# Introducing Lertap5, Version 11

A glimpse of the key features found in Lertap5 for Microsoft Excel
(Windows and Macintosh)

Larry R Nelson, [Curtin University](), Australia

This presentation was last updated on: 25 August 2022.

S2: Example 1, The MathsQuiz dataset

These results are from a quiz given to high school maths classes in a Chicago school district.

The students were presented with 15 multiple-choice questions (or "items").

Each question had four options.

This slide shows the answers given by the first 16 students.

In total, 999 students took the quiz.

*This is how Lertap5 typically looks in Excel. The data above are from a "**cognitive**" instrument, a test. NOTE: it's possible to download this dataset if you want to play along: get it here. Download Lertap itself from here.*

S3: Example 2, The Stigma of Mental Illness Scale

The Data worksheet seen here is typical of survey data (often referred to as "**affective**" data).

It shows the responses of 15 people to the first 9 questions on an instrument, the "IMSI29", having a total of 29 questions.

Many times surveys are anonymous, and that was the case here: there are no ID numbers in the first column.

The gender of each respondent was coded in column 30 (not seen here). It's possible to breakdown results by gender (see this link).

S4: The Elements of a Data Worksheet

1 points to the Lertap **tab** on the Excel **ribbon**.

Note that the icons on the tab are organised in sections: "Basic options", "New menu", "Run menu", "Graphics trio", and "Other menus".

2 points to **worksheets** called "Data" and "CCs". Excel is currently showing the contents of the first 18 rows and 16 columns of the Data sheet.

Student 4 selected option **2** on item **I8**.

S5: The CCs worksheet for the MathsQuiz data.
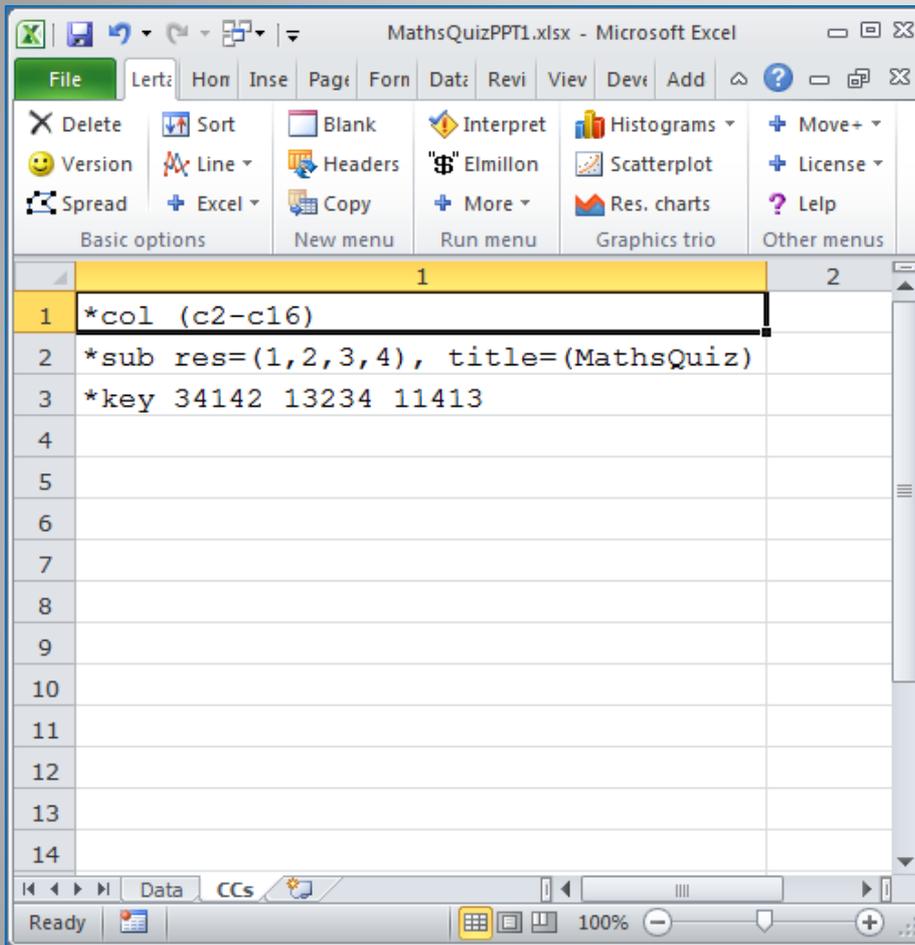
The CCs sheet is used to set up a Lertap analysis. CCs means "Control Cards".

