Dr. Roger Graves Fall 2005

UNIVERSITY OF VICTORIA
DEPARTMENT OF PSYCHOLOGY

PSYCHOLOGY 518 (F01) PSYCHOMETRIC METHODS

COURSE PROJECT: INSTRUCTIONS FOR THE PAPER

Diskette

Please include with your paper a diskette containing both your original (with all subjects and all items) and final SPSS data files (.SAV) and the final Winsteps control (.CON) and data (.DAT) files. There should be four files in total, filenames should unambiguously identify which file is which, e.g. RG51S42Q.SAV, where RG is your initials, 51S identifies the number of subjects, 42Q identifies the number of questions, and SAV identifies the filetype. If you needed to recode reverse scored items, all the files should have the properly recoded data.

This is a required part of the submission - the paper will not be graded without these files.

Paper

The paper should be prepared in APA format and should include the following sections.

Title page

Introduction

Include background on the construct, importance of the construct, previous methods of measurement with references, critique of previous methods, description of rationale and method you took to improve on previous methods that includes an explanation of how questions were chosen/constructed.

Method

Participants Describe your sample (population, N, any demographics you collected).

Materials Briefly describe the full questionnaire (with reference to the appendix that has the full questionnaire as given to the participants with all items and instructions).

Test Development **Briefly** (the details appear in Results) outline the criteria for the desired test psychometric properties (e.g., criteria for alpha, Rasch fit) that you set as the goal for your final test form and the strategy for deleting items that you planned to use to achieve these criteria. For the desired test properties, you might say something like: "The goal was to obtain a set of items for which Cronbach's alpha was at least .85 with no item exceeding fit criteria for the Rasch model (defined as both Infit MNSQ statistic larger than 1.3 and standardized Z statistic greater than 3.0)."

Results

Report the names of all software products used (e.g., SPSS version 11.0, Winsteps version 3.50) and identify which analyses each was used for. Include references to these programs in References.

Please organize this section with subheadings for each step. In this section you should document everything you did for data analysis and deletion of items/persons and why you did it.

Step 1 Report any recoding of reverse scored items and results for the initial analyses (with all persons and all items included) - this should include raw score statistics (mean, SD, max, min, Cronbach's alpha, SEm, and SEe), number of misfitting persons (use infit MNSQ > 1.3 AND Z > 3.0 plus report how many had Z>=4.0), number of misfitting items (for this step report how many items had infit MNSQ < .70 and how many had MNSQ >1.30, and of these, how many had Z <3.0 or >3.0

for comparison with the results after the optimizing process). Identify any persons or items that these results led you to select for deletion. Clearly explain the rationale for this selection (why the persons or items fail to meet your criteria). Also, for step 1, report a Spearman Rank correlation between your predicted order of difficulty and the observed order of difficulty of items and comment on any extreme discrepancies. Provide a table as an appendix that shows these two orders.

Step 2 Describe what you did for optimizing fit and report results. Report number of misfitting persons (use infit MNSQ > 1.3 AND Z > 3.0 plus report how many had Z>=4.0), number of misfitting items (for this step report how many items had infit MNSQ < .70 and how many had MNSQ >1.30 and of these, how many had Z >3.0 or <3.0 for comparison with the results before the optimizing process). Identify any persons or items that these results led you to select for deletion at the next step. Clearly explain the rationale for this selection (why the persons or items fail to meet your criteria).

Step 3, etc. Describe what you did re. deletion of persons and items at this step. (Note you should not delete more than 10% of your sample, not more than 5 for most of you.) Report results of analyses after deletion of any persons and/or items. This should only include misfit results unless this turns out to be the final step. Identify any persons or items that these results led you to select for deletion at the next step. Clearly explain the rationale for this selection (why the persons or items fail to meet your criteria). If you shortened the test, clearly describe the rationale you used for selection of items as well as the results at each step.

Final step Report results as for previous step plus also:

A Summary of why the results now meet your criteria.

