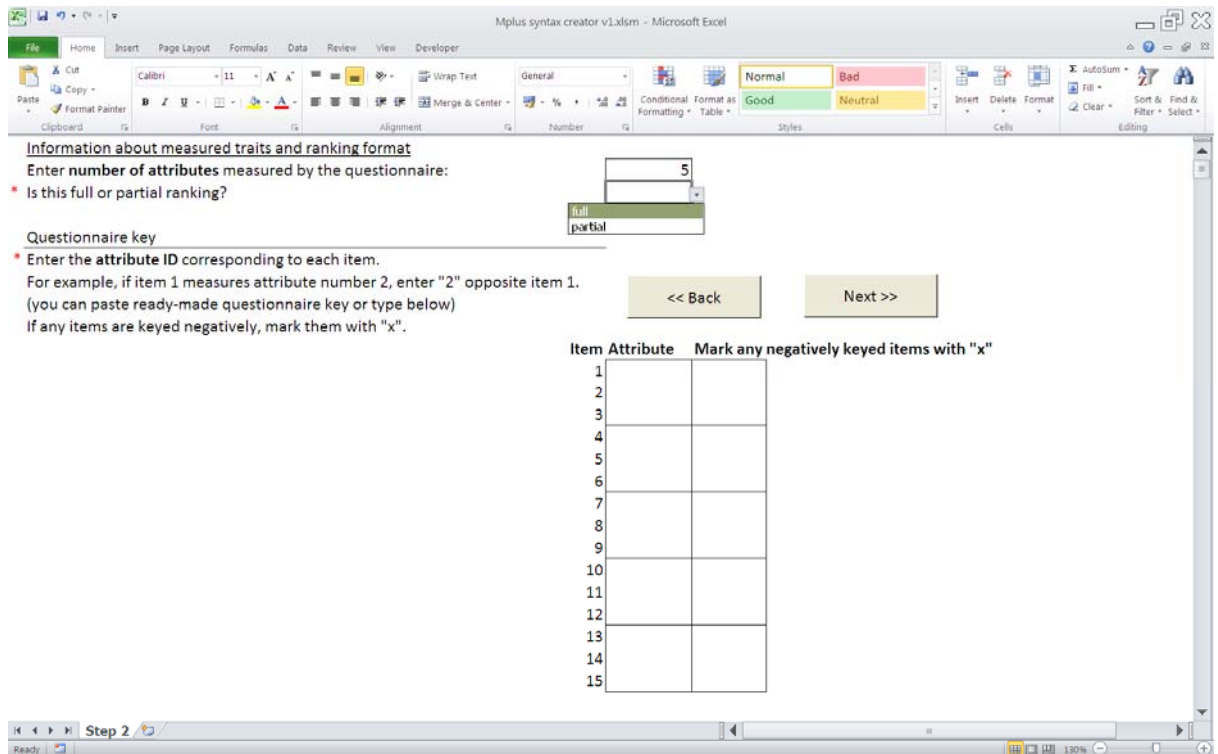
Next, there is a table for inserting the questionnaire "key". The "key" is simply a numbered list of all questionnaire items, and the user has to indicate which attribute (referred to by its number) each item measures. The macro also has an option to indicate any negatively keyed items. These are items designed to represent low attribute scores, such as "I keep in the background" to indicate Extraversion. This information is optional and is only used for assigning better (negative) starting values for factor loading parameters.

Pressing the 'Next' button takes the user to the 'Step 3' page.

STEP 3. ENTERING EXPECTED CORRELATIONS BETWEEN ATTRIBUTES

Step 3 is optional and enables the user to provide starting values for the attribute correlation matrix.

STEP 4. VIEWING AND COPYING MPLUS SYNTAX

With all previously supplied information, the Excel macro creates the full Mplus syntax, which can be viewed immediately in Excel, and also copied to a ready-to-execute Mplus input.

To copy the syntax, highlight all lines of syntax (it is only written in one column) and copy it to the clipboard. The copied text can be pasted directly into an Mplus input file (.inp) and saved.

The user may now close the excel application, or, alternatively press the 'Start over' button to return to the Step 1 and create a new Mplus syntax.

