Item Response Plots with Lertap 5.10.5
Larry Nelson
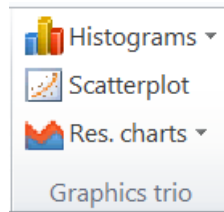
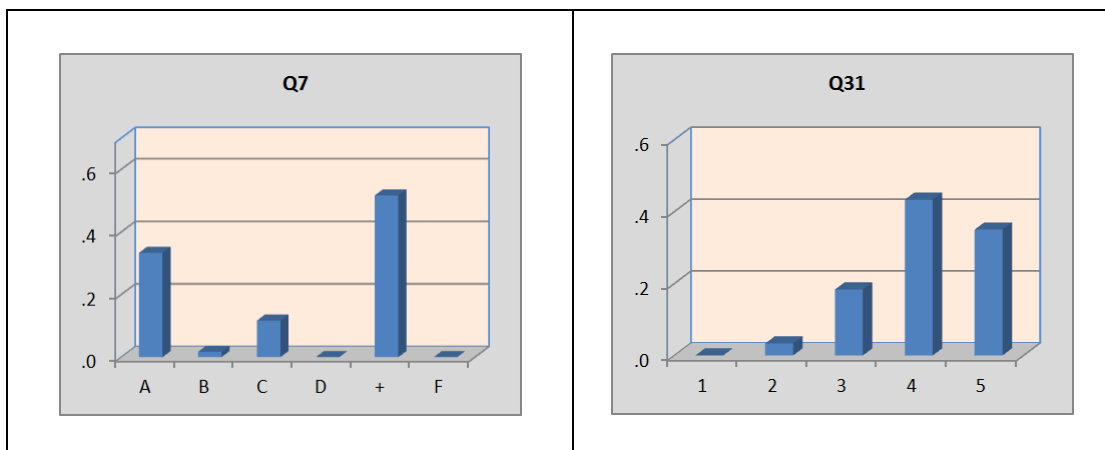--- Lertap.com ---

Document date: 1 March 2015

_____

Item response plots review

Lertap 5 emerged in the year 2001. It was a radical departure from previous versions, built from the ground up as an Excel "app". Earlier versions of Lertap, stretching back to 1973, were written in the Fortran and Visual Basic programming languages and operated as what we simply called "a computer program".

Excel, as most readers will know, has a powerful charting "engine" – it'll make some neat charts. There's an option on the Lertap 5 ribbon which is one entry into "response plots", and it has been available since the beginning of Lertap 5:
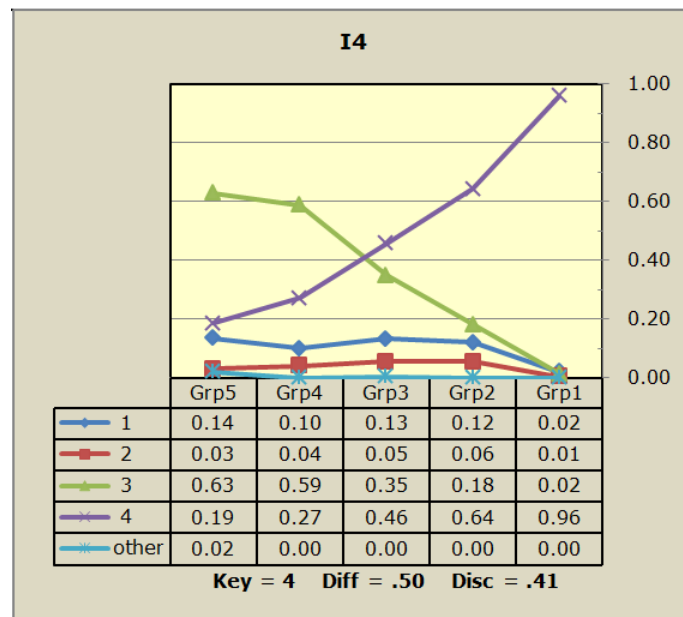

If I'm looking at a "Stats-b" summary from Lertap 5, such as a Stats1b report, and I click on the "Res. charts" option, I am rewarded with a simple chart summarising the relative popularity of each item option. Two examples are shown below; the chart on the left is from a cognitive item with seven options, of which the sixth was the keyed-correct response, represented as a + sign. The chart on the right is from a Likert question using the rather common 5-option scale, Strongly Disagree to Strongly Agree. More about these charts may be seen at this webpage.

