

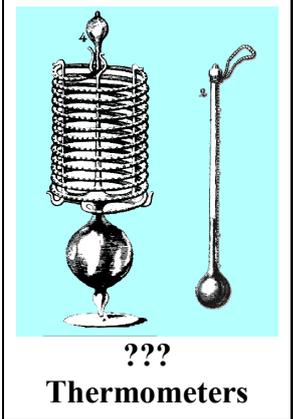
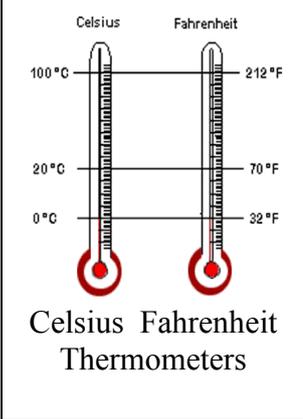
**1. Tutorial 4. Test Equating**  
**This week is a quick overview.**

- Test equating: separate and concurrent
- Prettifying output

If you don't know the meaning of a word, then please look at the "Glossary".  
<http://www.winsteps.com/winman/glossary.htm>

**2. A. Equating and Linking Methods**

**3.** We are familiar with Celsius and Fahrenheit thermometers. They use different numbering systems, but we can convert from one temperature-measurement scale to the other. The thermometers are "equated".  
 If we had a third type of thermometer we could equate it to our familiar thermometers by using "linking items", such as melting ice and boiling water. Using these items, we could calibrate the third thermometer to make it comparable to Celsius and Fahrenheit thermometers.



**4. For reference:**

<b>Linking Terminology: Raw Score and Rasch</b>		
<b>Term</b>	<b>Raw Score CTT meaning</b>	<b>Rasch meaning</b>
<b>Linking</b>	general term for making the results of different tests comparable	enabling the data to be analyzed together in one analysis (if desired) to construct one overall set of measures
<b>Equating</b>	correspondence of raw scores between tests	putting the measures in the same frame of reference
<b>Calibration</b>	putting the scores in the same frame of reference	constructing item measures in the internal frame of reference
<b>Projection</b>	scores on one test weakly predict scores on another test	(a height-weight situation)
<b>Moderation</b>	equivalences based on matching up sample statistics	(Fahrenheit-Celsius equating)
<b>Anchoring (fixing)</b>	-	measures obtained from one analysis (or construct theory) imposed on another to place it in the same frame of reference.
<b>Local origin</b>	zero score or sample mean	reference location from which to measure along the latent variable

from <http://www.rasch.org/rmt/rmt204b.htm>

5. *For reference:*
- Equating designs** for two tests:
- Baseline equating:** estimates from one analysis are anchor values in another analysis
- Parallel equating:** two analyses are performed separately, and then equated by a fit line.
- Common-item equating:** some (linking) items in the two tests are the same: *best method*.
- Common-person equating:** some (linking) persons who respond to the two tests are the same
- Linking tests:** a third test is constructed which contains items from the two tests. This is used as a common-item-equating link between the two tests.
- Concurrent (or one-step) equating:** both tests are analyzed together as one dataset.
- Virtual equating:** common-item equating based on items with similar (not identical) characteristics
- Vertical equating:** the two tests are intended to differ in difficulty, usually by at least a grade-level.
- Horizontal equating:** the two tests are intended to have the same difficulty
- Polytomous equating:** the two tests share rating scale structures
- Separate-estimation equating:** the two tests are analyzed separately, then the estimated measures are used for performing the equating.
- Random-equivalence equating:** two tests (or two examinee samples) are declared to produce randomly-equivalent measures because their items (or persons) have been allocated to the tests at random. So that the means and S.D.s of the two sets of item (or person) measures can be equivalenced.
- Alternate-forms equating:** two tests are declared to be equivalent, so that their measures are assumed to be in the same frame of reference.
- Anchored-form equating:** the tests included pre-calibrated items which force the two tests to report measures in the frame of reference of the item-anchor values.
- Common scale:** the measurement scale on which the measures from the two tests are expressed. This is usually the measurement scale of one of the tests.
- Equating constant:** when tests are equally discriminating, the amount to add to measures from one test to equate them with measures from the other test.
- Scaling constant or Equating slope:** when tests have measures with different discriminations (Celsius-Fahrenheit), the amount with which to multiply measures from one test to make them equally discriminating with those of the other test.
- Equating error:** loss of precision of the measures from one test when converted to become measures on the other test.
- Equipercetile equating:** a non-linear CTT technique based on equivalencing distributions. Not used in Rasch equating.
- Item drift:** change in the difficulty of an item from one test (or test administration) to another test (or test administration).

For more about equating methods, see Winsteps Help “Equating and Linking Tests”  
<http://www.winsteps.com/winman/equating.htm>

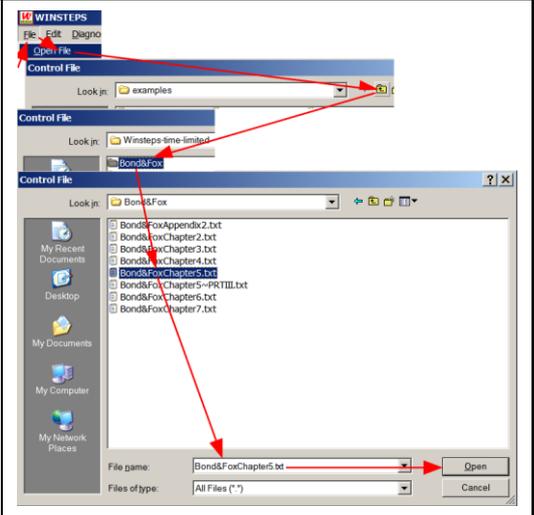
**6. B. Separate-estimation Common-person Test Equating**

7. Let's try equating two of the Bond & Fox datasets, as discussed in B&F Chapter 5. This will use "Common Person" equating, because the instruments were administered to the same people. The basic requirement is that the latent variable be effectively the same, "invariant", across the two instruments to be equated or linked. Then we can make one-to-one conversions between the measures on the two instruments.

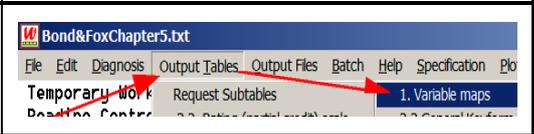
8. First we need the person ability measures from the "B&F chapter 5" data.  
Launch Winsteps



9. Click on "File"  
Click on "Open File"  
Click up one level to the "Winsteps" folder  
Click on "Bond&Fox"  
Double-Click on "Bond&FoxChapter5.txt"  
  
Report Output? Press Enter  
Extra Specifications? Press Enter  
  
The standard analysis is performed ....



10. Winsteps menu bar  
Click on "Output Tables"  
Click on "1. Variable Maps"



11. Table 1 displays in a Notepad window  
Scroll down to Table 1.1

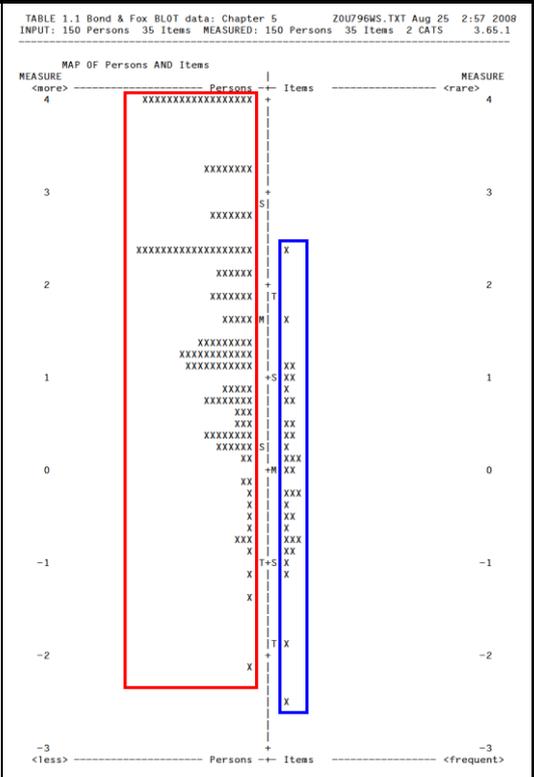
Red box: The summary map shows the ability measures for 150 persons