This CCs sheet has three "cards": *col, *sub, and *key. In this case they're for a test.

The *col card says that answers start in column 2 of the Data sheet, and continue out to column 16.

The res= control "word" on the *sub card says that this test's items used options 1, 2, 3, 4 . This "word" is not needed when item options are A,B,C,D.

The *key card gives the "keyed-correct" answer for each item. There are 15 entries on this card, one for each of the 15 items.

CCs sheets will always have a minimum of two "control cards". For cognitive instruments (quizzes, tests) *col and *key cards are always needed, while *col and *sub cards are always required for surveys and scales (affective instruments). Often other "cards" are included. Read more about them here.

S6: The CCs worksheet for the IMSI-29 survey.

This CCs sheet was used to process responses on the 29-question version of the Internalized Stigma of Mental Illness Scale, the "IMSI-29" (Slide S3).

The *col card says that answers start in column 1 of the Data sheet, and continue on to column 29.

The aff= control "word" on the *sub card says that the data are "affective" responses (indicating survey items).

The res= specification indicates the response codes used by each of the survey questions.

Read more about CCs cards here.

Having now peeked at the Data and CCs sheets, we're set to get results. The 1st step: click on the "Interpret" option (see next slide).

S7: The Freqs "report"

Click on the ⚠️ Interpret option on the Run menu. If the CCs cards are all okay, Excel will create a new worksheet called "Freqs", short for frequencies.

Here we see results for items I1 and I2. The most popular option on I1 was 3, while for I2 it was 4.

What's a "response" of 9 doing here? In this dataset, a 9 was used if a student did not answer an item. What about a 5? (Six students had a response of 5 on I2.)

These 5s were errors. Bubble sheets were used to get student answers, and some students, probably intending to shade in bubble 4, made a mistake and shaded in the adjacent bubble, 5.

Click here for more about Freqs.

I1 (c2)

Option	n	/999
1	27	2.7%
2	85	8.5%
3	843	84.4%
4	38	3.8%
9	6	0.6%

I2 (c3)

Option	n	/999
1	7	0.7%
2	6	0.6%
3	5	0.5%
4	971	97.2%
5	6	0.6%
9	4	0.4%

I3 (c4)

Option	n	/999
1	695	69.6%

The Freqs report provides the opportunity to check on data quality. Once satisfied that the data are sufficiently free from errors, you'll move on to the 2nd step, the "Elmillon" option, to get statistics.

S9: The Stats1b table for quizzes, tests, exams.

The 2nd step in a Lertap analysis: click the Elmillon option (blue arrow). This will get Excel to create a variety of statistical summaries.

One of these summaries, the "Stats1b" report, is highlighted when Elmillon finishes. It has two parts: a table with response percentages (seen on the left), and a chart of the item "diff" and "disc" values (next slide).

The last column in the table, **?**, is for item "flags". This column has entries when the responses to an item do not follow what might be expected. In this example, two of the options for I11 have been flagged, options **1** and **4**. (Click here for more info.)

The "b" in Stats1b means "brief". This report is a simple, concise summary of results. It's easy to have Excel print this report; comments on printing are here.

It is possible to combine the two steps (Interpret and Elmillon) so that only the first (Interpret) is needed. This is done by using a setting in the System worksheet. Click here to read more.