In 2003 a "quantile plot" capability was added to Lertap 5 and seemed to very quickly gain popularity.

If I'm looking at a "Stats-ul" report, such as Stats1ul, and I click on the "Res. charts" option, charts such as the following will be produced:

I4

	Grp5	Grp4	Grp3	Grp2	Grp1
1	0.14	0.10	0.13	0.12	0.02
2	0.03	0.04	0.05	0.06	0.01
3	0.63	0.59	0.35	0.18	0.02
4	0.19	0.27	0.46	0.64	0.96
other	0.02	0.00	0.00	0.00	0.00

Key = 4   Diff = .50   Disc = .41

I generally refer to these charts as "plots". They concern themselves with cognitive test items, and trace, or plot, the relative performance of each item option over "groups" of students. In this example, there were five groups, from "Grp5" on the left, over to "Grp1" on the right.

In earlier versions of Lertap 5, the group on the left was called the "Lower" group as it is always comprised of the students with the lowest test scores, while the group on the right was referred to as the "Upper" group as its members are always the students with the top, or "upper", test scores. (The "ul" part of Stats1ul means "upper-lower".) When there are five groups, these plots are referred to as "quintile plots".

The example above indicates that the correct answer, the "Key", was 4. If I look at the 4 row in the little table below the plot (referred to as the "data table"), I see that 19%[1] of the students in the lowest group, Grp5, got this item correct. Then, staying on the 4 row and moving to the right, I see that as students become more capable, more "proficient" we might say, they tend to be more likely to be able to identify the right answer. In the strongest group, Grp1, 96% got I4 correct. (I4 is the title of the item in this example.)

The lines in the actual plot reflect what we see in the table. In this example, we see that the most popular distractor[2] was option 3. Sixty-three percent of the least-proficient students thought this was the correct answer, compared to just 2% of the students in the strongest group.

Much is said in the literature about having test items which allow us to "discrimi-nate" among the students, that is, items which provide us with the means to pick out the least- and most-capable students. We might then devise special revision materials or classes for the weakest, while patting the strongest on the back and opening the door for them to proceed to the next level of instruction.

---

[1] The table shows proportions. Multiple these by 100 to get a percentage figure.
[2] Wrong answers to multiple-choice questions are generally referred to as "distractors".

A discriminating item will have a plot like that seen above for I4. The trace line for the correct answer will be low on the left and then *climb* steadily as we move to the right, while the lines for the distractors will be high on the left and slope *down* as we move to the right.
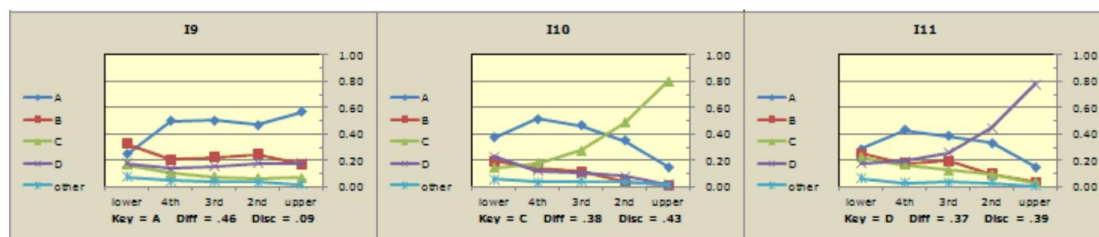
Several years ago I wrote a paper subtitled "Visual eye-tem analysis" in which it was suggested that these response plots can perhaps serve as a primary item evaluation tool for educators and test developers who may not have the statistical background useful for interpreting Lertap 5's Stats1f results. This paper has stood the test of time – it has been on the top-ten hit list almost since birth[3].

A reflection of the popularity of trace line plots is evidenced and exemplified today in Haladyna and Rodriguez (2013, pp. 345-355) who referred to them as " … *very effective for communicating item discrimination results* …" and "… *easy-to-read summaries* …". In 2010 my good friends at Assessment Systems Corporation saw to it that these plots were incorporated as a major tool in the fourth edition of their Iteman program.

(To be noted is that I am not claiming to have invented the idea of trace lines; indeed they date back many years. Lazarsfeld (1950) may have been the first to use the term *trace line*. I might suggest that Lertap 5 has popularised their use, and, if so, it's a reflection of advances in the power and graphical capabilities of contemporary computer hardware.)