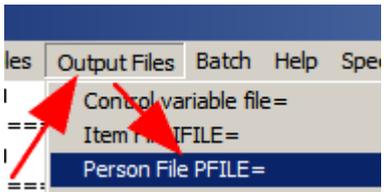
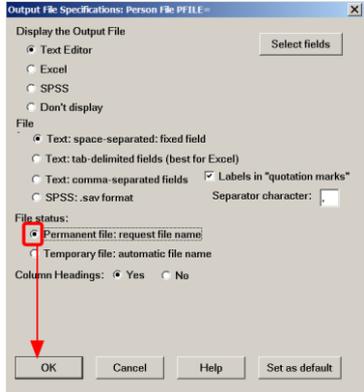
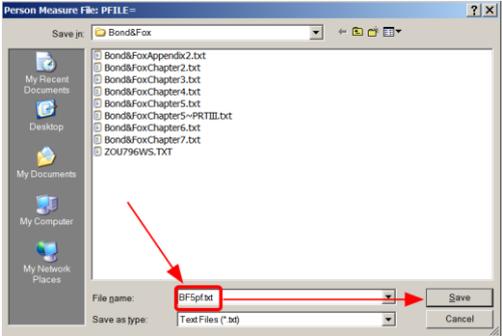
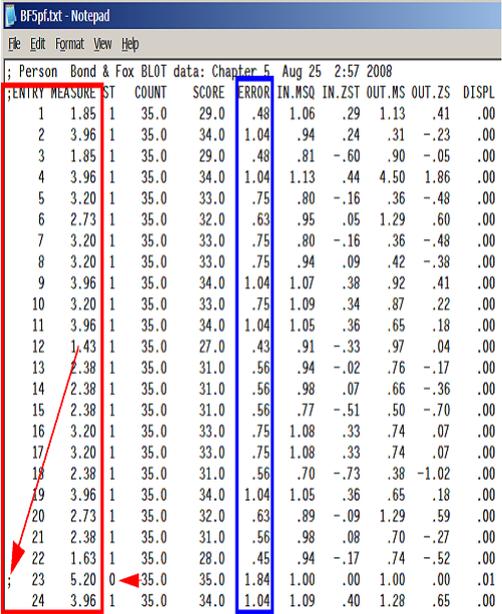
Blue box: and also the difficulties for 35 items.

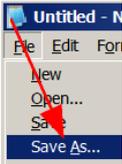
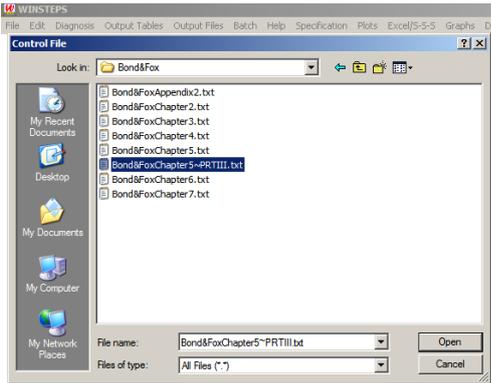
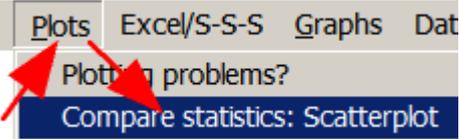
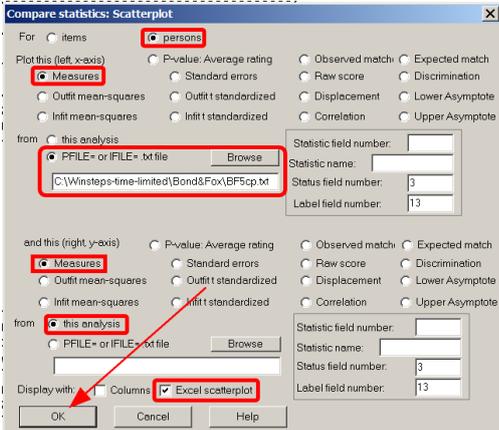
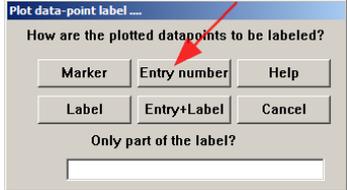
The 150 persons also responded to another test. So these are our "common persons".

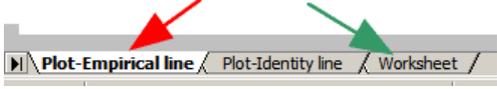
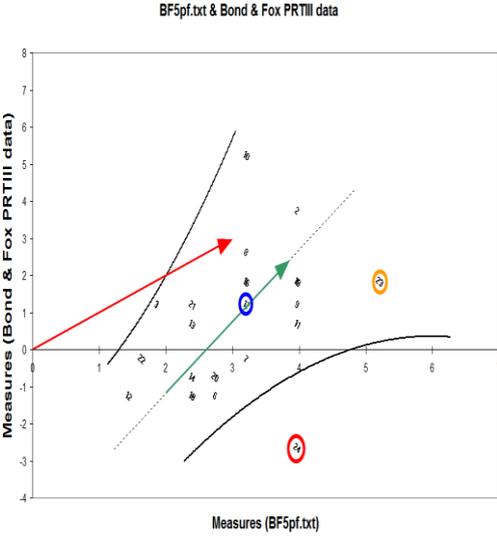
We need their measures for the equating process.

But let's pretend that only the first 24 persons are common persons, we can then check the effectiveness of the equating for the other 126 persons!

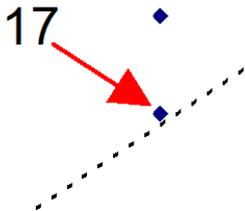
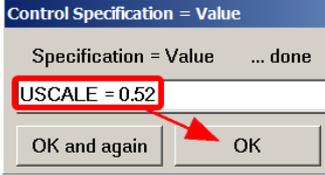
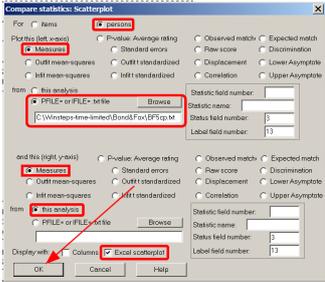
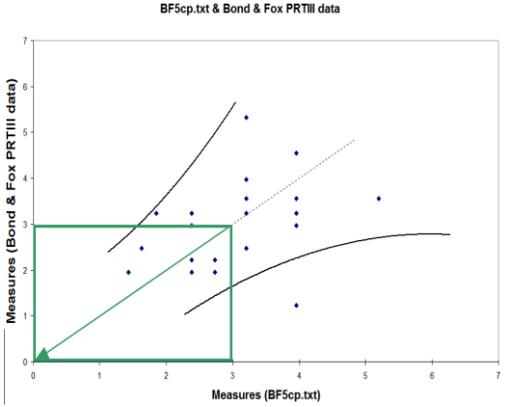
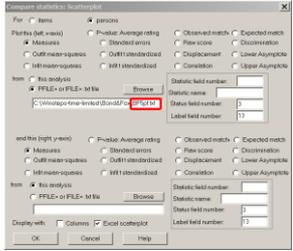