S10: The Stats1b plot for quizzes, tests, exams.
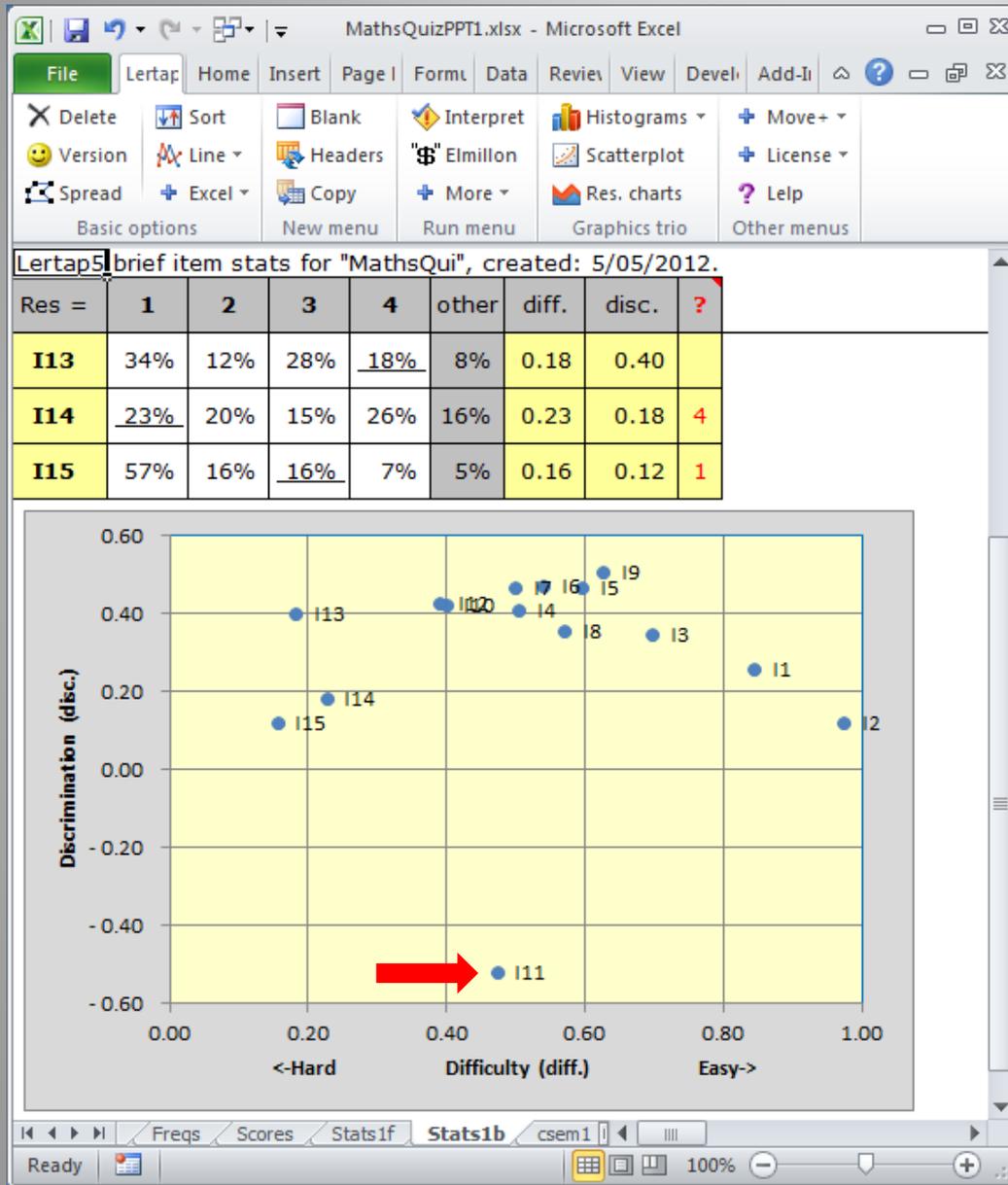
Lertap's Stats1b plot of item difficulty, "diff.", by discrimination, "disc.", is shown to the left.

When a test is meant to pick out the strongest students, this plot should not have any disc. values below 0.00.

But this one does. I11's discrimination value (disc.) was -0.52. This stems from an error in the *key card, a "mis-key", as discussed here.

For more plots of this type, see if your mouse will click here.

Checking for mis-keyed items is obviously important. Errors of this sort reduce a test's reliability and, if not corrected, will lead to further errors when it comes to interpreting and grading student results.

S11: The Stats1b table for surveys and scales.

This slide shows what the Stats1b report looks like when affective items are processed.