In 2012 an option was added to Lertap 5 which makes it possible to get "packed plots". These provide an item performance gestalt. They can be used as an efficient way to quickly identify those test items that are not discriminating.

Here's a small example, extracted from a selected webpage:

Picture yourself with a large computer monitor or display – you're able to simultaneously see a virtual cornucopia of these packed response plots, as many as 40 or 50 of them at a time. Weak items, those not discriminating, will become most obvious. They'll stick out.

In this little example of packed plots, the item on the left, "I9", is punching below its weight as the trace for the correct answer (option A) does not rise as we want. The inability of the correct answer to complete its rise is due to the popularity of two of I9's distractors; I know it's a bit hard to see, but there are two distractor trace lines, the ones for options B (red) and D (purple), which indicate that these answers were regarded as best by a combined total of almost 50% of the students in the 2nd highest group, and about 40% of those in the top group.

The way the upper-lower groups are defined in Lertap 5 is usually based on a test score, and that score is most often just the number of test items a student got right.

---

[3] As measured by AWSTATS and Webalizer internet hit counters on Lertap websites.

In any quantiles analysis, Lertap gets Excel to sort all the student test scores from highest to lowest, storing them in a temporary internal array.

If we're after a quintiles breakdown, the top 20% of the array's scores, the "uppers", are marked as being in Grp1. The next 20% go into Grp2, and so on until we get to the bottom 20%, the "lowers", which will be referred to as Grp5, the bottom quintile. In a deciles analysis, there will be ten groups, each with 10% of the students. The top group will again be Grp1, the bottom Grp10.

It has always been possible to use another score as the basis for defining the groups. This is done by using Lertap 5's "external criterion" ("EC") option. Read on and you'll see that it is now possible to add more results to Lertap 5's quantile plots, including detailed EC statistical summaries.

Many changes, ring in the new (!)

As wondrous as the original Lertap 5 plots are, there have been requests from users for improvements. Could we make it possible to have more than five quantiles? Could we control for page breaks, so that the plots would not be spread over pages when printed? Would it be possible to change the number of "upper-lower" groups without having to open Lertap 5's System worksheet and alter the corresponding setting? Could options be added for controlling the size of each plot?
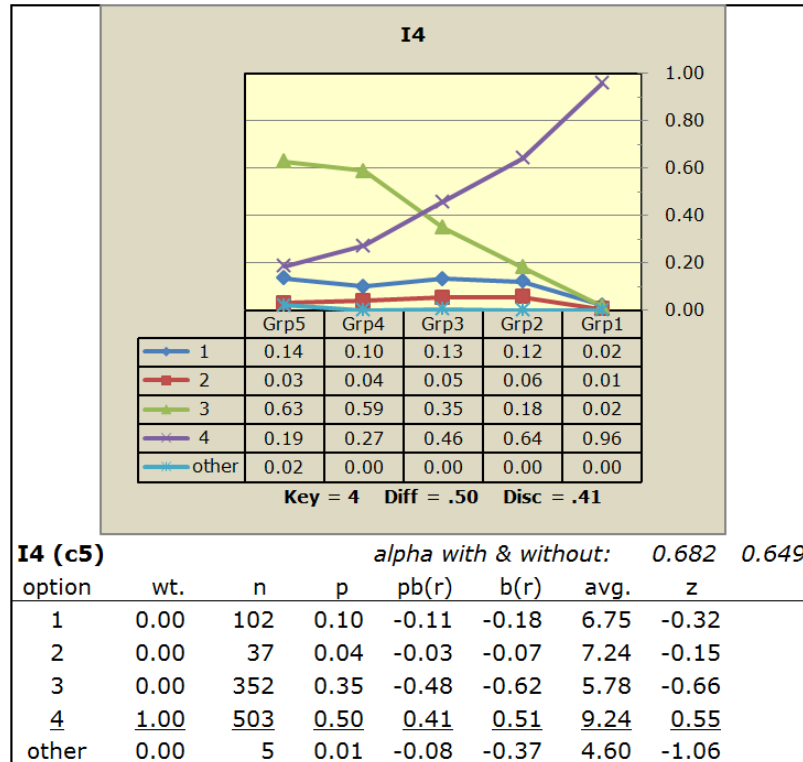
Could we perhaps add more information to the plots, such as the statistics found in Stats1f and ECStats1f reports?

