An introduction to Lertap 5.10

A glimpse of the key features found in Lertap for Excel 2010, 2013, & 2016

Larry Nelson, Curtin University, Australia

This presentation was last updated on: 22 April 2018.

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.

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 Lertap 5 typically looks in Excel 2010. The following slides will show how Lertap is used to get results.

The Data Worshsheet

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**.

The CCs worksheet
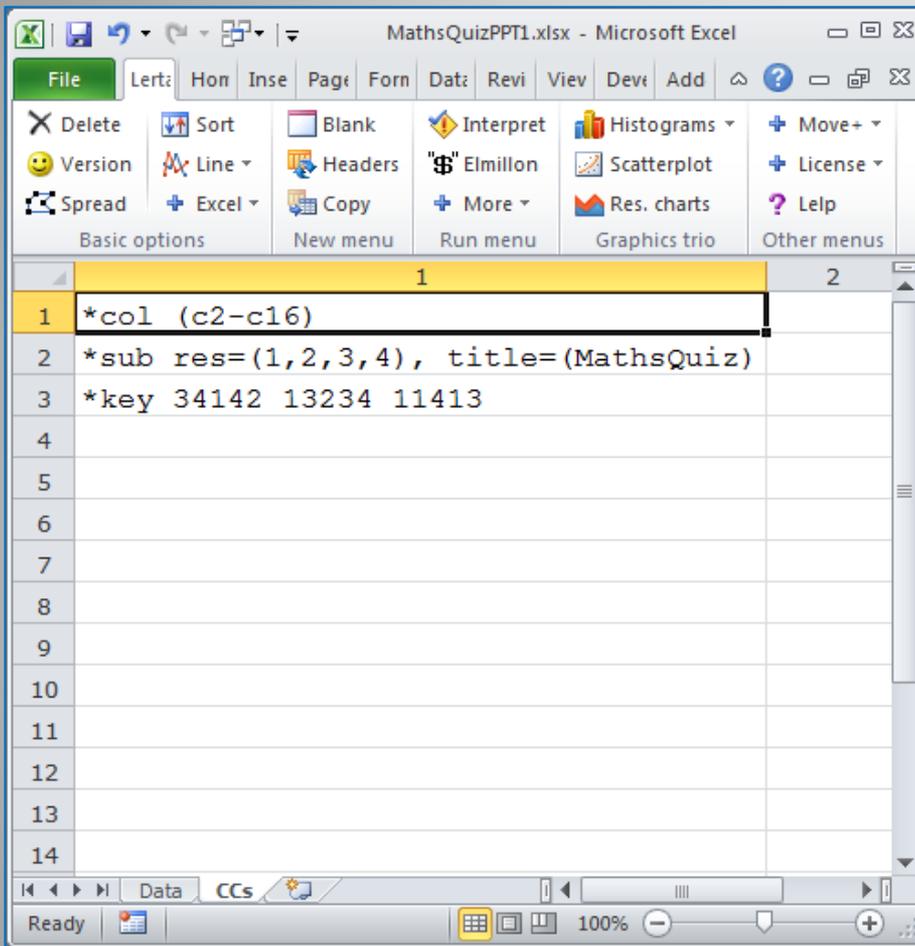
The CCs sheet is used to set up a Lertap analysis. CCs means "Control Cards".

This CCs sheet has three "cards": *col, *sub, and *key.

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

The res= control "word" on the *sub card says that this test's items use options 1, 2, 3, 4 .

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.

Read more about CCs cards here.

Having peeked at the Data and CCs sheets, we're set to get results. The 1st step: click on the "Interpret" option (next slide). NOTE: it's possible to download this dataset if you want to play along: get it here.

The Freqs "report"

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%

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 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.

The Freqs report gives you the chance to check on data quality. Once you're satisfied that the data are sufficiently free from errors, you'll move on to the "Elmillon" option to get statistics.

The Stats1b table

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.