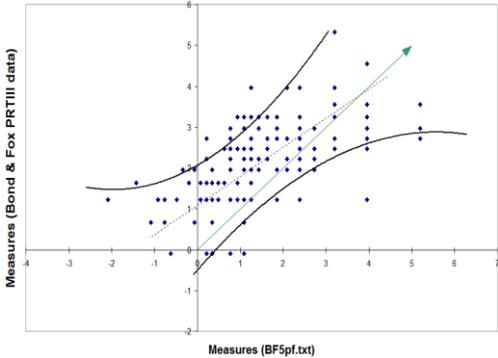
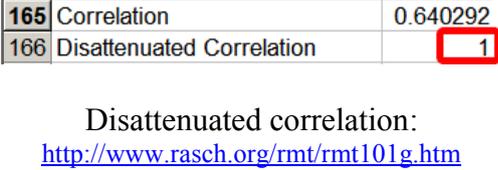
<p>12.</p>	<p>Now let's write out the person measures. Click on Output Files Click on Person File PFILE=</p>																																																																																																																																																																																																																																																																																				
<p>13.</p>	<p>Output File Specifications Click on Permanent file Click OK</p>																																																																																																																																																																																																																																																																																				
<p>14.</p>	<p>Enter a memorable file name. Mine is: BF5pf.txt Then Save</p>																																																																																																																																																																																																																																																																																				
<p>15.</p>	<p>The person file statistics are displayed in a Notepad window. This procedure would be exactly the same if we were doing a Common-item equating. A Rasch model does not know what is a “person” and what is an “item”, so treats them essentially the same. We do know which is which, so we use that insight to interpret the numbers.</p> <p><i>Red box:</i> the person entry numbers and their ability measures. These are what we need. The common-person entry numbers must be the same in the analysis of the test to be equated with this one.</p> <p><i>Blue box:</i> the standard errors. These will help us determine how stable the person measures are.</p> <p><i>Red arrow:</i> this indicates an extreme score: “;” (if shown) also status 0 or -1.. The measures corresponding to an extreme score are less secure for equating than the measures for non-extreme scores.</p>	 <table border="1"> <thead> <tr> <th>ENTRY</th> <th>MEASURE</th> <th>ST</th> <th>COUNT</th> <th>SCORE</th> <th>ERROR</th> <th>IN.MSQ</th> <th>IN.ZST</th> <th>OUT.MS</th> <th>OUT.ZS</th> <th>DISPL</th> </tr> </thead> <tbody> <tr><td>1</td><td>1.85</td><td>1</td><td>35.0</td><td>29.0</td><td>.48</td><td>1.06</td><td>.29</td><td>1.13</td><td>.41</td><td>.00</td></tr> <tr><td>2</td><td>3.96</td><td>1</td><td>35.0</td><td>34.0</td><td>1.04</td><td>.94</td><td>.24</td><td>.31</td><td>-.23</td><td>.00</td></tr> <tr><td>3</td><td>1.85</td><td>1</td><td>35.0</td><td>29.0</td><td>.48</td><td>.81</td><td>-.60</td><td>.90</td><td>-.05</td><td>.00</td></tr> <tr><td>4</td><td>3.96</td><td>1</td><td>35.0</td><td>34.0</td><td>1.04</td><td>1.13</td><td>.44</td><td>4.50</td><td>1.86</td><td>.00</td></tr> <tr><td>5</td><td>3.20</td><td>1</td><td>35.0</td><td>33.0</td><td>.75</td><td>.80</td><td>-.16</td><td>.36</td><td>-.48</td><td>.00</td></tr> <tr><td>6</td><td>2.73</td><td>1</td><td>35.0</td><td>32.0</td><td>.63</td><td>.95</td><td>.05</td><td>1.29</td><td>.60</td><td>.00</td></tr> <tr><td>7</td><td>3.20</td><td>1</td><td>35.0</td><td>33.0</td><td>.75</td><td>.80</td><td>-.16</td><td>.36</td><td>-.48</td><td>.00</td></tr> <tr><td>8</td><td>3.20</td><td>1</td><td>35.0</td><td>33.0</td><td>.75</td><td>.94</td><td>.09</td><td>.42</td><td>-.38</td><td>.00</td></tr> <tr><td>9</td><td>3.96</td><td>1</td><td>35.0</td><td>34.0</td><td>1.04</td><td>1.07</td><td>.38</td><td>.92</td><td>.41</td><td>.00</td></tr> <tr><td>10</td><td>3.20</td><td>1</td><td>35.0</td><td>33.0</td><td>.75</td><td>1.09</td><td>.34</td><td>.87</td><td>.22</td><td>.00</td></tr> <tr><td>11</td><td>3.96</td><td>1</td><td>35.0</td><td>34.0</td><td>1.04</td><td>1.05</td><td>.36</td><td>.65</td><td>.18</td><td>.00</td></tr> <tr><td>12</td><td>1.43</td><td>1</td><td>35.0</td><td>27.0</td><td>.43</td><td>.91</td><td>-.33</td><td>.97</td><td>.04</td><td>.00</td></tr> <tr><td>13</td><td>2.38</td><td>1</td><td>35.0</td><td>31.0</td><td>.56</td><td>.94</td><td>-.02</td><td>.76</td><td>-.17</td><td>.00</td></tr> <tr><td>14</td><td>2.38</td><td>1</td><td>35.0</td><td>31.0</td><td>.56</td><td>.98</td><td>.07</td><td>.66</td><td>-.36</td><td>.00</td></tr> <tr><td>15</td><td>2.38</td><td>1</td><td>35.0</td><td>31.0</td><td>.56</td><td>.77</td><td>-.51</td><td>.50</td><td>-.70</td><td>.00</td></tr> <tr><td>16</td><td>3.20</td><td>1</td><td>35.0</td><td>33.0</td><td>.75</td><td>1.08</td><td>.33</td><td>.74</td><td>.07</td><td>.00</td></tr> <tr><td>17</td><td>3.20</td><td>1</td><td>35.0</td><td>33.0</td><td>.75</td><td>1.08</td><td>.33</td><td>.74</td><td>.07</td><td>.00</td></tr> <tr><td>18</td><td>2.38</td><td>1</td><td>35.0</td><td>31.0</td><td>.56</td><td>.70</td><td>-.73</td><td>.38</td><td>-1.02</td><td>.00</td></tr> <tr><td>19</td><td>3.96</td><td>1</td><td>35.0</td><td>34.0</td><td>1.04</td><td>1.05</td><td>.36</td><td>.65</td><td>.18</td><td>.00</td></tr> <tr><td>20</td><td>2.73</td><td>1</td><td>35.0</td><td>32.0</td><td>.63</td><td>.89</td><td>-.09</td><td>1.29</td><td>.59</td><td>.00</td></tr> <tr><td>21</td><td>2.38</td><td>1</td><td>35.0</td><td>31.0</td><td>.56</td><td>.98</td><td>.08</td><td>.70</td><td>-.27</td><td>.00</td></tr> <tr><td>22</td><td>1.63</td><td>1</td><td>35.0</td><td>28.0</td><td>.45</td><td>.94</td><td>-.17</td><td>.74</td><td>-.52</td><td>.00</td></tr> <tr><td>23</td><td>5.20</td><td>0</td><td>35.0</td><td>35.0</td><td>1.84</td><td>1.00</td><td>.00</td><td>1.00</td><td>.00</td><td>.01</td></tr> <tr><td>24</td><td>3.96</td><td>1</td><td>35.0</td><td>34.0</td><td>1.04</td><td>1.09</td><td>.40</td><td>1.28</td><td>.65</td><td>.00</td></tr> </tbody> </table>	ENTRY	MEASURE	ST	COUNT	SCORE	ERROR	IN.MSQ	IN.ZST	OUT.MS	OUT.ZS	DISPL	1	1.85	1	35.0	29.0	.48	1.06	.29	1.13	.41	.00	2	3.96	1	35.0	34.0	1.04	.94	.24	.31	-.23	.00	3	1.85	1	35.0	29.0	.48	.81	-.60	.90	-.05	.00	4	3.96	1	35.0	34.0	1.04	1.13	.44	4.50	1.86	.00	5	3.20	1	35.0	33.0	.75	.80	-.16	.36	-.48	.00	6	2.73	1	35.0	32.0	.63	.95	.05	1.29	.60	.00	7	3.20	1	35.0	33.0	.75	.80	-.16	.36	-.48	.00	8	3.20	1	35.0	33.0	.75	.94	.09	.42	-.38	.00	9	3.96	1	35.0	34.0	1.04	1.07	.38	.92	.41	.00	10	3.20	1	35.0	33.0	.75	1.09	.34	.87	.22	.00	11	3.96	1	35.0	34.0	1.04	1.05	.36	.65	.18	.00	12	1.43	1	35.0	27.0	.43	.91	-.33	.97	.04	.00	13	2.38	1	35.0	31.0	.56	.94	-.02	.76	-.17	.00	14	2.38	1	35.0	31.0	.56	.98	.07	.66	-.36	.00	15	2.38	1	35.0	31.0	.56	.77	-.51	.50	-.70	.00	16	3.20	1	35.0	33.0	.75	1.08	.33	.74	.07	.00	17	3.20	1	35.0	33.0	.75	1.08	.33	.74	.07	.00	18	2.38	1	35.0	31.0	.56	.70	-.73	.38	-1.02	.00	19	3.96	1	35.0	34.0	1.04	1.05	.36	.65	.18	.00	20	2.73	1	35.0	32.0	.63	.89	-.09	1.29	.59	.00	21	2.38	1	35.0	31.0	.56	.98	.08	.70	-.27	.00	22	1.63	1	35.0	28.0	.45	.94	-.17	.74	-.52	.00	23	5.20	0	35.0	35.0	1.84	1.00	.00	1.00	.00	.01	24	3.96	1	35.0	34.0	1.04	1.09	.40	1.28	.65	.00
ENTRY	MEASURE	ST	COUNT	SCORE	ERROR	IN.MSQ	IN.ZST	OUT.MS	OUT.ZS	DISPL																																																																																																																																																																																																																																																																											
1	1.85	1	35.0	29.0	.48	1.06	.29	1.13	.41	.00																																																																																																																																																																																																																																																																											
2	3.96	1	35.0	34.0	1.04	.94	.24	.31	-.23	.00																																																																																																																																																																																																																																																																											
3	1.85	1	35.0	29.0	.48	.81	-.60	.90	-.05	.00																																																																																																																																																																																																																																																																											
4	3.96	1	35.0	34.0	1.04	1.13	.44	4.50	1.86	.00																																																																																																																																																																																																																																																																											
5	3.20	1	35.0	33.0	.75	.80	-.16	.36	-.48	.00																																																																																																																																																																																																																																																																											
6	2.73	1	35.0	32.0	.63	.95	.05	1.29	.60	.00																																																																																																																																																																																																																																																																											
7	3.20	1	35.0	33.0	.75	.80	-.16	.36	-.48	.00																																																																																																																																																																																																																																																																											
8	3.20	1	35.0	33.0	.75	.94	.09	.42	-.38	.00																																																																																																																																																																																																																																																																											
9	3.96	1	35.0	34.0	1.04	1.07	.38	.92	.41	.00																																																																																																																																																																																																																																																																											
10	3.20	1	35.0	33.0	.75	1.09	.34	.87	.22	.00																																																																																																																																																																																																																																																																											
11	3.96	1	35.0	34.0	1.04	1.05	.36	.65	.18	.00																																																																																																																																																																																																																																																																											
12	1.43	1	35.0	27.0	.43	.91	-.33	.97	.04	.00																																																																																																																																																																																																																																																																											
13	2.38	1	35.0	31.0	.56	.94	-.02	.76	-.17	.00																																																																																																																																																																																																																																																																											
14	2.38	1	35.0	31.0	.56	.98	.07	.66	-.36	.00																																																																																																																																																																																																																																																																											
15	2.38	1	35.0	31.0	.56	.77	-.51	.50	-.70	.00																																																																																																																																																																																																																																																																											
16	3.20	1	35.0	33.0	.75	1.08	.33	.74	.07	.00																																																																																																																																																																																																																																																																											
17	3.20	1	35.0	33.0	.75	1.08	.33	.74	.07	.00																																																																																																																																																																																																																																																																											
18	2.38	1	35.0	31.0	.56	.70	-.73	.38	-1.02	.00																																																																																																																																																																																																																																																																											
19	3.96	1	35.0	34.0	1.04	1.05	.36	.65	.18	.00																																																																																																																																																																																																																																																																											
20	2.73	1	35.0	32.0	.63	.89	-.09	1.29	.59	.00																																																																																																																																																																																																																																																																											
21	2.38	1	35.0	31.0	.56	.98	.08	.70	-.27	.00																																																																																																																																																																																																																																																																											
22	1.63	1	35.0	28.0	.45	.94	-.17	.74	-.52	.00																																																																																																																																																																																																																																																																											
23	5.20	0	35.0	35.0	1.84	1.00	.00	1.00	.00	.01																																																																																																																																																																																																																																																																											
24	3.96	1	35.0	34.0	1.04	1.09	.40	1.28	.65	.00																																																																																																																																																																																																																																																																											