Raw score statistics (mean, SD, max, min, SEm, and SEe),

Report of:

a) the range of item difficulty in logits from Table 13.1 (e.g., -1.5 to +2.0)

b) the range of category difficulty from Table 3.2, this is from the "Category Measure"column (e.g., -2.16 to +2.34). Use the results for Group 1.

c) the total range covered, which is the item range extended in each direction by the category range in that direction (e.g. -3.66 to +4.34).

The Rasch person and item reliabilities (use the 'real' values) and separations from Table 3.1.

Refer to appendices that contain Tables 3.1, 6.1, 10.1, 12.2, and 20.1 (both the score table and the figure part of this table). (Note, you can cut and paste. You may need to select the whole table and reduce the font size to make it fit in your pages.)

For step 1 and for the final step, the raw score statistics (mean, SD, max, min, and alpha) can generally be obtained from Winsteps Table 3.1. If, however, you had any extreme scores (subjects with the minimum or maximum possible score (that is, they used the lowest or highest category for all questions), then Winsteps deletes the data for these persons before computing the statistics in Table 3.1. You know when this happens if Table 3.1 says something like "Input 50 persons, Analyzed 48 persons". To get the correct raw score statistics in this case, you will have to use SPSS. (Create a new variable, e.g., "Total" that is the sum of the scores on all items by using "Transform, compute" on the SPSS menus (note you can use SUM (var00001 to var00050), then get statistics for this new variable.) Use Cronbach's alpha to calculate SEm and SEe. (Somewhere you should provide the

formulae used for these calculations and state that you are using alpha.)

Summarize the test development process at the end of Results. There should be a clear description of which items (by number) from the original questionnaire were ultimately retained (or deleted) to make the final questionnaire. Refer to appendices that contain the final person/item logit map, fit tables, score table and graph, etc. (Each appendix should have a statement in the text that refers to it.)

Discussion

Include a descriptive summary of the results, conclusions about how well the new questionnaire worked, discussion of deleted items (hypotheses for why these items came to be deleted, problems with them, etc.), suggestions for further research outlining how the questionnaire could be improved, validated, used, etc.

References

Include the following references along with your own.

Linacre, J. M., & Wright, B. D. (2001). Winsteps. Chicago: MESA Press.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institute for Educational Research. (Expanded edition, Chicago: The University of Chicago Press, 1980.)

Appendices

Include at least the following.

- Initial questionnaire.
- Table of predicted and obtained order of item difficulty. Ties are allowed for your predictions, but there should be at least 3 levels of difficulty.
- Winsteps tables for the final analysis (not for any intermediate steps) including:
  - Table 3.1, summary stats
  - Table 6.1, person misfit
  - Table 10.1, item misfit
  - Table 12.2 person and item map graphic
  - Table 20.1, correspondence between raw total score and Rasch (logit) person trait measure, include both the table text and the graphic under the table on two separate pages.

UNIVERSITY OF VICTORIA
DEPARTMENT OF PSYCHOLOGY

PSYCHOLOGY 518 (F01): PSYCHOMETRIC METHODS

DATA ANALYSIS LAB #1

Setup for entry of data

Open SPSS

START, PROGRAMS, SPSS

Set default format for new variables

EDIT, OPTIONS, DATA

set format for width of 1, decimal places of 0, close window

Set up for entry of data

Click on "Variable View" tab

Create a new variable SUBJECT, variable type string, width 3

Create a new variable BLANK, variable type numeric (width should automatically be 1)

Create N new variables, where N is the number of items in your questionnaire

Click on number of the highest item you have plus 2 (e.g. 62 if you have 60 items)

DATA, INSERT VARIABLE (this creates 60 new variables)

Click on "Measure" for the first variable and change to Ordinal, right click, copy

Highlight all other variables' measure column by dragging, right click, paste (this changes all variables to ordinal)

If you wish, you can rename all variables (e.g., Q1, Q2, ...), click, type in new name

Create two new variables, PREDRANK and OBSRANK, width 3, decimals 0, type ordinal

The above has been done for you. Open the data file.