Here results are displayed for the initial 10 questions on the "IMSI-29" scale as collected from a sample of U.S. adults suffering from depression (Slide S3).

The items on this scale all used {1,2,3,4} as their response codes. The "other" column will have entries whenever missing responses are found.

An item will have a minus sign in the "pol." column if it has been reverse scored.

The "cor." column indicates the degree to which an item's responses correlate with the other items in the scale. The next slide displays a plot of the item "cor." figures.

Lertap5 brief item stats for "IMSI29US", created: 22-Apr-21.

Res->	1	2	3	4	other	pol.	mean	s.d.	cor.	h
Q1	16%	30%	39%	15%		+	2.52	0.93	0.70	
Q2	18%	40%	30%	11%		+	2.35	0.91	0.65	
Q3	13%	35%	41%	12%		+	2.52	0.86	0.47	
Q4	16%	31%	38%	15%		+	2.51	0.93	0.63	
Q5	22%	27%	37%	15%		+	2.45	0.99	0.62	
Q6	22%	30%	36%	11%		+	2.37	0.95	0.65	
Q7	25%	46%	25%	4%		+	2.09	0.82	0.50	
Q8	66%	29%	4%	1%		+	1.41	0.64	0.46	
Q9	60%	35%	4%	0%		+	1.44	0.59	0.24	
Q10	60%	32%	6%	2%		+	1.49	0.68	0.55	

S12: The Stats1b plot for surveys and scales.
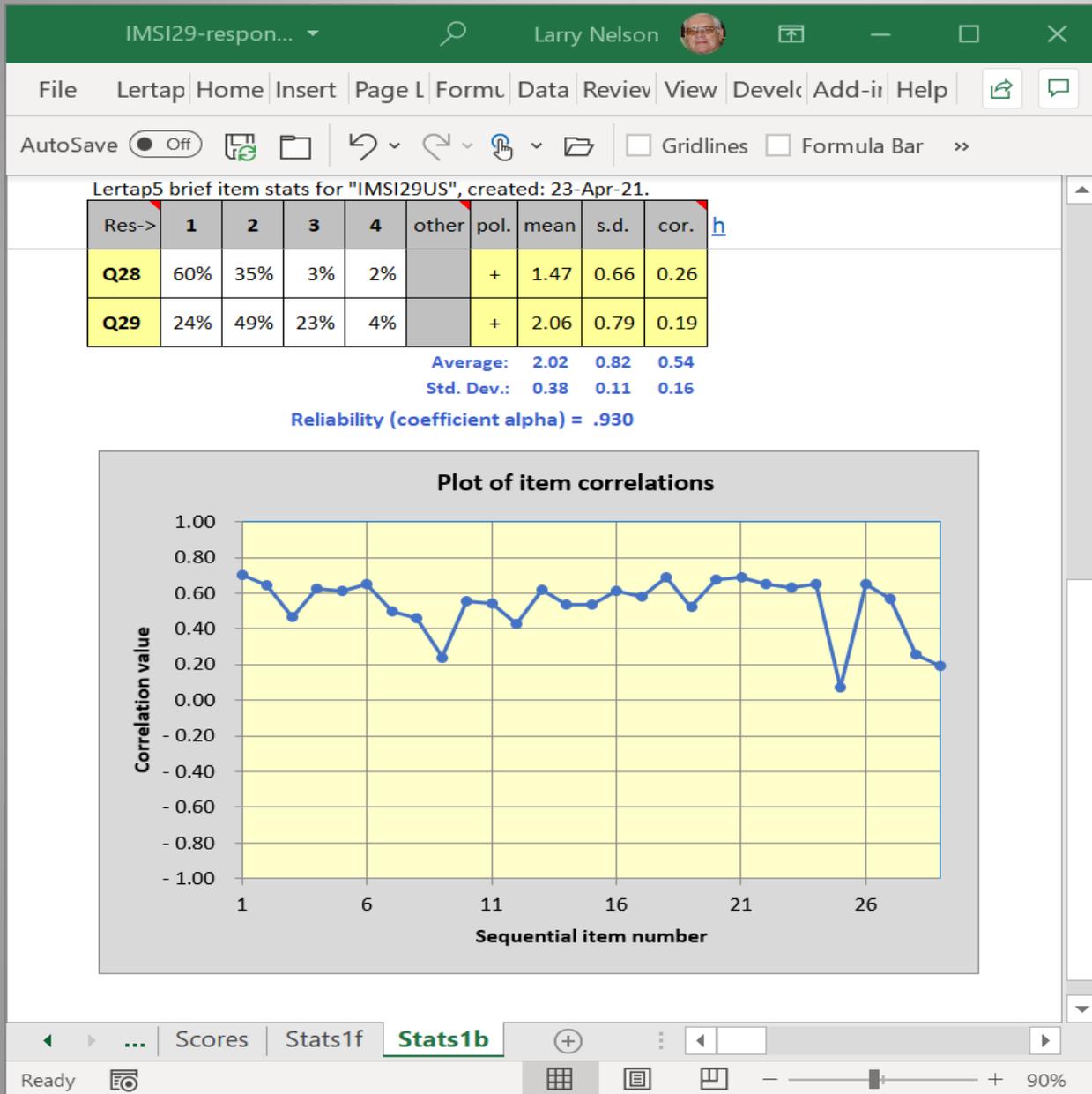
This snapshot shows the plot seen at the bottom of every Stats1b report for affective items.