<p>16.</p>	<p>Only persons 1-24 are common.          Notepad Window          Delete everyone after person 24 in this file.          Click on “File”          Click on “Save As” the Notepad file          Type in: <i>BF5cp.txt</i> (cp = common persons)</p>	
<p>17.</p>	<p>Now let’s analyze the dataset we want to equate with this one.          Winsteps menu bar          Click on File          Click on Start another Winsteps</p>	
<p>18.</p>	<p>In the new Winsteps window          Click on File          Click on Open File          Click on ....PRTIII.txt          Click on Open          Report Output? Press Enter          Extra Specifications? Press Enter            The standard analysis is performed</p>	
<p>19.</p>	<p>Now let’s compare the person measures:          Winsteps Menu Bar          Click on Plots menu          Click on Compare Statistics</p>	
<p>20.</p>	<p>In the Compare Statistics control box,          Click on for the x-axis:          persons          Measures          PFILE          Browse and select BF5cp.txt            Click on for the y-axis:          Measures          this analysis            Check-mark: Excel scatterplot          Click on OK</p>	
<p>21.</p>	<p>Plot data-point label ....          Click on Entry number ... we want to easily identify the outliers</p>	

<p>22.</p>	<p>The Excel plot displays.</p> <p>We do want the Empirical trend-line, because we have no reason to suppose that the logit measures defined by the BLOT test (BF5) have the same substantive “length” as the logit measures defined by the PRTIII test. If you don’t follow this, read “Length of a Logit” <a href="http://www.rasch.org/rmt/rmt32b.htm">http://www.rasch.org/rmt/rmt32b.htm</a></p>	
<p>23.</p>	<p>And here’s the plot. If you have more than 24 points, then you forgot to Save BF5cp.txt after deleting all the lines after person 24.</p> <p><i>Red circle:</i> We can see that person 24 is an outlier. Not surprising because the curved control lines are set to <math>p &lt; .05</math>, i.e., only 1 point in 20 is expected to lie outside them, but maybe it would be better to omit this person from the equating set as being too variable.</p> <p>For successful common-person equating, we need the person to behave the same way on both instruments. If not, the person becomes a “new and different” person.</p> <p><i>Red arrow:</i> This is the identity line, through the origin, with equal increments on both axes. This would be the trend line if the tests produced the same person measures.</p> <p><i>Green arrow:</i> A best-fit line, a dashed line, is not parallel to the identity line. So the instruments have different discrimination and different average item difficulty, relative to the average person ability.</p> <p>From a measurement perspective, the conversion between the BLOT and the PRTIII is like that between Celsius and Fahrenheit.</p>	 <p><i>Blue circle:</i> Person 17 is on a “best fit” trend line.</p> <p><i>Orange circle:</i> Person 23 had an extreme (perfect) score on the BLOT, so this measure is somewhat arbitrary. This is also an outlying influential point, suggesting that it should be omitted from the equating set.</p>
<p>24.</p>	<p>If the best-fit line (<b>green arrow</b>) is parallel to the identity line (red arrow), then it is usually easiest to analyze all the data together. In each person record, the items on each instrument are placed in adjacent columns, so that they look like one long test. This is called “racking” the data.</p>	

<p><b>25.</b> Before we move, look closely at the plot. There seem to be two trends among the persons. One group (perhaps the boys) follows the red arrow. The other group (perhaps the girls) follows the blue arrow. If we follow an automatic linking-person selection rule (such as Frederic Lord's), we will eliminate the minority arrow (fewer persons, probably the red arrow), and will base our equating on the majority arrow (more persons, probably the blue arrow). But "majority" and "minority" may be an accident of the sampling of our common persons, so it would be better to identify which of the two arrows (or what proportion of each) better represents a meaningful commonality between the two tests .</p>											
<p><b>26.</b> Back to our trend line. Our job is to find the conversion factors. Fortunately EXCEL has done the hard work for us. It has plotted a best-fit trend line for us. This is shown by the dashed black line on your plot. Click on "Worksheet" below the plot</p>											
<p><b>27.</b> Scroll down to the bottom of the Worksheet Do you see the Empirical intercept and slope? They convert PRTIII y-axis measures into the BLOT x-axis measures or the reverse:  <math>BLOT(x\text{-axis}) * slope + y\text{-axis intercept} = Predicted\ PRTIII</math>  <math>PRTIII(y\text{-axis}) / slope + x\text{-axis intercept} = Predicted\ BLOT</math></p> <p>This conversion applies to both the item difficulties and the person measures. We can convert all measures from one analysis onto the measurement scale of the other analysis.</p>		<table border="1" data-bbox="975 943 1481 1081"> <tr> <td>36</td> <td>Empirical intercept with x-axis</td> <td>2.60375</td> </tr> <tr> <td>37</td> <td>Empirical intercept with y-axis</td> <td>-5.03994</td> </tr> <tr> <td>38</td> <td>Empirical slope</td> <td>1.935648</td> </tr> </table> <p>If we delete any points in the worksheet, then Excel will redraw the plots, and recompute the slopes and intercepts.</p>	36	Empirical intercept with x-axis	2.60375	37	Empirical intercept with y-axis	-5.03994	38	Empirical slope	1.935648
36	Empirical intercept with x-axis	2.60375									
37	Empirical intercept with y-axis	-5.03994									
38	Empirical slope	1.935648									
<p><b>28.</b> It's always wise to check this type of result. Notice that person 17 (blue circle) is almost exactly on the trend line. In BF5pf.txt - the BLOT person 17 measure is 3.20 logits. That should be x-axis position of person 17. Yes.</p>											
<p><b>29.</b> For the actual y-axis position: PRTIII analysis (the one we are doing) Click on the Winsteps menu bar Click on Output Tables Click on 18. PERSON: entry</p>											
<p><b>30.</b> Scroll down to person 17 The PRTIII person measure is 1.23 logits. This is the y-axis position for person 17 in our plot. Yes! Looks like it .....</p>											

<p><b>31.</b></p>	<p>So now let's check the conversion formulae for person 17:          BLOT (x-axis)*slope + y-axis intercept = Predicted PRTIII  <math>3.20 * 1.94 + -5.04 = 1.17</math> (observed: 1.23)          PRTIII (y-axis)/slope + x-axis intercept = Predicted BLOT  <math>1.23 / 1.94 + 2.60 = 3.23</math> (observed: 3.20)          which are close enough .... see the adjacent Figure.</p>	
<p><b>32.</b></p>	<p>But how well does this work?          We are doing a PRTIII analysis, let's rescale the PRTIII measures into the predicted BLOT measures:           Winsteps analysis menu bar          Specification          Type in: UIMEAN = 2.60 (the x-axis intercept)          Click on "OK and Again"</p>	
<p><b>33.</b></p>	<p>Type in USCALE = 0.52 (= 1/SLOPE)          Click "OK"</p>	
<p><b>34.</b></p>	<p>Winsteps menu bar          Click on "Plots"          Click on "Scatterplots"          Same again .....</p> <p>Click on "Markers"</p>	
<p><b>35.</b></p>	<p>Good! The conversion has worked.           The trend line is now the identity line.          This plot is in BLOT units.</p>	
<p><b>36.</b></p>	<p>But how has this worked for all the other persons?          Scatterplot again, but change the file to  <b>BF5pf.txt</b>          which has all the persons ....          Click OK           Click on "Marker"</p>	