FILE, OPEN, DATA, browse to O:\PSYCHOLOGY\PSYC518\518DataTemplate.SAV

Edit this file to suit your questionnaire by deleting the unneeded question variables.

Save the file FILE, SAVE AS (e.g., T:\YourNetlinkID\RG00S60Q) This creates an SPSS .SAV data file.)

Entry of data

When you have your results, you can then enter data into this file.

Click on "Data View" tab

Enter a value for SUBJECT for Subject 1 (e.g. S01) and then all the responses for that subject. One approach is to use the numeric keypad with the right hand and the cursor keys with the left hand while a friend reads off the numbers. Enter data for all subjects.

Enter your predicted rank order for the questions. The question you think will be the least endorsed will have rank 1, the next least rank 2, etc. You could use "1" for a set of e.g. 10 questions that you think are equally least likely to be endorsed, "2" for the next set, etc.

Save the file FILE, SAVE AS (e.g., RG53S45Q.SAV). **Please use a label that indicates the number of subjects and number of items.**

Check for any data entry errors. (Doing a Descriptives for all item variables will show if the range of values is correct at least.)

Recode any reverse coded items

Example of a method to do this.

Open the raw data file FILE, OPEN, DATA, browse to O:\PSYCHOLOGY\PSYC518\LIKERT_RAWDATA.SAV

This is a sample data file of 30 Likert scale items, responses coded 1-4. Questions 3, 7, 9 in this file are reverse coded and must be recoded.

TRANSFORM, RECODE, INTO SAME VARIABLES

Click on q3, no hold down the CTRL key and click on q7 and q9, click on the arrow to move those three variables to the right side box.

Click on "Old and New Values" button.

Type 1 in the Old Value box, 4 in the new value box, click on Add,

repeat for 2 to3, 3 to 2, 4 to 1.

Click on Continue

Click on OK. (This will change the values for these three items.)

Save file with a new name,

FILE, SAVE AS (e.g., T:\YourNetlinkID\LIKERT_RECODED.SAV)

Save the data in "fixed ASCII" format (with a .DAT filetype) to use in Winsteps.

Example: Open the raw data file

FILE, OPEN, DATA, browse to O:\PSYCHOLOGY\PSYC518\LIKERT_EXAMPLE.SAV

Now save it in the new format.

FILE, SAVE AS

In the "Save as type" field select "Fixed ASCII (*.dat)"

Provide a file name such as T:\YourNetlinkID\RG29S30Q.DAT

Please use a label that indicates the number of subjects and number of items.

Check that the format is correct by opening this .DAT file in a text editor (e.g., WordPad). It should look like the lines below, with no extra spaces between the response numbers.

```
S1 22223233232333232232221113321
S2 41333342221433331131311111121
```

Create a Winsteps control file.

Open WordPad (This is better than Word for this type of file.)

Open the sample control file

FILE, OPEN, DATA, browse to
O:\PSYCHOLOGY\PSYC518\RATING_SCALE_EXAMPLE.CON

Edit the sample control file to make it suit your data.

TITLE= line should have an informative title for your questionnaire,

NI= entry should have the correct number of items,

CODES= should have the correct item response codes,

DATA= should have the correct path and file name for your ASCII data file,

The lines after the CFILE* line should have the correct response categories for your questionnaire,

Item labels should be present for each item, and be in the correct order. Keep the labels short so they will fit on the output tables.

Save the control file with a new filename. For example, RG529S30Q.CON. Remember to save this file in ASCII, text, format.

UNIVERSITY OF VICTORIA
DEPARTMENT OF PSYCHOLOGY

PSYCHOLOGY 518 (F01): PSYCHOMETRIC METHODS

DATA ANALYSIS LAB #2

SPSS

Check on data entry

Start SPSS

Open your data file. This should be available in your personal folder on the T drive.

ANALYZE, DESCRIPTIVE STATISTICS, DESCRIPTIVES

Highlight all the item variables, click on the arrow to move them to the right. OK

Look at the N (should be the same for all) and the range of values (should be within the range you used).