The items in this case are those from the IMSI-29 mental illness scale.

Coefficient alpha reliability was good at 0.93.

Most items had correlations above 0.40. Alpha would be higher if four items (9, 25, 28, 29) were excluded from the scoring (should that be desired).

Lertap5 also computes McDonald's omega reliability. See this topic.

S13: The Stats1f summary for both tests and scales.

Summary statistics

number of scores (n):	999	
lowest score found:	0.00	(0.0%)
highest score found:	14.00	(93.3%)
median:	8.00	(53.3%)
mean (or average):	**7.67**	(51.1%)
standard deviation:	2.88	(19.2%)
standard deviation (as a sample):	2.89	(19.2%)
variance (sample):	8.32	

number of subtest items:	15	
minimum possible score:	0.00	
maximum possible score:	15.00	
reliability (coefficient alpha):	**0.68**	
coefficient alpha 95% CI:	0.65 < ---- > 0.71	
index of reliability:	0.83	
standard error of measurement:	**1.63**	(10.8%)

"Stats1f" is another standard report made by Elmillon. The "f" stands for "full". The results here pertain to the MathsQuiz cognitive test, Slide S2.

Read mode about Stats1f reports here – you'll see that they have several sections, with far more information than the summary statistics shown in this slide.

The reliability of the MathQuiz, 0.68, was low. This was largely due to one of the items, I11, being mis-keyed. Fixing the *key line in the CCs worksheet (Slide S5) to correct this error would bring the reliability up to about 0.80 (see the next slide).

Lertap5 will also compute McDonald's omega reliability estimate. See this topic

Stats1b and Stats1f are made for all instruments, be they cognitive or affective in nature. A third report, Stats1ul, is also created for cognitive instruments – see the next slide.

S14: The Stats1ul report for tests and exams.

Lertap5 U-L stats for "MathsQui", created: 5/05/2012.

Res =	1	2	3	4	other	U-L diff.	U-L disc.
I9 upper	0.01	0.01	0.97	0.00	0.01	**0.57**	**0.82**
2nd	0.07	0.03	0.90	0.00	0.01		
3rd	0.25	0.02	0.69	0.02	0.01		
4th	0.42	0.07	0.40	0.05	0.07		
lower	0.57	0.08	0.16	0.10	0.10		
I10 upper	0.01	0.14	0.02	0.84	0.00	**0.46**	**0.75**
2nd	0.04	0.33	0.05	0.59	0.00		
3rd	0.03	0.58	0.07	0.30	0.02		
4th	0.07	0.65	0.09	0.19	0.01		
lower	0.08	0.66	0.12	0.09	0.05		
I11 upper	0.16	0.00	0.01	0.84	0.00	**0.44**	**- 0.57**
2nd	0.34	0.01	0.03	0.63	0.00		
3rd	0.54	0.02	0.02	0.40	0.02		
4th	0.61	0.02	0.05	0.32	0.01		
lower	0.72	0.05	0.06	0.11	0.07		

Summary group statistics

	n	avg.	avg%	s.d.	min.	mdn.	max.
upper	199	11.8	79%	1.0	10	12	14
2nd	199	9.3	62%	0.7	8	9	10
3rd	203	7.6	50%	0.5	7	8	8
4th	199	5.8	39%	0.7	5	6	7
lower	199	3.8	25%	1.0	0	4	5
everyone	999	7.7	51%	2.9	0	8	14

"Stats1ul" is also a standard report made by Elmillon when cognitive items are involved.