EXAMPLES

EXAMPLE 1. BUILDING SYNTAX FOR BLOCKS OF 3 ITEMS USING FULL RANKING FORMAT

This is an example of creating Mplus syntax for testing forced-choice data arising from blocks of 3 items, using the **full ranking** format. In this case, the format is full ranking if the respondents are asked to rank order items, or to select one item that describes them most and on that describes them least. In both cases all outcomes of pairwise comparisons between items are known.

In this simple example, 3 attributes are measured by 4 blocks. The assignment of items to measured attributes is as follows:

Item	Attribute	Keying
1	1	+
2	2	+
3	3	+
4	1	-
5	2	+
6	3	+
7	1	+
8	2	+
9	3	-
10	1	+
11	2	-
12	3	+

The technical detail for this design can be found in Brown and Maydeu-Olivares (2012), Example 1.

First, we enter the basic detail of this design into the 'Step 1' form and press 'Next':
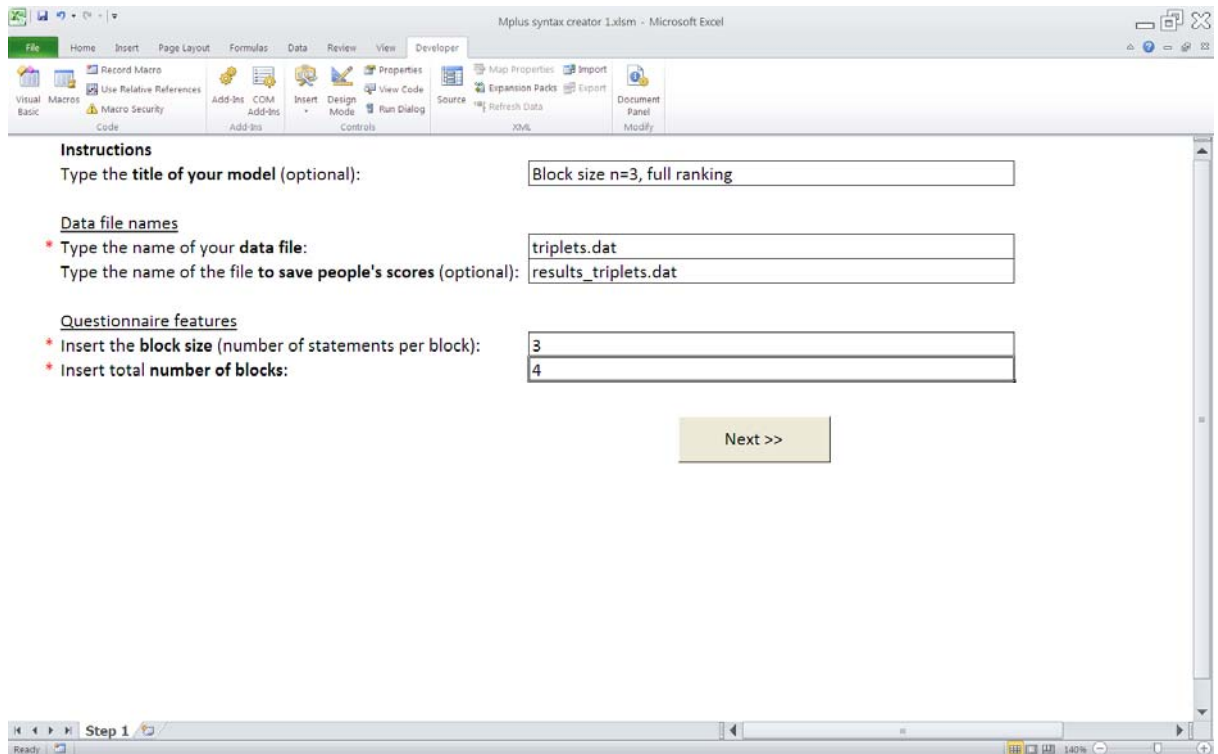
Next, we specify that the number of measured attributes is 3, the data is full ranking, and enter the questionnaire "key", as follows:
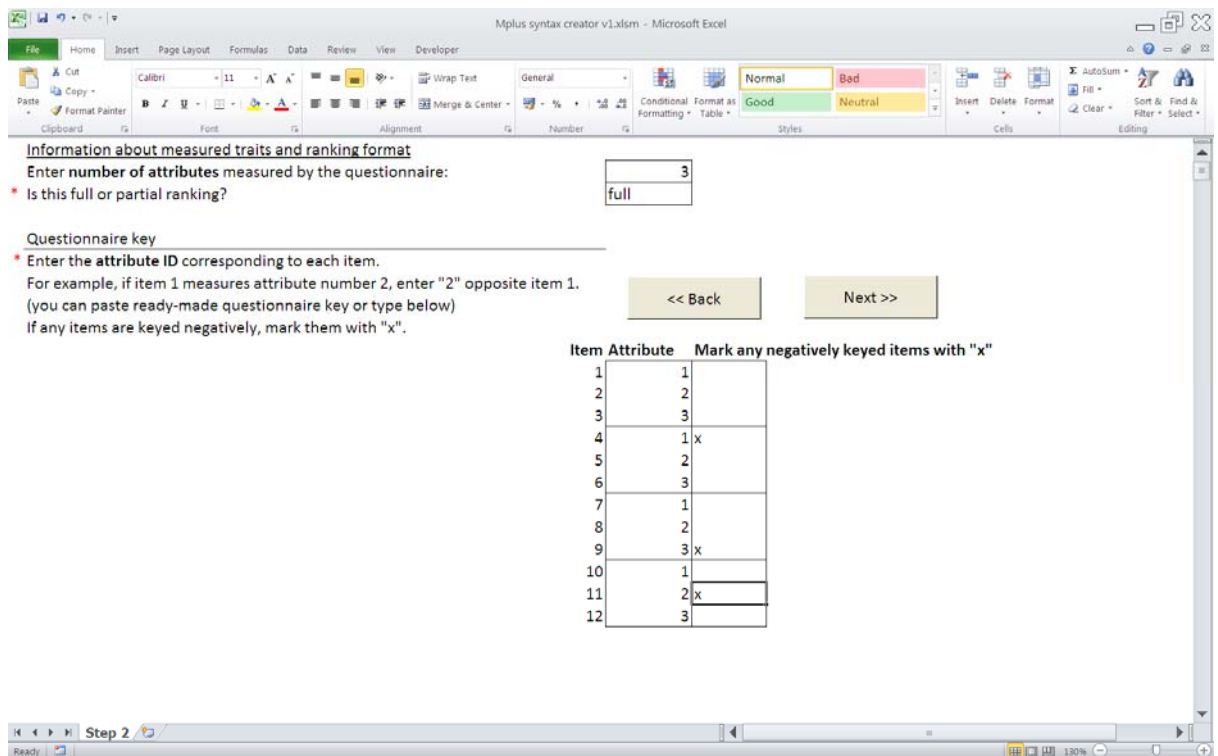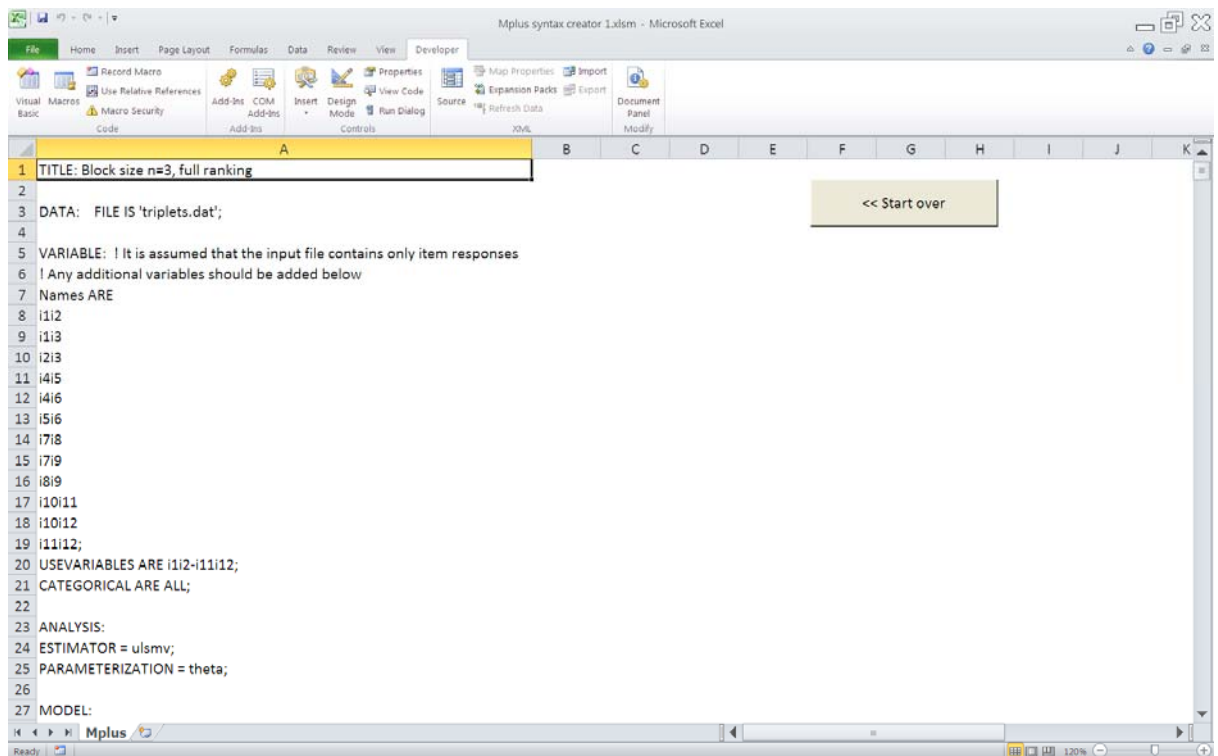