If there are any problems, correct them now. Save the corrected .SAV file and create a new .DAT file to replace the faulty one.

Make backups of your original data files

Make copies of your original (with all persons and questions) .SAV, .CON, and .DAT files, e.g., by adding .BAK to the original names. You will need to submit the original files with your paper, and you may need to go back to them to start over if your analysis runs into difficulty.

SPSS

Check for homogeneity - obtain Cronbach's alpha for the original dataset

ANALYZE, SCALE, RELIABILITY ANALYSIS

Highlight all the item variables, click on the arrow to move them to the right box.

Choose Model = Alpha, Click on OK. (Output shows alpha.)

Winsteps

Load your control file

Start Winsteps

Open your control file. This should be available in your personal folder on the T drive. To open the control file:

Press Enter (a file window will open)

Browse to the desired control file

Click Open

Look at the control file

FILE, EDIT CONTROL FILE ... (This opens the control file in Wordpad.)

Confirm that the DATA= statement provides the correct path to your data on the T drive.

Edit and save if necessary.

Run a Rasch analysis

Press the Enter key twice. (The analysis should now run.)

Check for fit to the Rasch model

OUTPUT TABLES, Choose Table 10. Look at 10.1, Question statistics

First, look at the PTMEAS. CORR. column. This shows the correlation (across persons) between each item score and the person's total score. For a unidimensional test, all these correlations should be positive. Any negative correlation suggests a major problem with the item (e.g., failure to recode a reverse coded item, problem with wording, etc.). A low positive correlation (e.g., less than .20) suggests a possible problem with that item (but item could be OK if item difficulty is at either high or low extreme). If there are more than 3 items with a negative correlation, consult the instructor.

For any item with a negative item/total correlation, look at the wording of the item. Should this item be reverse coded? If so, go back to SPSS, recode any such items, and save the SPSS file. You can use the original filename since you are correcting an error. Then create a new Winsteps .DAT data file and save that. Then obtain Cronbach's alpha.

If the item wording does not indicate a reverse coding problem. The negative correlation items should be deleted. We will do that later.

OUTPUT TABLES, Choose Table 6. Look at 6.1, Person statistics.

Look at the INFIT column, check if any person has Infit MNSQ > 1.3 AND Z>3.0. (Z >=3.1 provides a Bonferroni corrected $p = .05$, 1-tail, alpha level based on 50 subjects.) This indicates misfit, e.g., the person was responding randomly or was influenced by additional factors. Make a note of how many persons, and which persons, misfit by this criterion. If there are more than 10%, please consult with the instructor.

OUTPUT TABLES, Choose Table 10. Look at 10.1, Question statistics

Look at the INFIT column, check if any item has infit MNSQ outside the range .70-1.30.

Write down the item numbers, the infit MNSQ, and Z values for all items outside this range.

EVERYTHING UP TO THIS POINT IS PART OF "STEP 1" FOR YOUR PAPER. At this point all persons and items have been retained, data entry appears to be correct, you have a value for alpha, and you have noted which persons and items fail the misfit criteria.

Optimize Rasch model fit.

The Winsteps analysis you just did used a "Rating Scale Model" in which the distance (in logit units) between response categories is the same for all items. This may not be realistic for all your items. You do not have enough subjects to use the "Partial Credit Model", which would allow Winsteps to establish a unique distance between response categories for each item individually. One could simply delete any items that misfit because their response category distances are different from the bulk of the items. Indeed, if there are only a few misfitting items and you have more than you need, this would be

the simplest approach. A compromise approach is to allow Winsteps to estimate the distances for 2 or 3 sets of items, more than 1 set for Rating Scale, but less than 40 or 50 sets for Partial Credit.

If there are one or more items with Infit MNSQ > 1.3, form a new "group" consisting of all these items. Do the same for any items with Infit MNSQ < .7.

Edit your control file.