The Stats1b plot
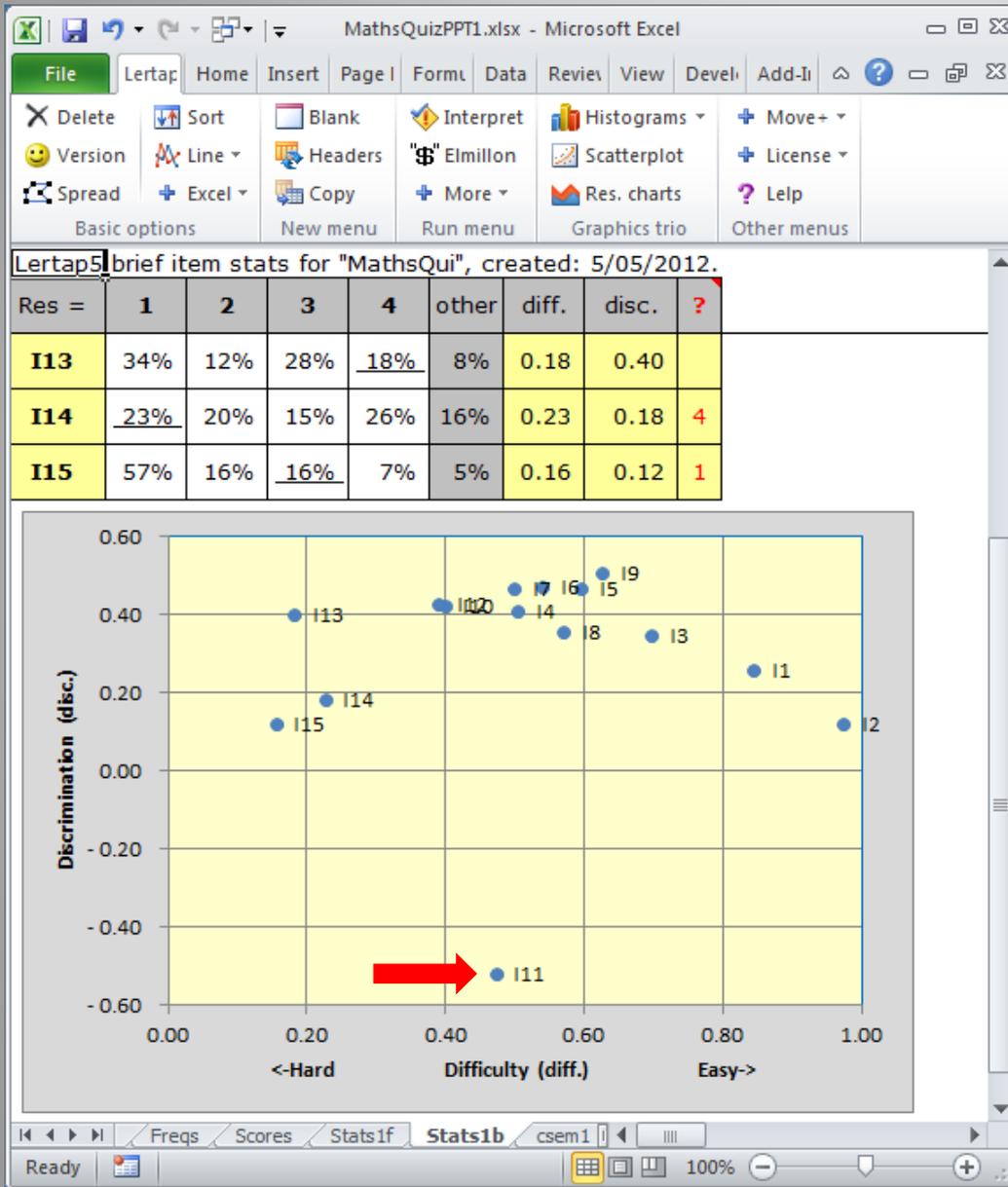
Lertap's Stats1b plot of item difficulty, "diff.", by discrimination, "disc.", is shown to the left.

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

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.

The Stats1f summary

"Stats1f" is another standard report made by Elmillon. The "f" stands for "full".

This report has the "Summary statistics" section shown here. The % figures correspond to the adjacent number, expressed as a percentage of the maximum possible test score (15.00). The median of 8.00, for example, corresponds to a percentage score of 53.3%.