<p>37.</p>	<p>No, this doesn't look good - even though we have done our best to equate the two sets of person measures.  <i>Green arrow:</i> we expected the dashed trend line to be along the identity line.          But the identity line is within the confidence intervals, so we can't reject the hypothesis that we have the correct equating line. The first 24 persons were not a good equating sample from the persons. They are too central. There are many techniques for selecting the best equating sample, and choosing the best line. We do the best that we can do, but it is not perfect.</p>	
<p>38.</p>	<p>Look at the bottom of the Excel Worksheet:          We see that the correlation between the PRTIII and BLOT measures is 0.64, which means that only 40% of their variances are shared. But this correlation is "attenuated" by measurement error.</p>	
<p>39.</p>	<p>"Attenuation" of correlations is counter-intuitive.           Imagine we have two measures that are perfectly correlated = 1.0. But we cannot measure them perfectly. We measure them with measurement error. So the correlation is "attenuated" (=reduced) by measurement error. Correlation = 0.9.   <b>Every correlation we observe is attenuated by measurement error, because we can never measure perfectly.</b>           Then we discover how to remove ("<b>disattenuate</b>") the measurement error. After the measurement error is removed, the correlation becomes 1.0. The "disattenuated" correlation is always higher (more extreme) than the observed correlation.</p>	
<p>40.</p>	<p>Look at the bottom of the Excel Worksheet:          We see that the correlation between the PRTIII and BLOT measures is 0.64.          If we remove the measurement error, then the Disattenuated Correlation is 1.0. The "true" relationship between the two sets of measures could be perfect!</p>	
<p>41.</p>	<p><i>Conversion tables - Table 20 - USCALE UMEAN</i>          Here is a general instruction for "common item" (same items on two tests) or "common person" (same persons respond to two tests). We need to put both sets of measurements onto the same "ruler" (frame-of-reference). We usually choose one test as the "baseline", and then adjust the measures on the other test to match it. UIMEAN= and USCALE= are a convenient way of making the measures match each other.</p> <ol style="list-style-type: none"> <li>1. Choose one instrument as the baseline.</li> <li>2. Its Table 20 (or SCFILE=) is the score-to-measure conversion.</li> <li>3. For the other instrument, set USCALE=(equating slope) UMEAN=(equating constant).</li> <li>4. Analyze the data.</li> <li>5. Check that the numerical difficulties of the common items are almost the same in both analyses.</li> <li>6. Table 20 (or SCFILE=) for this analysis is the equated score-to-measure conversion.</li> </ol>	
<p>42.</p>	<p>Please close all windows</p>	

--	--

**43. C. Equating with Common Items**

**44.** “Test Equating, Scaling, and Linking: Methods and Practices” by Michael J. Kolen, Robert L. Brennan, (Springer, 2004) uses two illustrative datasets. They are dichotomous (0-1) responses to 36 items by 1600+ persons. The two datasets have 12 common items. Let’s equate these two datasets, called Mx.txt and My.txt

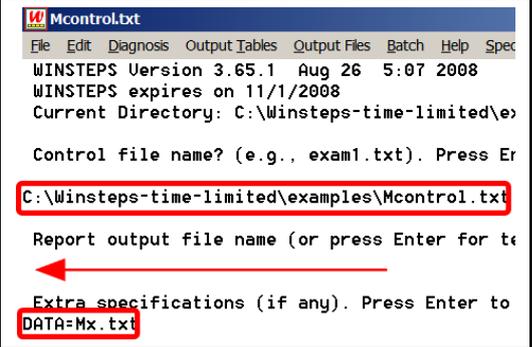
```

111110111010100110100100110000100100018 4
1100111011010100000000001010010014 3
1001100010110110000000000000000000 9 3
1110010110110000011001101000100000015 5
01111111110111101111011111101110030 9
1110110100110001000000000000000011 5
11001001001000000100000000000000 8 4
00000000000000000000000000000000 3 0
100011001000100000000010010000010010 9 2
11111111110111101101010101010000125 8
0111010000010000000110100001000000 9 4
110111011110110110010101010000000019 6
1111111001101100010101000000000017 6
101001010100110000000100100010000011 6
1110100101010000110111000001100001017 6
011000010000011010000001000010000 9 4
1100110100001000001001000001001011 3
    
```

**45.** Launch Winsteps



**46.** We will analyze file, Mx.txt, using control file Mcontrol.txt  
Mcontrol.txt has no data, so we will need to specify the data file.  
  
Open File:  
**Mcontrol.txt**  
Extra Specifications:  
**Data=Mx.txt**



**47.** Estimation completes.  
Mx.txt has 1655 persons and 36 items

Data from Kolen and Brennan (2004)

PERSONS	1655	INPUT	1655	MEASURED	INFIT	OUTFIT
MEAN	15.8	36.8		-.31	.41	1.00
S.D.	6.5	.0		1.00	.95	.15
REAL RMSE	.41	ADJ. SD	.91	SEPARATION	2.19	PERSON RELIABILITY
						.831

ITEMS	36	INPUT	36	MEASURED	INFIT	OUTFIT
MEAN	127.3	1655.0		.06	1.00	-.5
S.D.	294.6	.0		.98	.01	.09
REAL RMSE	.06	ADJ. SD	.97	SEPARATION	16.13	ITEM RELIABILITY
						1.001

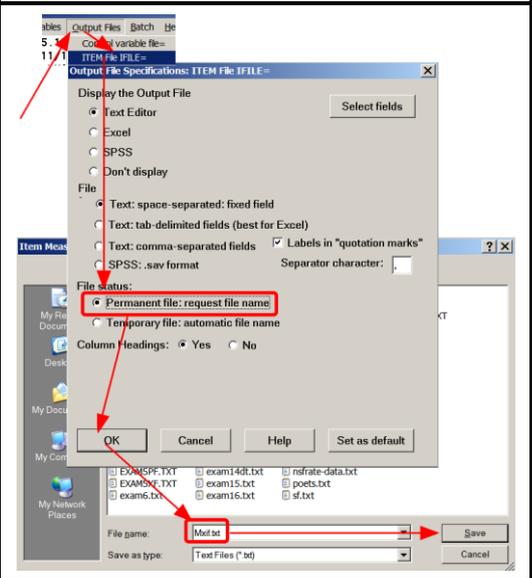
Output written to C:\Winsteps-time-limited\examples\200844WS.TXT  
CODES= 01  
features constructed: use "Diagnosis" and "Output Tables" menus

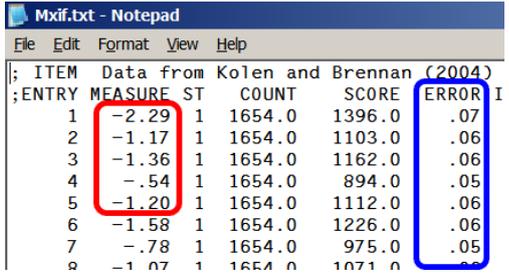
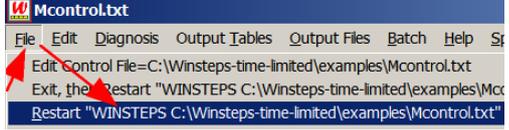
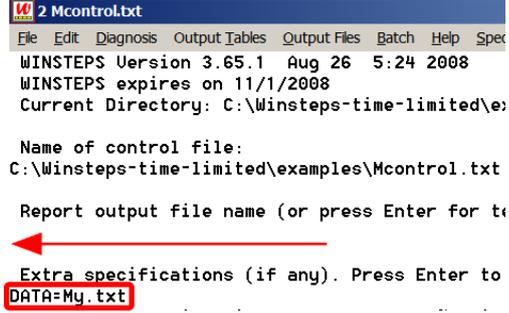
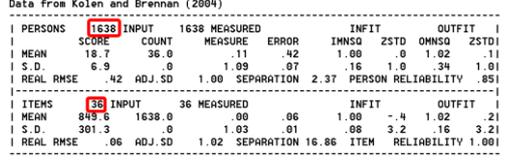
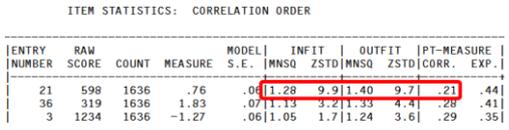
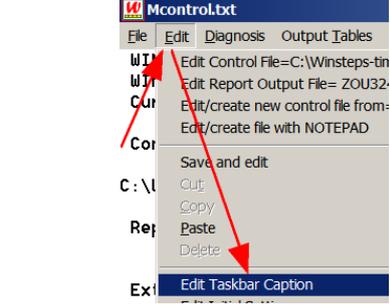
**48.** As usual, confirm that everything is OK in Mx.txt  
Diagnosis menu.  
A. Item Polarity  
Yes, the worst item looks acceptable ....