Yes. Over a dozen options have been added to the System worksheet:

	The settings below are the standard ones for the Excel 2010 and Excel 2013 versions of Lertap.	Present setting:	Allowed settings:	Usual setting:
2				
80	**Additional settings for 'quantile' plots**			
81	Should the plots include Stats1f results?	yes	yes / no	yes
82	include coefficient alpha figures too?	yes	yes / no	yes
83	Should the plots include ECstats1f results (if they're available)?	yes	yes / no	yes
84	Number of worksheet rows for each plot	14	3 or more	14
85	Number of blank rows after each plot	0	0 or more	0
86	height of each row (should be an integer, no decimal point)	15	5 to 20	15
87	Number of worksheet columns for each plot	7	3 or more	7
88	column where plots start	3	2 or more	3
89	width of each column	5.5	5 to 15	5.5
90	Add a frame around the plots?	yes	yes/no	yes
91	Turn on the number of 'upper-lower' groups adjuster?	yes	yes/no	yes
92	Use Lertap's page break controller?	yes	yes/no	yes
93	Use custom page break setting?	no	yes/no	no
94	Set a page break at row number:	48	30 - 60	44
95	Use page margin settings in rows 64 - 69 above?	yes	yes/no	no

H ◀ ▶ H   Comments / Data / CCs **System** / Syntax / OldCCs

Let the plots thicken

Here's an example of a quantile which resulted from using the settings seen above[4]:

I4 (c5)				alpha with & without:			0.682	0.649
option	wt.	n	p	pb(r)	b(r)	avg.	z	
1	0.00	102	0.10	-0.11	-0.18	6.75	-0.32	
2	0.00	37	0.04	-0.03	-0.07	7.24	-0.15	
3	0.00	352	0.35	-0.48	-0.62	5.78	-0.66	
4	1.00	503	0.50	0.41	0.51	9.24	0.55	
other	0.00	5	0.01	-0.08	-0.37	4.60	-1.06	

The plot includes the Stats1f results for I4. In fact, the "plot" now involves information taken from all three of Lertap 5's standard statistical summaries. The shaded data table below the plot is taken from the Stats1ul summary. The line below the shaded table, with Key, Diff, and Disc, comes from the Stats1b summary, and many readers will probably recognize the I4 (c5) table as coming from the Stats1f summary[5].

The "*alpha with & without*" information also comes in from Stats1f. It quickly indicates if an item is correlating positively with the other items in the test. The more test items intercorrelate, the higher alpha will be, alpha being Lertap's primary indicator of test reliability[6].

Ordinarily, alpha will drop should an item be withdrawn from the test. If, on the other hand, alpha were to *increase* without the item, that would be an indication that the item was not contributing to test reliability and could perhaps be excluded from the test. In the case of I4, alpha would go from 0.682 down to 0.649 if it were removed from the test, suggesting that I4 is okay and should remain part of the test.

Next up is an item titled Q25615, an item included as one of ten trial items in a 50-item test. Trial items are sometimes included in a test to see how well they're

---

[4] Row 15 of the System sheet, was set to "Yes" so that the data table would be included.
[5] The (c5) refers to column 5 in the Data worksheet, where the responses to I4 are found.
[6] In the case of "mastery" tests, such as certification exams, alpha takes a back seat, replaced by Lertap 5's dependability and consistency estimates.

going to function; they're not scored, they don't contribute to a student's overall score (students usually do not know which items, if any, are trials).