After pressing the 'Next' button, we can enter expected correlations between the measured attributes. This step is optional. In the below screen shot, we do not change default zero correlations.
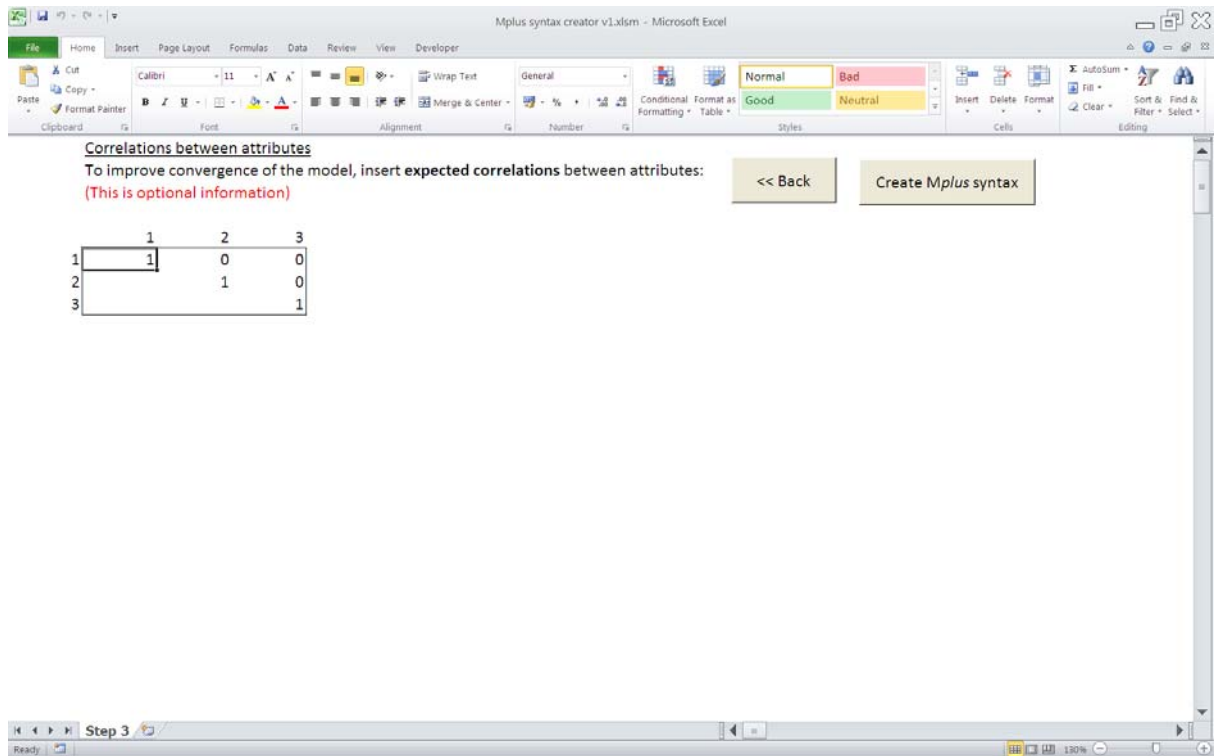
After pressing the 'Next' button, we can view the Mplus syntax, and copy and paste it to an Mplus input file ready for execution.