Students have been grouped according to test score. The top 20% are in the "upper" group, while the bottom 20% are in the "lower". Altogether there are 5 groups in this report.

This provides another way of looking at item quality. If an item is meant to pick out the top students, then we'd want the pattern seen for I9, whose correct answer was 3 (underlined). In the upper group, 97% got this item correct, dropping to just 16% in the lower group. I10's pattern is similar (4 was the correct answer on I10).

But I11 has a problem. The correct answer appears to be 1. Wrong! An error was made on the CCs *key line: the right answer for I11 is 4, not 1.

Stats1b or Stats1ul. Which to use? Either. Both. They look at the same thing, item quality, from unique angles. Have a read of Chapter 2 in the manual when you have a chance.

S15: Quintile plots for cognitive items.
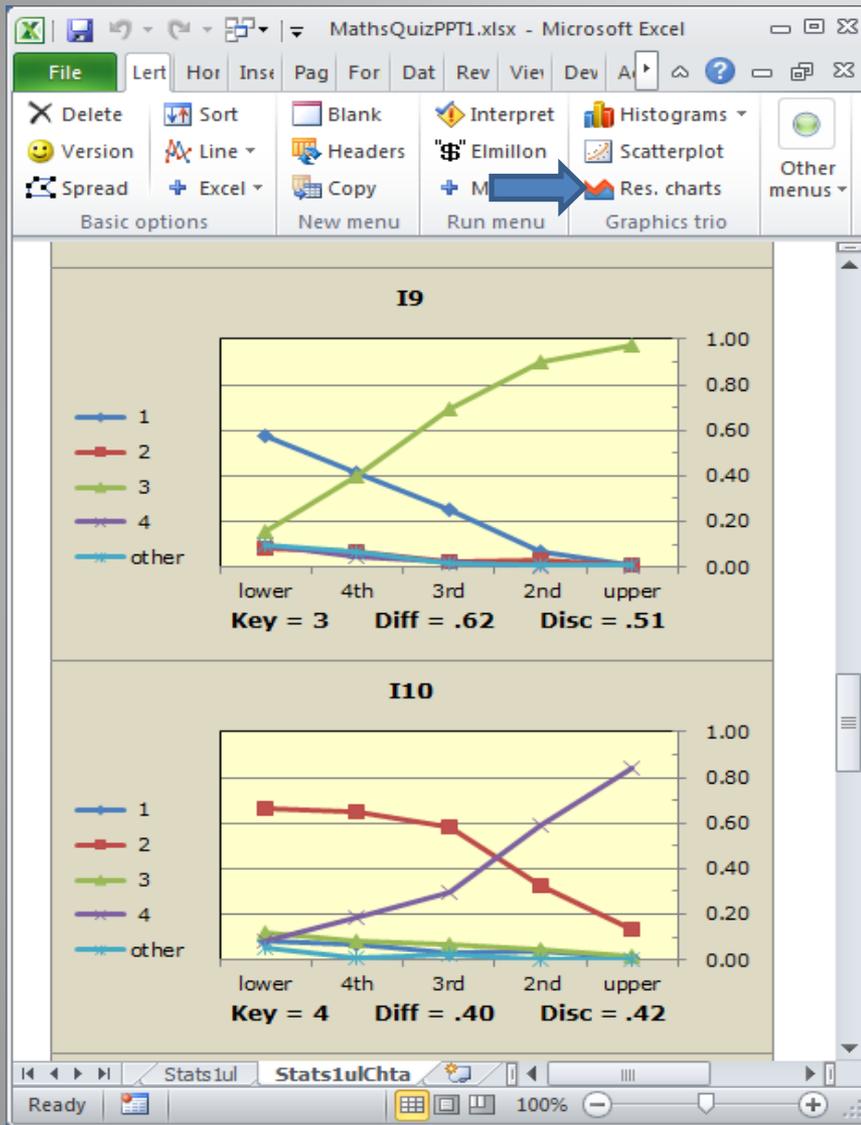
See where the blue arrow is in the picture to the left? It's pointing to the "Res. Charts" option (Response Charts).