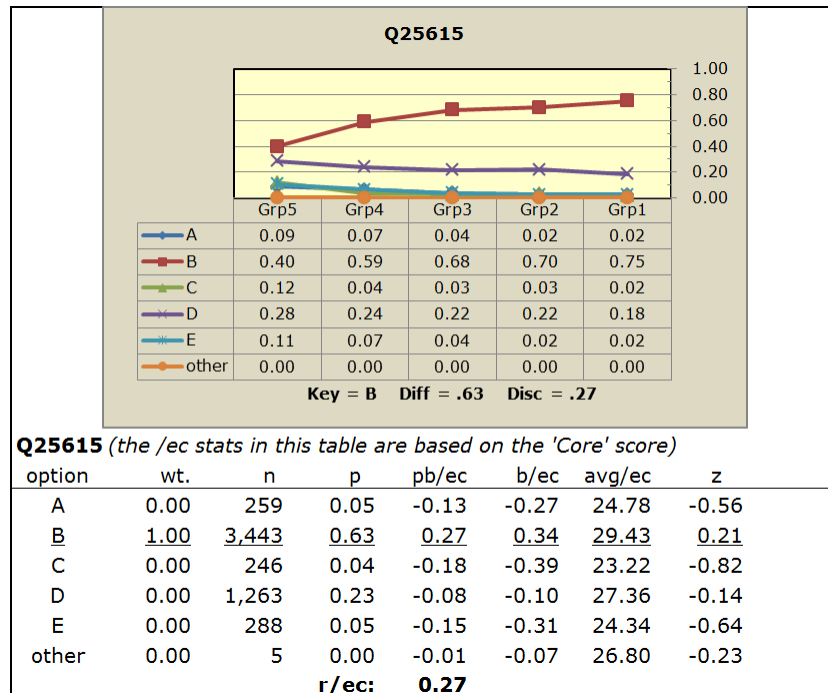The CCs lines used to process student responses are shown below[7]:

```
*col (c5-c54)
*sub res=(A,B,C,D,E), title=(Core), name=(Core items only)
*key ADAEE DDADA EBEDB BDCCE DAEAC EBCDB CDCBB DACAC EDBEE BDDCC
*exc (c10,c19,c20,c25,c27,c33,c35,c39,c43,c53)
*col (c10,c19,c20,c25,c27,c33,c35,c39,c43,c53)
*sub res=(A,B,C,D,E), title=(Trial), name=(Trial items only)
*key DBBDEDCBAC
```

The trial items have been lumped together as the second subtest in these lines. The title of this subtest is "Trial", and the name is "Trial items only".

There is no real interest in this so-called "subtest" as a true test of any sort. Nonetheless, Lertap will produce the usual statistical summaries much as if it were indeed the real deal, and I'll get Stats2f, Stats2b, and Stats2ul summaries. (The "2" in these names reflects the fact that "Trial" is the second subtest.)

What I'd like to see is how well the trial item titled "Q25615", a member of the second "subtest", correlates with the score from the 40 core items. To do this, I'll set up an external criterion analysis and let student scores on the first subtest, "Core", serve as the criterion for assessing the performance of each trial item, as found in the second subtest, "Trial". I will <u>not</u> be interested in including Stats2f results in my quantile plot, so I'll go the System sheet and set row 81 to "No" before clicking on the "Res. charts" option.

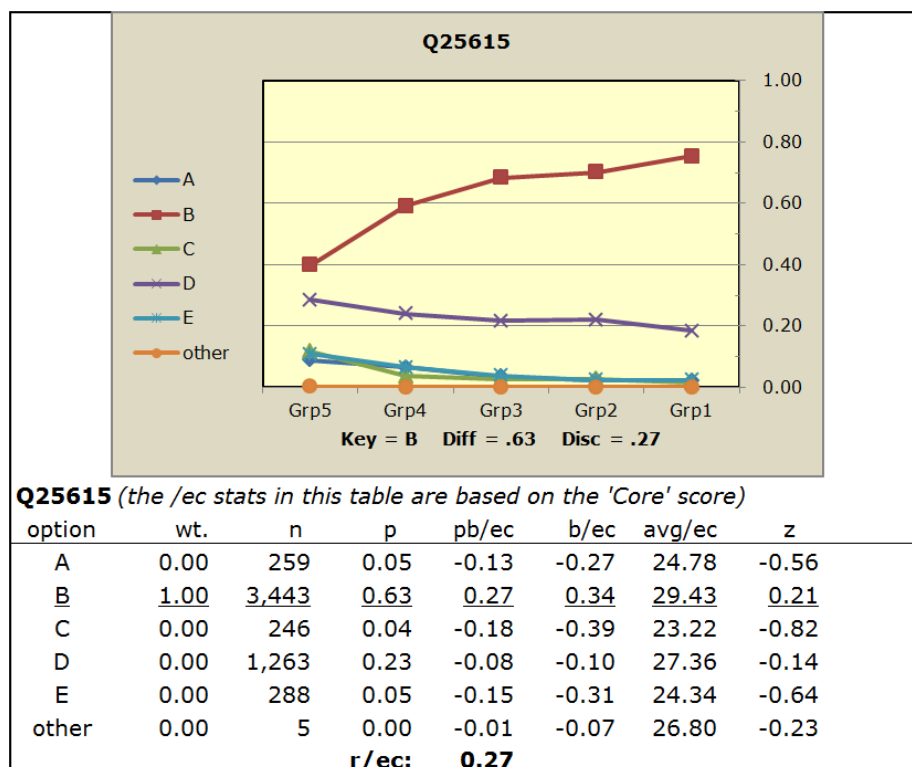Here is Lertap 5's corresponding quantile summary:

Q25615

	Grp5	Grp4	Grp3	Grp2	Grp1
A	0.09	0.07	0.04	0.02	0.02
B	0.40	0.59	0.68	0.70	0.75
C	0.12	0.04	0.03	0.03	0.02
D	0.28	0.24	0.22	0.22	0.18
E	0.11	0.07	0.04	0.02	0.02
other	0.00	0.00	0.00	0.00	0.00

Key = B   Diff = .63   Disc = .27

Q25615 *(the /ec stats in this table are based on the 'Core' score)*

option	wt.	n	p	pb/ec	b/ec	avg/ec	z
A	0.00	259	0.05	-0.13	-0.27	24.78	-0.56
B	1.00	3,443	0.63	0.27	0.34	29.43	0.21
C	0.00	246	0.04	-0.18	-0.39	23.22	-0.82
D	0.00	1,263	0.23	-0.08	-0.10	27.36	-0.14
E	0.00	288	0.05	-0.15	-0.31	24.34	-0.64
other	0.00	5	0.00	-0.01	-0.07	26.80	-0.23
			r/ec:	**0.27**			

---

[7] This example is based on a real testing situation; more information is here.

The Key and Diff values seen above have been picked up from the Stats2b summary, while the Disc value is now equal to the correlation between the item and the external criterion, that is, "**r/ec**". ("Diff" means item difficulty, while "Disc" stand for item discrimination.)
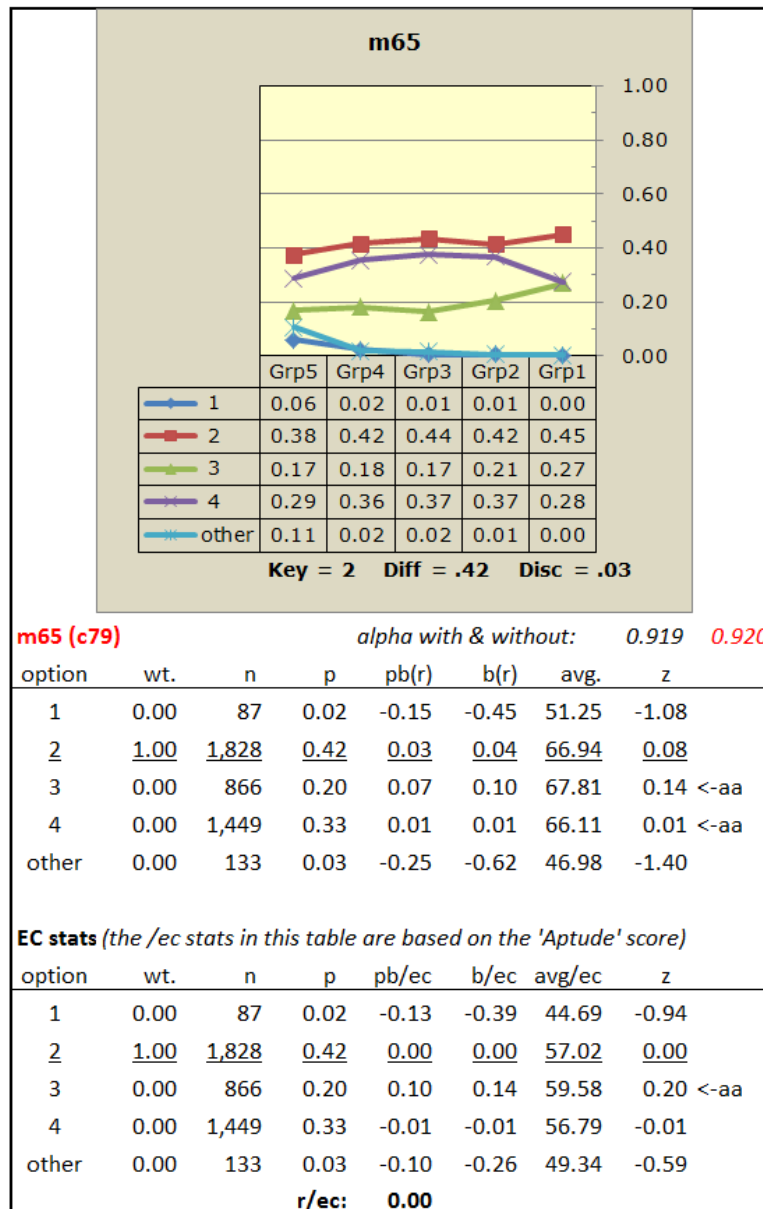
The five groups have now been defined on the basis of the "Core" test score. The item statistics for Q25615, to the right of the "p" column, are based on using the external criterion (ec) score, which in this case comes from the forty "Core" items.