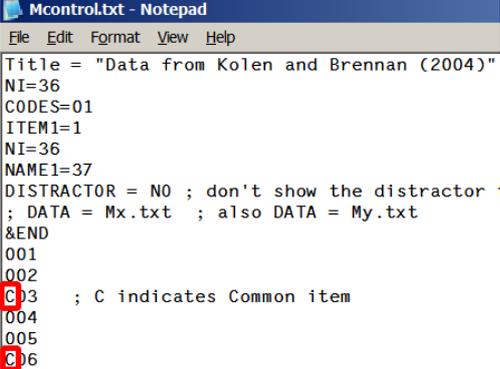
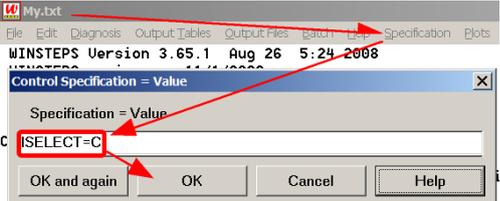
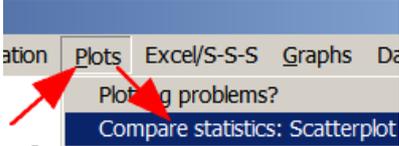
ITEM STATISTICS: CORRELATION ORDER

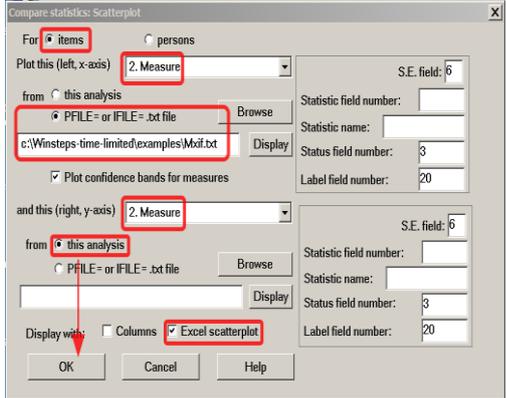
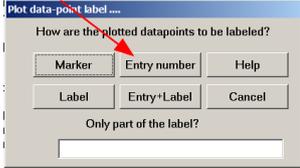
ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	S.E.	INFIT	OUTFIT	IPT-MEASURE	[EXACT MATCH]	ITEM
35	213	1654	1.94	.08	1.19	3.41	1.50	5.01	.16
34	300	1654	1.47	.07	1.22	4.71	1.32	4.51	.19
21	563	1654	-.45	.06	1.21	7.61	1.26	6.81	.23
31	506	1654	.64	.08	1.15	5.21	1.27	6.21	.26
1	1396	1654	-2.29	.07	.97	-7.11	1.10	1.21	.29
36	253	1654	1.71	.07	1.06	1.31	1.21	2.61	.29

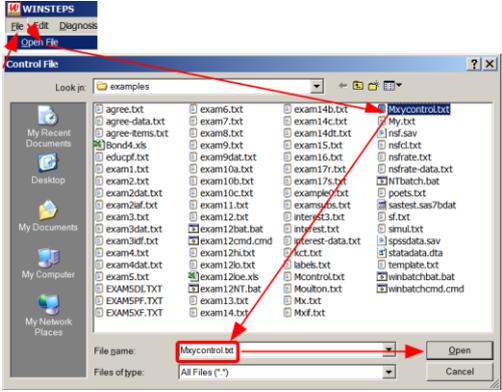
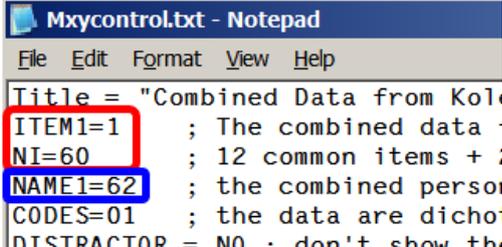
**49.** Save the item difficulties ....  
  
Click on “Output Files” menu  
Click on “Item File: IFILE=”  
  
Click on “Permanent file”  
Click “OK”  
  
Save as Mxif.txt  
  
These are the difficulties for the 36 items in Mx.txt

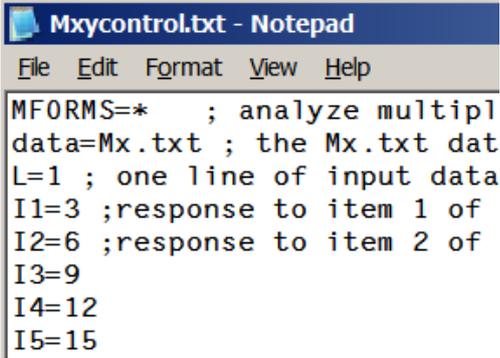


<p><b>50.</b> The item file displays in a Notepad window.</p> <p><i>Red box:</i> here are the logit item calibrations</p> <p><i>Blue box:</i> these are the standard errors (precision) of the measures. They are small because we have 1654 observations of each item.</p>		
<p><b>51.</b> Now for file My.txt Winsteps menu bar “Restart Winsteps ... Mcontrol.txt”</p>		
<p><b>52.</b> We are using the same control file, but specifying a different data file.</p> <p>Extra specifications? DATA=My.txt</p> <p>Then run the analysis.</p>		
<p><b>53.</b> In My.txt there are 1638 person records, and 36 items.</p>		
<p><b>54.</b> Confirm that everything is OK in My.txt Diagnosis menu. A. Item Polarity</p>		
<p><b>55.</b> We have two Winsteps analyses on the Windows task bar. It is easy to get them confused. Let’s give them more distinctive names. Click on the first one, Mcontrol.txt - the analysis of Mx.txt</p>		
<p><b>56.</b> Winsteps menu bar Click on Edit Click on Edit Taskbar Caption</p>		
<p><b>57.</b> In the “Edit Taskbar Caption” dialog box: Type the name of the data file: <b>Mx.txt</b></p> <p>Click on Ok</p>		

58.	Now on the Windows taskbar we have Mx.txt	
59.	Click on other Winsteps analysis Edit menu Edit Taskbar Caption My.txt OK Then on the Windows taskbar .....	
60.	Click on My.txt Winsteps menu bar Edit menu Edit control file: Mcontrol.txt	
61.	Mcontrol.txt displays in a Notepad window.  The items are numbered 1 to 36. They are different items in the two data files, except for the 12 common items.  Conveniently they are in the same columns in both data files.  I have put a "C" as the first letter of the item labels for the common items.	
62.	12 is a generous number of common items. The minimum number is around 5, uniformly distributed across the expected range of person ability. Any less, and we may not have enough items if some seriously misfit in our analyses or their item difficulties change (drift). In Rasch methodology, it is the number of items, rather than the proportion (percentage) of items which is crucial.	
63.	<i>Let's equate them using separate estimation!</i> For My.txt analysis, Winsteps menu bar Click on Specification We want to look at the Common Items, which have a "C" in the first character of the item label: ISELECT=C	
64.	The Winsteps analysis window confirms we have selected the 12 common items	<p style="text-align: center;"><b>ISELECT=C</b> <b>ITEMS SELECTED: 12</b></p>
65.	Now to scatterplot .... My.txt analysis: Winsteps menu bar Click on Plots Click on Scatterplots	

<p>66.</p>	<p>In the “Compare Statistics” Dialog box:</p> <p>Click on “Items”</p> <p><i>We want the item difficulties from Mxif.txt on the x-axis</i></p> <p>“Measures”</p> <p>Browse to Mxif.txt</p> <p><i>and My.txt on the y-axis</i></p> <p>“Measures”</p> <p>“This analysis”</p> <p>“Excel scatterplot”</p> <p>Click on OK</p>																
<p>67.</p>	<p>“Data point” dialog box:</p> <p>Click on Entry Number</p>																
<p>68.</p>	<p>What an amazingly well-behaved plot!</p> <p>None of the difficulties are conspicuously outside the very tight confidence intervals.</p> <p><i>Red box:</i> This is one logit wide and one logit high. The empirical line is the diagonal of the red box, so it appears to be parallel to the identity line.</p>																
<p>69.</p>	<p>Confirm the slope of the empirical line in the worksheet:</p> <p>1.02, as near to 1.0 as real data ever get.</p> <p>At this point, we can apply the conversion formulae we used above. But this time it will be easier to do a concurrent analysis</p>	<table border="1" data-bbox="975 1223 1481 1402"> <tr> <td>24</td> <td>Empirical intercept with x-axis</td> <td>-0.06798</td> </tr> <tr> <td>25</td> <td>Empirical intercept with y-axis</td> <td>0.0692</td> </tr> <tr> <td>26</td> <td>Empirical slope</td> <td>1.017983</td> </tr> <tr> <td>27</td> <td>Correlation</td> <td>0.994431</td> </tr> <tr> <td>28</td> <td>Disattenuated Correlation</td> <td>1</td> </tr> </table>	24	Empirical intercept with x-axis	-0.06798	25	Empirical intercept with y-axis	0.0692	26	Empirical slope	1.017983	27	Correlation	0.994431	28	Disattenuated Correlation	1
24	Empirical intercept with x-axis	-0.06798															
25	Empirical intercept with y-axis	0.0692															
26	Empirical slope	1.017983															
27	Correlation	0.994431															
28	Disattenuated Correlation	1															
<p>70.</p>	<p>Please close all windows</p>																