Click on it to get pictures of how the test items have performed, as exemplified here for items I9 and I10 (two of the best-performing items in our little maths quiz).

The information in these handy little snap-shots is taken directly from Stats1ul results.

Items which are meant to pick out the best students should have graphs like these. In the lower groups, few students will have the nous required to pick out the right answer, while the top, or "upper", students should have little trouble.

Bone up on these plots by having a read of this stuff.

S16: Scores!
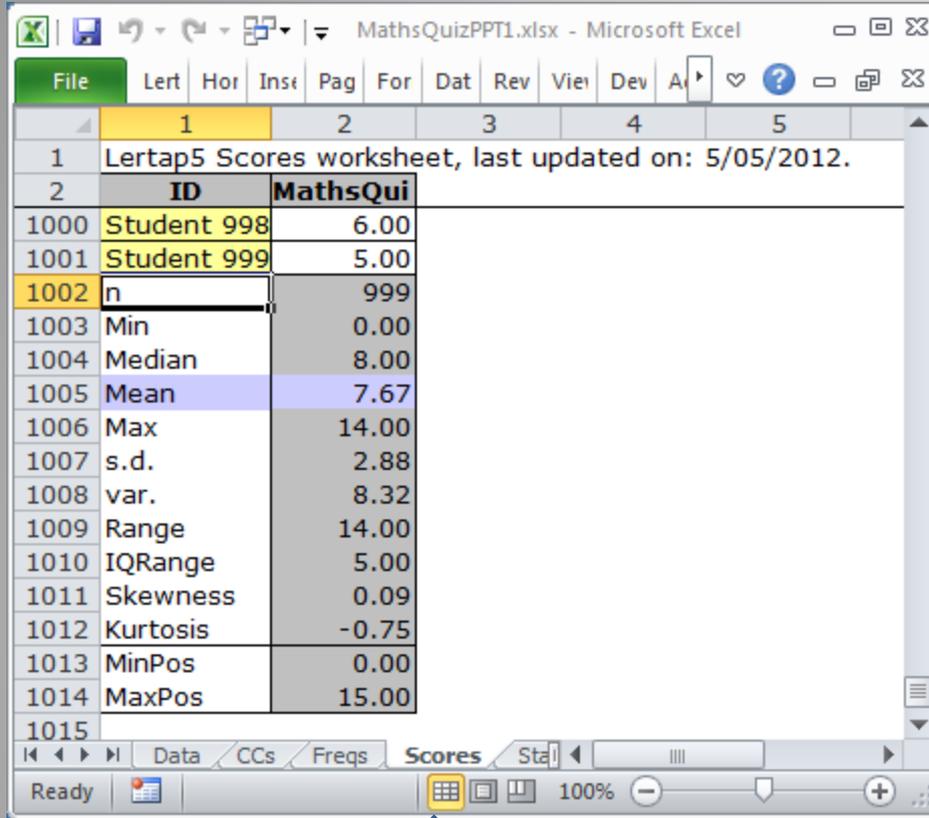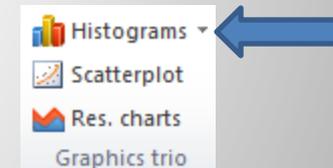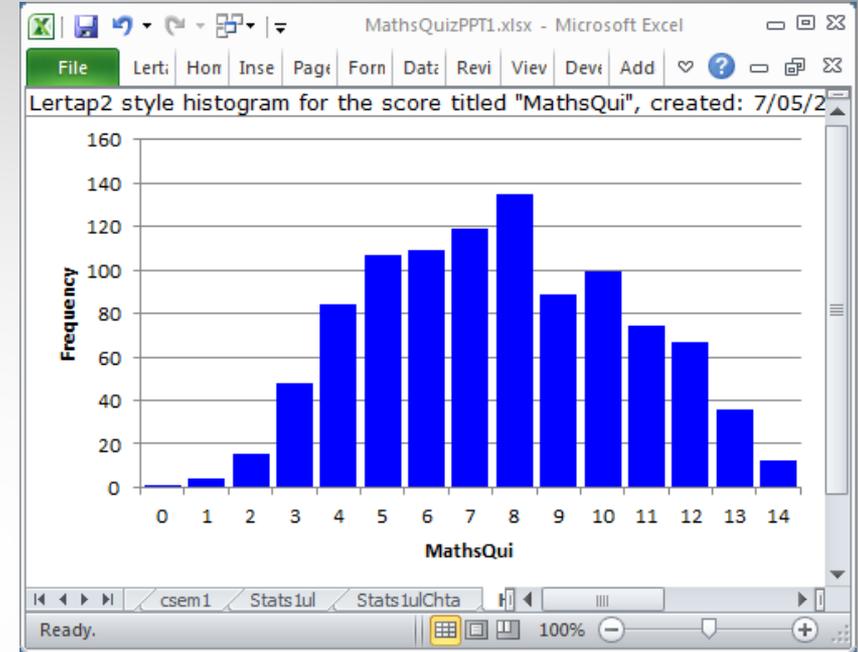
The "Scores" worksheet is another Emillon product. The histogram on the right was made by clicking on the "Histograms" option (natch!). Scores are made for both cognitive and affective instruments.