The keyed-correct answer to Q25615 was B. How did the item perform? Well, the trace line for B rises rather slowly, but steadily. It starts out a bit high on the left: 40% of the weakest students got the item right. And, over on the right, distractor D has kept it from going above 0.80 in the top group. I also note that the 3,443 students who chose option B had the highest average Core score (avg/ec = 29.43). And, best of all perhaps, I find Q25615 to have a fairly healthy correlation with the criterion measure (r/ec = 0.27). I would be inclined to say that Q25615 has perhaps passed the trial, but still I'd ask the item writer to review it with an eye to possibly revising distractor D – we don't want distractors to be pulling in 18% of the strongest students.

The same plot is seen below without the data table – whether or not this table appears may be controlled in two ways: by setting row 15 of the System sheet to "Yes" if it's wanted, otherwise to "No". It may also be toggled on and off by using a menu option.

Q25615 (the /ec stats in this table are based on the 'Core' score)

option	wt.	n	p	pb/ec	b/ec	avg/ec	z
A	0.00	259	0.05	-0.13	-0.27	24.78	-0.56
B	1.00	3,443	0.63	0.27	0.34	29.43	0.21
C	0.00	246	0.04	-0.18	-0.39	23.22	-0.82
D	0.00	1,263	0.23	-0.08	-0.10	27.36	-0.14
E	0.00	288	0.05	-0.15	-0.31	24.34	-0.64
other	0.00	5	0.00	-0.01	-0.07	26.80	-0.23
			r/ec:	0.27			

Okay, let me now bring in one more example, "**m65**", a member of a 94-item high school mathematics test developed by a very experienced national testing service. A maths and science aptitude test had also been created.

m65

		Grp5	Grp4	Grp3	Grp2	Grp1
	1	0.06	0.02	0.01	0.01	0.00
	2	0.38	0.42	0.44	0.42	0.45
	3	0.17	0.18	0.17	0.21	0.27
	4	0.29	0.36	0.37	0.37	0.28
	other	0.11	0.02	0.02	0.01	0.00

Key = 2   Diff = .42   Disc = .03

m65 (c79) alpha with & without: **0.919** *0.920*

option	wt.	n	p	pb(r)	b(r)	avg.	z	
1	0.00	87	0.02	-0.15	-0.45	51.25	-1.08	
2	1.00	1,828	0.42	0.03	0.04	66.94	0.08	
3	0.00	866	0.20	0.07	0.10	67.81	0.14	<-aa
4	0.00	1,449	0.33	0.01	0.01	66.11	0.01	<-aa
other	0.00	133	0.03	-0.25	-0.62	46.98	-1.40	

EC stats *(the /ec stats in this table are based on the 'Aptude' score)*

option	wt.	n	p	pb/ec	b/ec	avg/ec	z	
1	0.00	87	0.02	-0.13	-0.39	44.69	-0.94	
2	1.00	1,828	0.42	0.00	0.00	57.02	0.00	
3	0.00	866	0.20	0.10	0.14	59.58	0.20	<-aa
4	0.00	1,449	0.33	-0.01	-0.01	56.79	-0.01	
other	0.00	133	0.03	-0.10	-0.26	49.34	-0.59	
			r/ec:	**0.00**				

The maths test was quite a good one. I looked hard to find poorly-performing items and found just four, with m65 being one.

The response proportions in the plot are based on the test's Stats1ul summary. The Key Diff Disc data are from the Stats1b summary. The m65 (c79) table is based on Stats1f. I ran an "ec" (external criterion) analysis using student scores on the aptitude test, which was titled "Aptude[8]". The EC stats results are from the ECStats1f summary.

Note that I could have based the plot's response proportions on the ECStats1ul summary instead of Stats1ul[9]. The only difference would be that the trace lines in the plot would be based on ECStats1ul response proportions instead of those in Stats1ul.

---

[8] Most of the titles used for things in Lertap are constrained to use no more than 8 characters.
[9] The external criterion option results in the creation of ECStats-f and ECStats-ul summaries.