71.	<b>D. Common-item Concurrent Equating</b>	
72.	<p>We have two tests, Mx.txt and My.txt. They have 12 items in common, and 24 unique items each. We want to analyze them together: 12+24+24 = 60 items. Mx.txt has 1655 persons. My.txt has 1638 persons. Combined they have 3293 person. The Winsteps control variable, MFORMS=, provides a convenient way to combine these data so that all the measures are reported in the same frame-of-reference. It is in Winsteps Help and at <a href="http://www.winsteps.com/winman/mforms.htm">http://www.winsteps.com/winman/mforms.htm</a></p>	<p><b>MFORMS= reformat input data and multiple data forms</b> <a href="#">Up Previous</a> <a href="#">Next</a></p> <p>MFORMS= supports the reformatting of input data records, and also <a href="#">equating</a> multiple input files in different formats, such as alternate forms of the same test. Data after <a href="#">END NAMES</a> or <a href="#">END LABELS</a> is processed first, as is data specified by <a href="#">DATA=</a> in the core control file.</p> <p>Data reformatted by MFORMS= can be accessed, viewed, edited and "saved as" permanently using the "Edit" pull-down menu. It has a file name of the form: ZMF.....txt</p> <p>Here is the layout:</p> <pre>mforms=* data=Forma.txt ; the name of an input data file L=2 ; there are 2 lines in input data file for each data record I1 = 20 ; response to item 1 of the test is in column 20 of the input data file I3-5 = 21 ; items 3, 4, 5 are in columns 21, 22, 23 of the input data file I16-20=11 ; items 16, 17, 18, 19, 20 are in columns 11, 12, 13, 14, 15 P1=9 ; the first character of person label is in column 9 of the input data file</pre>
73.	Please launch Winsteps	
74.	<p>Winsteps menu bar Click on Open file Double-click on "Mxycontrol.txt" in the examples folder</p> <p>Report ...? Press Enter Extra specifications? Press Enter Run the analysis</p>	
75.	<p>The Analysis completes ...</p> <p>3293 persons = 1655 + 1638. <i>Correct!</i> 60 items = 12 + 24 + 24. <i>Correct!</i></p> <p>We have concurrently (one-step) equated the two tests. But how did we do it? And how can we use it?</p>	<pre>-----   PERSONS   3293 INPUT   3293 MEASURED             SCORE      COUNT      MEASURE  ERF   MEAN      17.2      36.0      - .08   S. D.     6.8       .0      1.05   REAL RMSE .42     ADJ. SD     .97 SEPARAT:  -----   ITEMS     60 INPUT    60 MEASURED   MEAN      946.2     1975.8      .00   S. D.     477.2     658.6      .99   REAL RMSE .06     ADJ. SD     .99 SEPARAT:  -----</pre>
76.	<p>Winsteps menu bar Click on Edit Click on Edit Control File</p>	
77.	<p>Mxycontrol.txt displays in a Notepad window We define the combined Mx.txt+My.txt data file: ITEM1=1 - the combined item responses start in column 1 NI=60 - there are 60 different items NAME1=62 - the person identification starts in column 62 <i>Leave a blank column between the item responses and the person label so that it easy to see that the layout is correct.</i></p>	

<p><b>78.</b> Scroll down to the item labels &amp;END - end of the control variables</p> <p><i>Red box:</i> the 12 common items. I have kept the same item labels as in Mx.txt and My.txt</p> <p><i>Blue box:</i> the 24 items unique to Mx.txt. I have changed the Mx.txt item labels to start “X”</p> <p><i>Green box:</i> the 24 items unique to My.txt. I have changed the My.txt item labels to start “Y”</p> <p>END LABELS (Notice that there are no data below)</p>		<pre> &amp;END C03 ; C - 12 Common items C06 C09 C12 C15 C18 C21 C24 C27 C30 C33 C36 X01 ; X - 24 items unique to Mx.txt X02 X04 X05 X07  Y32 Y34 Y35 END LABELS </pre>
<p><b>79.</b> Scroll back up to MFORMS= Now we have some clerical work - but much less than trying to rearrange the data using “rectangular copying” (such as Alt+Mouse in Word) or many other methods of rearranging the data. MFORMS= input data files with different layouts. data=Mx.txt - input the Mx.txt data file. L=1 - each person record is one line in Mx.txt I1=3 - the first item in the combined data file is the first common item, C03, this is in column 3 of Mx.txt I2=6 - the second item, C06, is in column 6 of Mx.txt</p>		 <pre> Mxycontrol.txt - Notepad File Edit Format View Help MFORMS=* ; analyze multiple data=Mx.txt ; the Mx.txt data L=1 ; one line of input data I1=3 ; response to item 1 of I2=6 ; response to item 2 of I3=9 I4=12 I5=15 </pre>
<p><b>80.</b> I12=36 - the twelfth item is the last common item C36 I13-14=1 - now we have the items unique to Mx.txt. Columns 13-14 of the combined data have items X01 and X02 (originally 001 and 002) of Mx.txt which are in columns 1 and 2 of Mx.txt.</p>		<pre> I12=36 I13-14=1 ; response to item I15-16=4 ; response to item I17-18=7 I19-20=10 </pre>
<p><b>81.</b> I35-36=34 - columns 35 and 36 of the combined file hold the last two unique items of Mx.txt, X34 and X35. P1=“X” - this puts character X as the first character of the person label P2-5=37 - X is followed by the 4 characters of person label from column 37 onwards of Mx.txt. # - end of instructions for Mx.txt</p>		<pre> I35-36=34 ; We have procedure P1="X" ; first character of P2-5=37 ; characters 2-5 of # </pre>
<p><b>82.</b> Same again, but for My.txt</p>		<pre> data=My.txt ; the My.txt data L=1 ; one line of input data I1=3 ; response to item 1 of I2=6 ; response to item 2 of I3=9 I4=12 I5=15 </pre>
<p><b>83.</b> I37-38=1 - Items unique to My.txt follow to the right of the unique items for Mx.txt</p>		<pre> I37-38=1 ; response to item I39-40=4 ; response to item I41-42=7 I43-44=10 I45-46=13 I47-48=16 </pre>

<p><b>84.</b> I59-60=34 - these are the last two unique items of My.txt  P1="Y" - first character of combined person label is Y for My.txt  P2-5=37 - the person label information from My.txt  * - end of MFORMS=  &amp;END - end of control variables</p>		<pre>I59-60=34 ; We ha P1="Y" ; first ch P2-5=37 ; charact * &amp;END</pre>
<p><b>85.</b> Now, what did this do?  Winsteps Analysis window, near the top  Red box: Mx.txt is processed  Blue box: My.txt is processed  Orange box: They are reformatted into a temporary file  Green box: The temporary file is processed as the data file.</p>		
<p><b>86.</b> The temporary file is available:  Winsteps menu bar  Click on Edit  Click on Edit MFORMS File=    Temporary files are automatically deleted when Winsteps closes.</p>		
<p><b>87.</b> The reformatted temporary file displays in a Notepad window.  The first records are from Mx.txt.  The X (P1="X") is the first character of the person label</p>		
<p><b>88.</b> Scroll down the file ...  We can see the transition from Mx.txt to My.txt.  The 12 common items. Then the unique Mx.txt items.  Then the unique My.txt items.  Notice that a wonderful feature of Rasch methodology is that the missing data doesn't matter! The measures are estimated from the observations. Fewer observations means less precision. There is no need to impute missing data nor to delete cases, items or whatever.</p>		
<p><b>89.</b> And then My.txt finishes at the bottom.  If you want to keep this file, then "Save As" a permanent file. Its control file would be Mxycontrol.txt, omitting the MFORMS= instructions</p>		
<p><b>90.</b> Equating of polytomous items is somewhat more complex. Which rating-scale structure is the decisive one?  We usually need to anchor this SAFILE= early in the equating study.  Test A or Test B or a joint analysis of Test A and Test B?  See the suggestions at  <a href="http://www.rasch.org/rmt/rmt101f.htm">http://www.rasch.org/rmt/rmt101f.htm</a></p>		

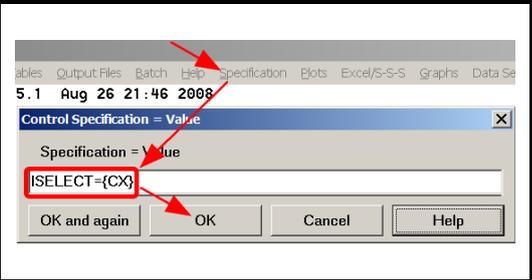
**91. E. Reporting the Concurrent-Equating Measures**

**92.** We have one analysis with all the items and persons in the same frame-of-reference.  
 Winsteps menu bar  
 Click on Output Tables  
 Click on 14. Items in Entry Order

There are all the items  
 Red box: 12 Common items  
 Blue box: 24 items unique to Mx.txt  
 Green box: 24 items unique to My.txt

This combined listing would be useful if we wanted to make an item-bank of 60 items.