As you can see, the reliability of this administration of the test was 0.68. This would be too low if the test was meant to let us pick out the best students with some confidence (in which case we'd probably want a reliability of at least 0.80).

The reliability of this test was low in part because one of the items, I11, was mis-keyed. Correcting this error will bring the reliability estimate up to around 0.80. More info here.

The Stats1ul report

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	<u>0.97</u>	0.00	0.01	**0.57**	**0.82**
2nd	0.07	0.03	<u>0.90</u>	0.00	0.01		
3rd	0.25	0.02	<u>0.69</u>	0.02	0.01		
4th	0.42	0.07	<u>0.40</u>	0.05	0.07		
lower	0.57	0.08	<u>0.16</u>	0.10	0.10		
I10 upper	0.01	0.14	0.02	<u>0.84</u>	0.00	**0.46**	**0.75**
2nd	0.04	0.33	0.05	<u>0.59</u>	0.00		
3rd	0.03	0.58	0.07	<u>0.30</u>	0.02		
4th	0.07	0.65	0.09	<u>0.19</u>	0.01		
lower	0.08	0.66	0.12	<u>0.09</u>	0.05		
I11 upper	<u>0.16</u>	0.00	0.01	0.84	0.00	**0.44**	**- 0.57**
2nd	<u>0.34</u>	0.01	0.03	0.63	0.00		
3rd	<u>0.54</u>	0.02	0.02	0.40	0.02		
4th	<u>0.61</u>	0.02	0.05	0.32	0.01		
lower	<u>0.72</u>	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

Stats1b csem1 **Stats1ul**

"Stats1ul" is also a standard report made by Elmillon. The "ul" stands for "upper-lower".

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, <u>97%</u> got this item correct, dropping to just <u>16%</u> 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.

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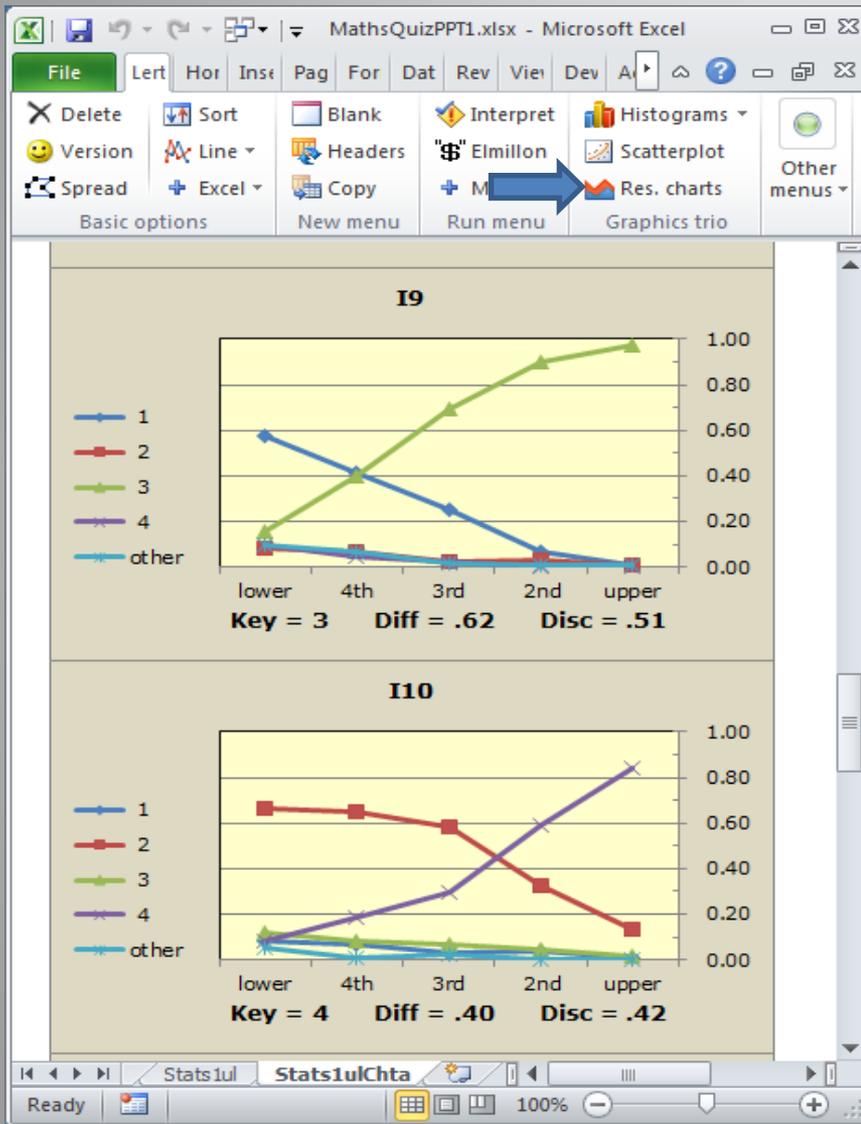
See where the blue arrow is in the picture to the left? It's pointing to the "Res. Charts" option (Response Charts).