FILE, EDIT CONTROL FILE ... (This opens the control file in Wordpad.)

Put a semicolon in front of the existing GROUPS=" " statement. This tells Winsteps to ignore it.

Add a new GROUPS=" " line to look like "GROUPS=11121111311112113", with the same number of digits as your number of items. "1" corresponds to items in the main group, "2" corresponds to the high Infit group, and "3" for items in the low Infit group. (In this example, items 1, 2 & 3 are in Group 1, item 4 in Group 2, item 9 in Group 3, etc.)

After you have changed the GROUPS= line, SAVE the control file and re-run the analysis.

FILE, EXIT THEN RESTART...

Before opening any tables, close all the old tables to avoid confusion (be sure you wrote down the info!)

Re-check the question statistics.

OUTPUT TABLES, Choose Table 10. Look at 10.1, Question statistics

Look at the INFIT column.

Items placed in the new groups should now have improved fit statistics. If this process did improve an item's infit MNSQ value (moved it closer to 1), leave the item in the new group. If infit worsened, put the item back into Group 1. If there are new items that have infit MNSQ outside the .70-1.30 range, add them to the new groups and re-run the analysis until no further improvement can be made. You are trying to have no items with infit MNSQ outside this range while using only 3 groups of items. When no further improvement can be made, note which items still fail the fit criteria.

Re-check the person statistics.

OUTPUT TABLES, Choose Table 6. Look at 6.1, Person statistics.

Make a note of which persons fail the fit criteria (Infit MNSQ > 1.3 AND Z > 3.0).

THIS COMPLETES STEP 2 FOR YOUR PAPER. You still have retained all persons and items but have optimized fit by forming three groups. You can report how many persons and items fail the fit criteria at this stage.

Delete person outliers

OUTPUT TABLES, Choose Table 6. Look at 6.1, Person statistics.

Look at the INFIT column, check if any person has Infit MNSQ >= 2.0 AND Z>3.0. We will consider anyone with a MNSQ this large to be an extreme misfit, e.g., the person was responding randomly or using some unique strategy. Make a note of the "Entry Number" of such persons.

Any such persons will be deleted before continuing so that their data do not bias the results, but do not delete more than 10% of your sample.

Edit the control file.

FILE, EDIT CONTROL FILE ... (This opens the control file in Wordpad.)

Add two new lines. These could be placed just above the DATA= line.

```
IDELQU=Y    ; This tells Winsteps to ask if you want to delete any items.
PDELQU=Y    ; This tells Winsteps to ask if you want to delete any items.
```

Now re-run Winsteps.

FILE, EXIT THEN RESTART...

When Winsteps asks if you want to delete any questions, respond with "N" to both queries.

When Winsteps asks if you want to delete any persons, respond "N" to the first query, "Y" to the second, then respond with the person entry numbers, end by entering zero. This is a temporary deletion, no change is made to the data file.

Before opening any tables, close all the old tables to avoid confusion.

Re-check the question statistics.

OUTPUT TABLES, Choose Table 10. Look at 10.1, Question statistics

Look at the INFIT column.

Confirm that no new items not in group 2 are above infit MNSQ 1.3 and no new items not in group 3 are below infit MNSQ .70. If there are, put them in group 2 or 3.

Re-check the person statistics.

OUTPUT TABLES, Choose Table 6. Look at 6.1, Person statistics.

Look at the INFIT column.

Confirm that no new person have infit MNSQ >=2.0. If so, re-run Winsteps, delete them and recheck. (Note you will have to enter ALL the person numbers to delete each time.)

Delete misfitting questions [Deleting questions with negative PTMEAS CORR values should be included.]

OUTPUT TABLES, Choose Table 10. Look at 10.1, Question statistics

Look at the INFIT column, check if any item has Infit MNSQ greater than 1.30 AND also has a Z>3.0. (Z >= 3.1 provides a Bonferroni corrected $p = .05$, 1-tail, alpha level based on 50 questions.) Note that we are not excluding questions with Infit MNSQ <.70. The presence of a few such questions that fit the model too well (less error variance than predicted) does not cause problems with parameter estimation and such questions are frequently retained.