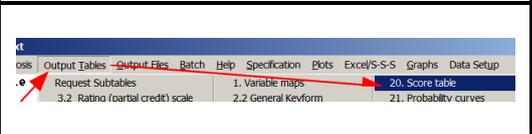
**93.** But suppose we need a score-to-measure Table for future administrations of the Mx.txt test.  
 Windows menu bar  
 Click on Specification  
 Type in: ISELECT={CX}  
 (Select item labels starting C or X)  
 Click on OK



**94.** In the Windows analysis window, we see that 36 items have been selected. *Correct!*

**ISELECT={CX}**  
**ITEMS SELECTED: 36**

**95.** Winsteps menu bar  
 Click on Output Tables  
 Click on 20. Score table

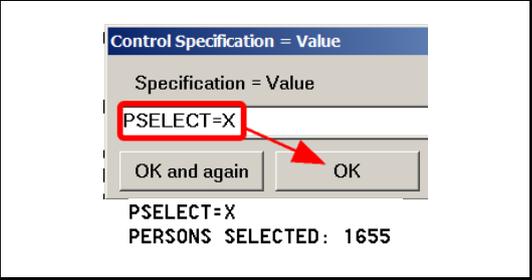


**96.** And, in a Notepad window, we have the measures corresponding to scores from 0 to 36 on the Mx.txt test.

SCORE	MEASURE	S.E.	SCORE	MEASURE	S.E.	SCORE	MEASURE	S.E.
0	-5.16E	1.84	13	-.62	.38	26	1.22	.40
1	-3.91	1.03	14	-.48	.37	27	1.39	.42
2	-3.16	.75	15	-.34	.37	28	1.57	.43
3	-2.69	.63	16	-.20	.37	29	1.76	.45
4	-2.35	.56	17	-.07	.37	30	1.98	.48
5	-2.06	.51	18	.07	.37	31	2.22	.51
6	-1.82	.47	19	.20	.37	32	2.50	.56
7	-1.61	.45	20	.34	.37	33	2.85	.63
8	-1.42	.43	21	.48	.37	34	3.31	.75
9	-1.24	.41	22	.61	.38	35	4.06	1.03
10	-1.07	.40	23	.76	.38	36	5.31E	1.84
11	-.91	.39	24	.90	.39			
12	-.76	.38	25	1.06	.39			

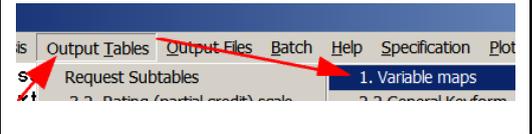
**97.** To report the persons in Mx.txt test:  
 Windows menu bar  
 Click on Specification  
 Type in: PSELECT=X  
 Click on OK

Persons selected = 1655. *Correct again!*

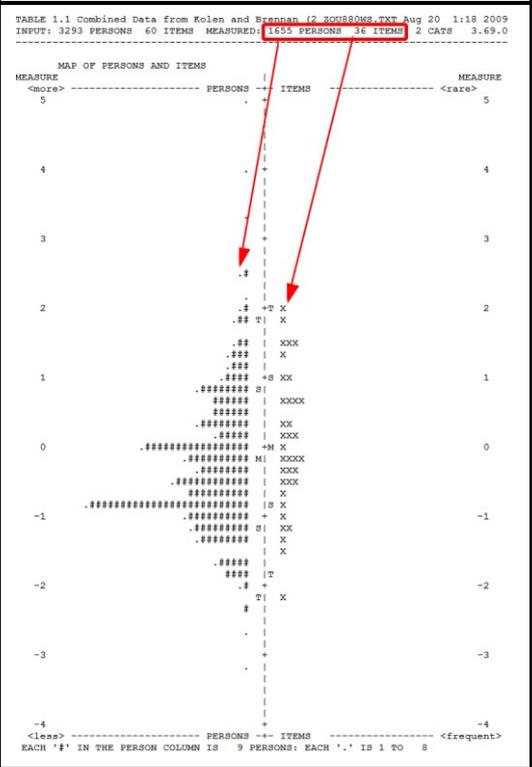


**98. F. Improving the Appearance of the Tables**

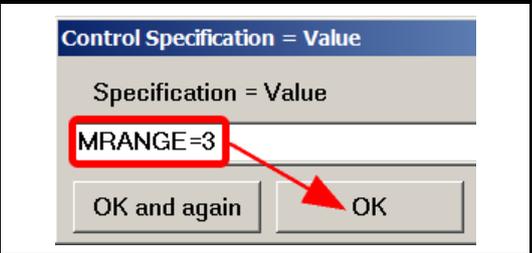
**99.** So now we can see the “map” for Mx.txt  
 Winsteps menu bar  
 Click on Output Tables  
 Click on 1. Variable maps



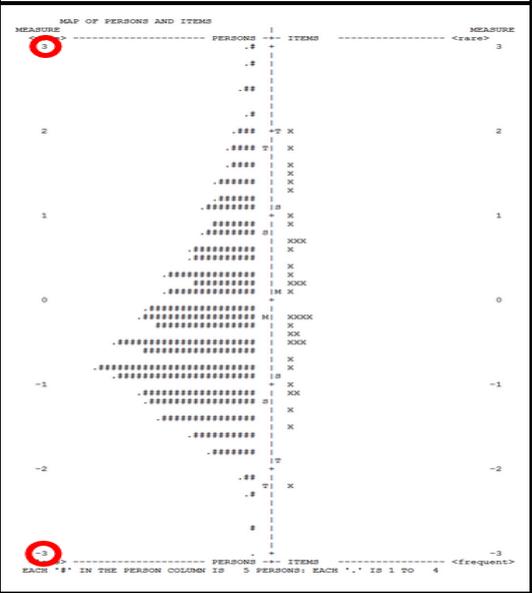
**100.** The Table 1. Variable Map displays in a Notepad window.  
 On the left are the 1655 persons. On the right are the 36 items.  
 Let’s “prettify” (beautify) the map ...  
 The range shown is +5 to -4. We only need +3 to -3  
*You could also change to a more user-friendly scaling with UIMEAN= and USCALE=*



**101.** Winsteps menu bar  
 Click on Specification  
 Type in: MRANGE=3  
 Click on OK  
 MRANGE= specifies the half-range to display around the local origin.



**102.** Winsteps menu bar  
 Click on Output Tables  
 Click on 1. Variable maps  
 This Table-Map is too long.  
 Let’s squash it.





**108.** Winsteps menu bar  
 Click on Output Tables  
 Click on 1. Variable maps

Elegant! The lines look much neater.  
 You can copy-and-paste this into a Word document, but you will need to set the font to Letter Gothic Line in Word.  
*Sorry. This does not work in some versions of Windows*

To make the Notepad font change permanent: see Winsteps Help - "Notepad"  
 So let's try the third display option .....

**109.** Winsteps menu bar  
 Click on Specification  
 Type in: ASCII=Webpage  
 Click on OK  
 This specifies HTML output

**110.** Cool! *Red box*: the variable map displays in your Internet Browser software.

You can copy-and-paste this direct to Word without losing the font or the formatting

**111.** Reinstate all the items:  
 Windows menu bar  
 Click on Specification  
 Type in: ISELECT=  
 Click on OK

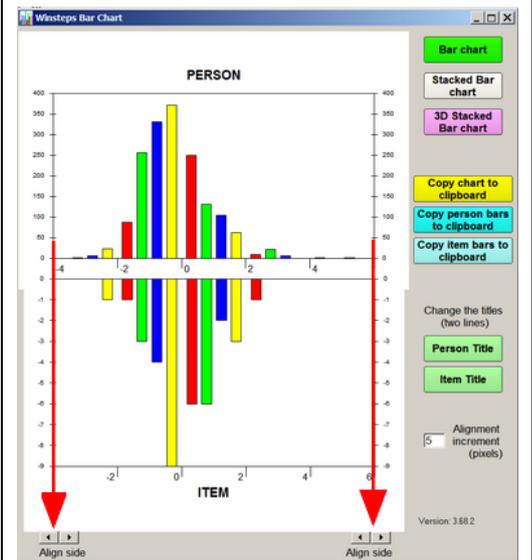
**ISELECT=**  
**ITEMS SELECTED: 60**

**112.** And a different way to see the same distribution:  
 Graphs menu  
 Click on: Person-item Barchart

113. You see a graphical distribution chart ....

If the sides at the top and bottom do not align correctly, please adjust them with the arrows at the bottom of the window.

This plot is an experimental prototype - a better-looking plot is coming!



114. *Oh no! Our time is up, and there is so much more ....*  
Please close all windows



115. Thank you for joining in this exciting exploration. We have all learned so much ....