EXAMPLE 2. BUILDING SYNTAX FOR BLOCKS OF 4 ITEMS USING 'MOST-LEAST' RANKING FORMAT

This is an example of creating Mplus syntax for testing forced-choice data arising from blocks of 4 items, using the **partial ranking** format. Partial ranking arises in blocks of 4 items if the respondents are asked to select one item that describes them most and on that describes them least. In this case, the outcome of one pairwise comparison is not known – this is the comparison between items that are not selected as 'most' or 'least'.

In this simple example, 4 attributes are measured by 3 blocks. The assignment of items to measured attributes is as follows:

Item	Attribute	Keying
1	1	+
2	2	-
3	3	+
4	4	+
5	1	-
6	2	+
7	3	+
8	4	+
9	1	+
10	2	+
11	3	-
12	4	+

The technical detail for this design can be found in Brown and Maydeu-Olivares (2012), Example 2.

First, we enter the basic detail of this design into the 'Step 1' form and press 'Next':


Next, we specify that the number of measured attributes is 4; the data is **partial** ranking, and enter the questionnaire "key", and indicate negatively keyed items (optional) as follows:
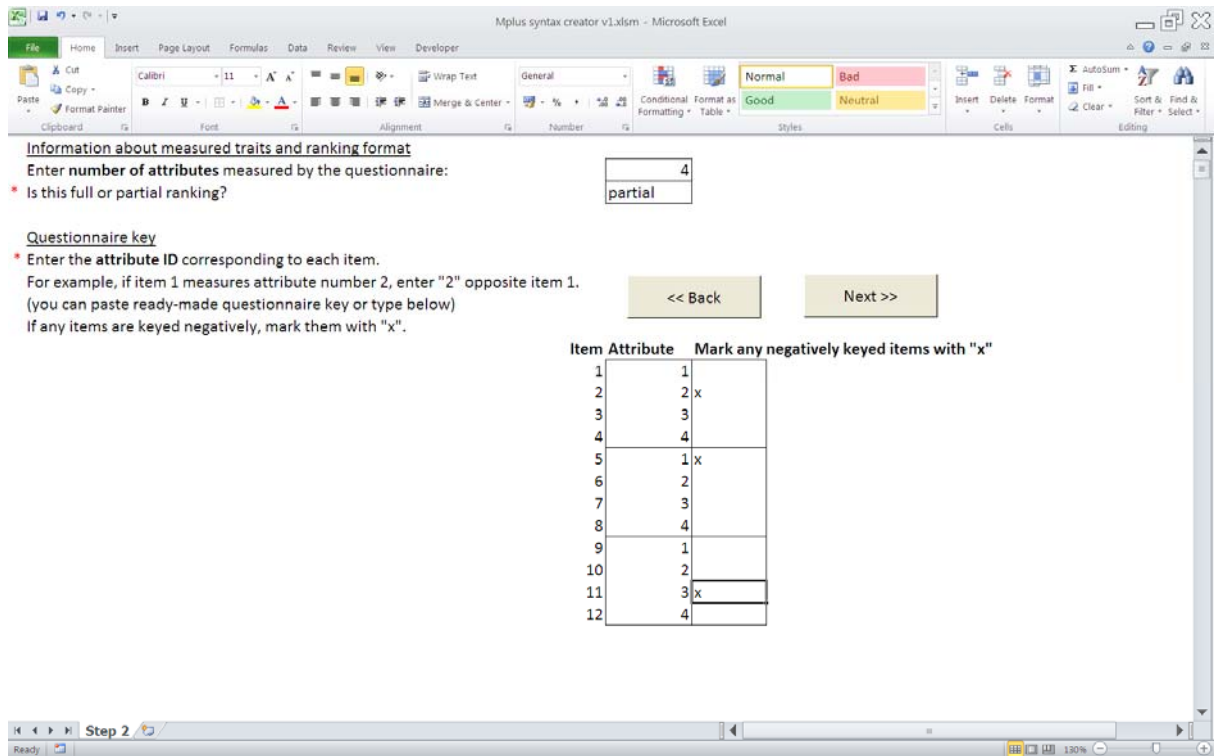