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 snapshots is taken directly from Stats1ul results.

Items which are meant to pick out the best students should have graphs like these. In the lower group(s), 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](#).

Scores!
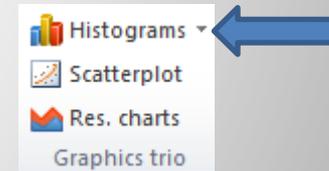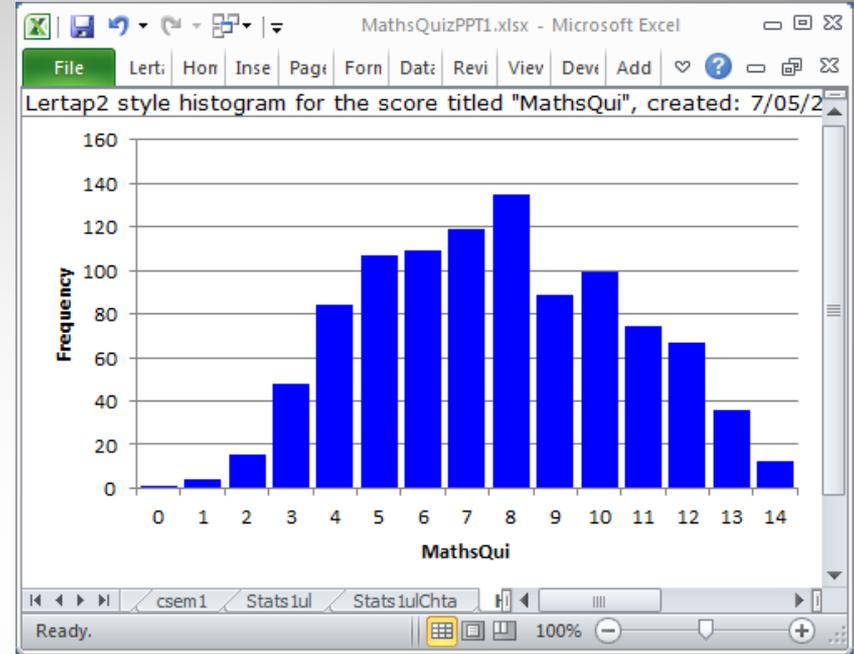
The "Scores" worksheet is another Emillon product. The histogram on the right was made by clicking on the "Histograms" option (natch!).

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.

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 "MathsQuizPPT.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.

- 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 (see it again).

- Next, I wowed you by getting some quintile plots. These provide one more way to see how the 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.)

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.

- These slides have looked at a cognitive test. Lertap will also process results from affective instruments, such as surveys and rating scales. Click here.

- When there are groups of students, such as males and females, or from various regions, 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 its first principal component extracted. More here. Bone up on coefficient omega, and using R with Lertap 5.

- Data may be imported from ASCII text files, and exported to programs such as SPSS, Bilog-MG, Iteman, and Xcalibre.

- More: recode data; formula scoring; data sampling; printing.

- I didn't rave on about other available resources, such as the main website at Lertap.com, a spiffy website with sample data, some neat tips and tricks, the manual, and comprehensive online help. The commercial outlet at Lertap.com is here.

- And there's me, larry@lertap.com.