Make a note of the "Entry Number" of which items fail this fit criterion.

Now re-run Winsteps.

FILE, EXIT THEN RESTART...

When Winsteps asks if you want to delete any questions, respond with "N" to the first query "Y" to the second. Then respond with the question entry numbers for all the misfitting questions, end by entering zero.

When Winsteps asks if you want to delete any persons, respond "N" to the first query, "Y" to the second, then respond with the person entry numbers, end by entering zero. This is a temporary deletion, no change is made to the data file.

Before opening any tables, close all the old tables to avoid confusion.

Re-check the person and question tables. Make a note of the entry numbers of any persons that may now exceed the outlier criterion (infit MNSQ >=2.0 AND Z > 3.0) and any questions that exceed the fit criterion (infit MNSQ > 1.3 AND Z > 3.0).

THIS COMPLETES STEP 3. You have temporarily deleted outlier persons and misfitting questions. You can report how many persons and items have been deleted and how many of the remaining ones fail the fit criteria at this stage.

Repeat the process involved in step 3 if there are remaining outlier persons or misfitting questions. Each repetition would constitute a new step for the paper.

THIS COMPLETES THE DATA DELETION STEPS. At this stage you have a set of persons with outliers deleted but perhaps with some still failing the lesser misfit criteria and you have a set of questions all of which satisfy the fit criteria.

Create new data and control files

SPSS

Delete the persons and items that were temporarily deleted in Winsteps.

Save the SPSS data file (.SAV) **using a new filename.** (e.g. RG50S45Q.SAV)

Save a Winsteps format data file (.DAT) **using a new filename.**

Obtain Cronbach's alpha for the new dataset.

Winsteps

Edit the control file.

Change the NI= line to correspond to the new number of items.

Change the GROUPS= line to correspond to the new number of items by deleting the digits that referred to the deleted items.

Delete the question labels for the deleted questions.

Change the DATA= line to correspond to the new .DAT filename.

Save the control file **using a new filename** (e.g., RG45S52Q.CON where the S and Q numbers match those on the .DAT file).

Now close Winsteps, restart it, and open the new control file.

Run Winsteps - this time you say "N" to all those data deletion questions.

Re-check the Person and Question tables to make sure they are the same as the last time.

Save these tables to your temporary folder. You will need them for the paper.

Obtain Winsteps variable map (Table 1.2) and Test Characteristic Curve and score table (Table 20.1)

Save these tables to your temporary folder. You will need them for the paper.

Obtain Winsteps' Summary Statistics (Table 3.1) and save this table to your temporary folder.

THIS COMPLETES THE FINAL STEP. You now have a set of questions that satisfies the fit criteria we used for fit to the Rasch model. Some of the subjects may be misfitting, but you can report what percentage of the original sample does satisfy the fit criteria (the number not deleted and not misfitting by our criteria (infit MNSQ > 1.3 and Z >3.0) as one type of evidence for the percentage of the population for which the test would provide a valid measure of your construct.

You could now (and this is recommended), experiment with shortening and improving the test. You would do this by selectively deleting items. Deleting items with low item/test correlations should preserve (or increase) alpha (and person reliability as reported by Winsteps). You would also use the item map and try to keep items at the extremes of the endorseability range (delete items in the middle, especially those where there are many others at the same approximate endorseability value). You might also use the Winsteps Principal Components Analysis (PCA) (Table 23) and select items that load high on what Winsteps terms "Factor 1" (which is the 1st residual factor after the main Rasch factor has been extracted - SPSS would call this factor 2). Table 23.1, which is at the bottom of the PCA tables, shows the highest positive and negative item intercorrelations. Items should be independent for IRT, thus any correlated items are redundant - you could choose to delete one of each pair of redundant items, selecting for deletion the item of the pair that has low item/test correlation, or is less extreme in endorseability, or has higher loading on the PCA factor 1.