After pressing the 'Next' button, we can enter expected correlations between the measured attributes. This step is optional. In the below screen shot, we enter expected correlations. These will be used as starting values in Mplus estimation.
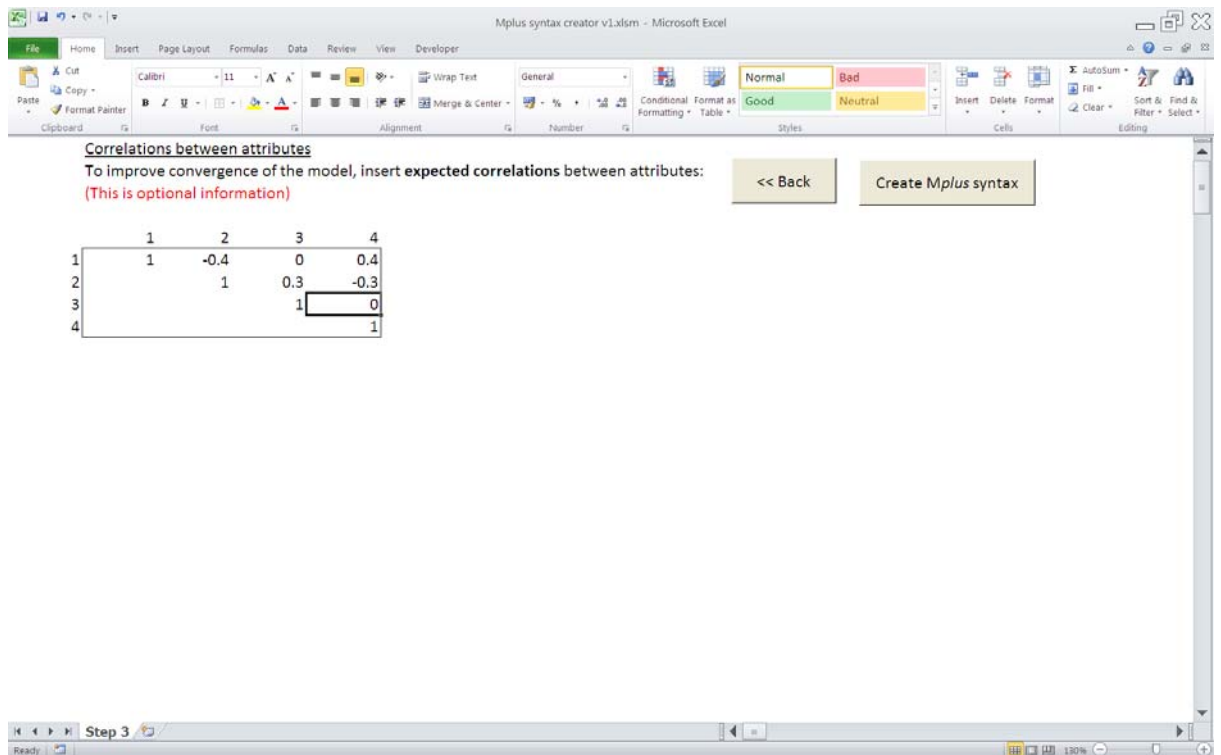
After pressing the 'Next' button, we can view the Mplus syntax, and copy and paste it to an Mplus input file ready for execution.
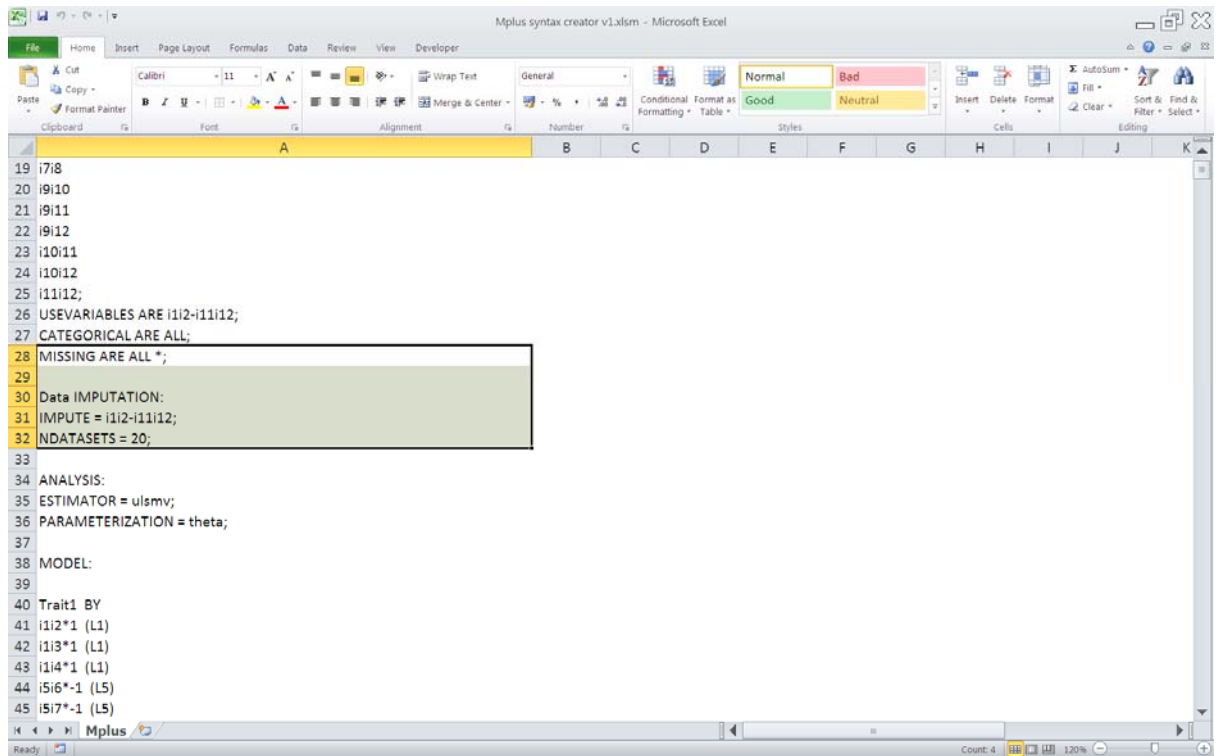
It can be seen that in this case of partial ranking, the multiple imputation facility provided by Mplus is used. By default, the symbol used for recognising missing outcomes in the syntax is "*".