When scanners are used to process student response sheets, it's not uncommon to find a student or two with a valid ID, but no item responses. These "bad records" will usually show up at the left of a histogram. Read this topic for an example.

S17: What's been done to here

- We started with a display of a typical Lertap dataset: an Excel workbook with two worksheets, "Data", and "CCs". The workbook was named "MathsQuizPPT1.xlsx".

- You didn't see me, but I clicked on the [Interpret] option. This is always the first step when results are wanted. This produced the "Freqs" report.

- Again behind the scenes, I clicked on the ["$" Elmillon] option. This added several new worksheets, or "reports", to the workbook: Stats1b, Stats1f, Stats1ul, and Scores.

- I showed you two views of the Stats1b report, the table view, and the plot view. These views provide a concise summary of how the items performed. Slides S11 and S12 are similar for affective instruments.

- Then we looked at the summary section of the Stats1f report to check on test reliability.

- The Stats1ul report provides another way to examine item performance; it's an alternative (or complement) to Stats1b for cognitive instruments (see it again).

- I wowed you by getting some quintile plots. These provide one more way to see how cognitive items functioned, and are very popular.

- Finally I had you look at the Scores worksheet, with a histogram (see it again).

Was the MathsQuiz a good test? Can its results be used to pick out the best students? No. An error was made when the CCs *key line was created. Item I11's key was not correct and, as a result, the reliability was low. Is it possible to fix this error? But of course. Easy. (Click here for more info.)

S18: What has *not* been done to here

- I did not show you how to process results from mastery, criterion-referenced, and pass-fail tests, often used in licensing and certification. Read about this here.

- When there are groups of students, such as males and females, or regional groups, Lertap has options to look at group differences at two levels: test scores, and item responses. When there are just two groups, Lertap will open the door so that the possibility of differential item functioning (DIF) may be investigated.

- Cheating detection is found in Lertap's "RSA" routine (response similarity analysis).

- An inter-item correlation matrix may be created, and eigenvalues computed. Tetrachoric correlations are supported. Scree plots are easy. An estimate of McDonald's omega reliability coefficient is also available. More here.

- Data may be imported from ASCII text files, and exported to programs such as SPSS, SAS, Bilog-MG, Iteman, JASP, Xcalibre, and a selection of R packages.

- More: recode data; data sampling; IRT analyses; coefficient omega.

- I didn't rave on about other available resources, such as the main website at Lertap5.com, a spiffy website with sample data, some neat tips and tricks, the manual, and comprehensive online help. Lertap may be downloaded from this page.

- And there's our support desk, ready, waiting, ever eager: lertap5@gmail.com.