> **MISSING ARE ALL *;**

If the missing outcomes were coded differently, the asterisk needs to be replaced with the appropriate code. For instance, if 99 is the missing indicator, use the following syntax

> **MISSING ARE ALL (99);**

EXAMPLE 3. BUILDING SYNTAX FOR BLOCKS OF 2 ITEMS (ITEM-PAIRS) WHEN THE NUMBER OF MEASURED ATTRIBUTES IS GREATER THAN 2

This is an example of creating Mplus syntax for testing forced-choice data arising from item-pairs, when more than 2 attributes are measured. In this case, no item uniqueness can be identified. It is convenient to assume that uniquenesses for bot item utilities involved in a comparison equal 0.5, so that uniqueness of each pair can be set to 1.

In this simple example, 3 attributes are measured by 6 item-pairs. The assignment of items to measured attributes is as follows:

Item	Attribute	Keying
1	1	+
2	2	+
3	3	+
4	1	+
5	2	+
6	3	+
7	1	+
8	2	–
9	3	+

10	1	-
11	2	+
12	3	-

The technical detail for this design can be found in Brown and Maydeu-Olivares (2012), Example 3.

First, we enter the basic detail of this design into the 'Step 1' form and press 'Next':

Next, we specify that the number of measured attributes is 3, and enter the questionnaire "key", as follows:
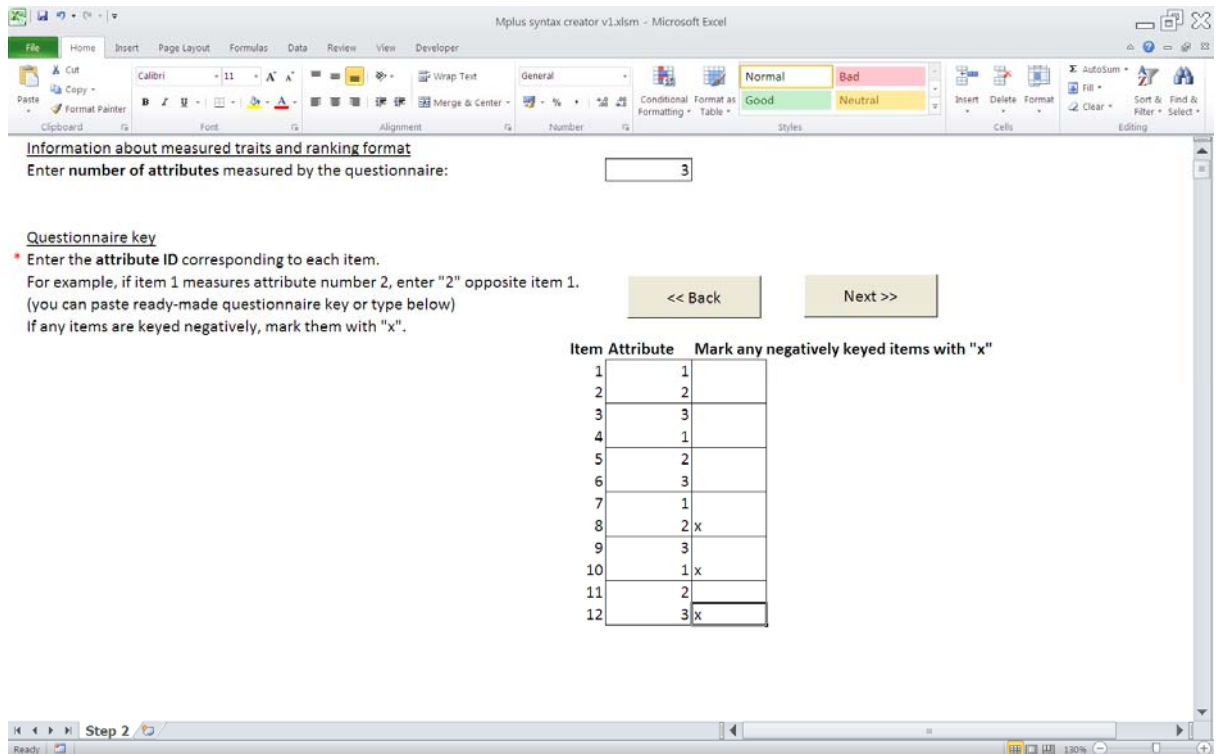
After pressing the 'Next' button, we can enter expected correlations between the measured attributes. This step is optional, and is not different from specifying these correlations in Example 1.

After pressing the 'Next' button, we can view the Mplus syntax, and copy and paste it to an Mplus input file ready for execution.

EXAMPLE 4. BUILDING SYNTAX FOR BLOCKS OF 2 ITEMS (ITEM-PAIRS) WHEN THE NUMBER OF MEASURED ATTRIBUTES IS EXACTLY 2

This is an example of creating Mplus syntax for testing forced-choice data arising from item-pairs, when exactly 2 attributes are measured. In this case, we have an exploratory two-factor analysis model with p binary variables.

In this simple example, 2 attributes are measured by 6 item-pairs. The assignment of items to measured attributes is as follows:

Item	Attribute	Keying
1	1	+
2	2	+
3	1	+
4	2	+
5	1	+
6	2	+
7	1	+
8	2	-
9	1	-
10	2	+
11	1	+
12	2	-

The technical detail for this design can be found in Brown and Maydeu-Olivares (2012), Example 4.

First, we enter the basic detail of this design into the 'Step 1' form and press 'Next':

Next, we specify that the number of measured attributes is 2, and enter the questionnaire "key", as follows:

After pressing the 'Next' button, we can enter expected correlations between the measured attributes. This step is optional.

After pressing the 'Next' button, we can view the Mplus syntax, and copy and paste it to an Mplus input file ready for execution.

R<small>EFERENCES</small>

Brown, A. & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement, 71*, 460-502. doi: 10.1177/0013164410375112

Brown, A. & Maydeu-Olivares, A. (2012). Fitting a Thurstonian IRT model to forced-choice data using Mplus. *Behavior Research Methods.*

Muthén, L.K. & Muthén, B.O. (1998-2010). *Mplus User's guide. Sixth edition*. Los Angeles, CA: Muthén & Muthén. Retrieved from www.statmodel.com