Was m65 a good item? Not particularly. The trace line for the correct answer, 2, does not rise adequately. The "avg." test score for the 866 students selecting option 3, a distractor, was above the test mean[10]. More than a thousand students chose another distractor, 4, and their avg. was also above the mean. Not good; we want avg. for the distractors to be below the test mean – when they're not, Lertap adds the little "aa" flags seen ("aa" = "above the overall test average").

The item has poor discrimination, Disc = .03. Its correlation with the aptitude score was zero (r/ec), and the average criterion score for the 866 selecting option 3 was above the mean Aptude score of 56.97.

m65 would not seem to warrant further use. What would happen to test reliability, as measured by alpha, if m65 were omitted? It would go up but just slightly, from 0.919 to 0.920. There were 93 other items and 90 of them performed very well, carrying alpha to a nice high. But I would suggest that m65 be taken "off line", and I'd also recommend that the students who selected distractor 3 or 4 might get some credit. (This is done by using *mws lines in the CCs sheet. An example is given here.)

Experiment

If the new quantiles don't look picture perfect to you, you'll likely want to fool with the System settings introduced above, and discussed in detail below.

I suspect a common reaction might have to do with the height of the trace line plot area. Either the plots will look too scrunched or too stretched. The main setting controlling this is found in row 84. Keep in mind that it's possible to first experiment with the height manually, before changing row 84. Just select the chart and tug at its handles.

To do this in an optimal manner, go to the "Basic options" section of the Lertap 5 ribbon, as mentioned in this topic, and then click on the "+ Excel" option as shown herewith:

Turn on the "Gridlines" option, and maybe also "Headings". Then, as you tug at the chart, keep track of the number of worksheet rows it occupies. When it seems to be what you want, then change the row 84 setting.

See the "Page breaks" option? It's new. If you opt to use the "page break controller" (System row 92), you're very likely to notice dotted lines on worksheet pages. Use this flashy new option to extinguish them, or, if you want to quickly see where page breaks are going to be, use this option to see 'em.

---

[10] Which was 65.91.

More comments on the new options

The row numbers which appear below refer to rows in the Lertap 5 System worksheet.

Row 81 Should the plots include Stats1f results?

> "Stats1f" refers to the full statistics for "subtest 1", the first subtest mentioned in the CCs worksheet. When the CCs worksheet defines more than one subtest, there will also be a "Stats2f" worksheet with results, and maybe "Stats3f" too, depending on the number of subtests.
>
> Row 81 means to refer to any of these Stats-f worksheets, not just Stats1f.
>
> Refer to this topic about subtests if you'd like to.

Row 82 include coefficient alpha figures too?

> Alpha is Lertap 5's main index of test reliability. It's found in the second section of a Stats-f report, as mentioned here. (When the "Mastery" option is in effect, alpha often takes a distant back seat to the dependability and consistency measures which will appear at the bottom of Stats1ul reports.)
>
> When this option is set to "Yes", the "*alpha with and without*" information will be displayed for each item (it can be seen in the examples above).
>
> This information can be used as a guide to how well an item is "hanging in there", that is, how it's correlating with the other items in the subtest. We'd usually want a subtest's alpha figure to be higher with the item included in the subtest – if its alpha is actually higher *without* the item then we'd have some reason to think that perhaps the item should be excluded from the subtest, or, perhaps, be examined for ambiguity and possibly double-keyed[11]. (Refer to this paper for a related discussion.)

Row 83 Should the plots include ECstats1f results (if they're available)?

> "ECstats" are produced when the "Use external criterion" option is selected from Lertap 5's "Run menu".

Row 84 Number of worksheet rows for each plot

> This setting may take you into the "Experiment" discussion above.

Row 85 Number of blank rows after each plot

> These will appear at the very bottom of a quantile plot, just above the bottom frame line. None of the examples above have blank rows. These rows might be used to include comments.

Row 86 height of each row (should be an integer, no decimal point)

---

[11] There are times when alpha is not of primary concern – in mastery and certification tests for example, in formative testing, and when a "test" is simply a trial run with only a few items.

This is usually measured in "points", there being 72 points to an inch. If you're thinking about changing the height of the plot area, I strongly suggest using row 84's setting instead of this one.

Row 87 Number of worksheet columns for each plot

If either the Stats1f setting (row 81) or the ECStats1f setting is used, then the number of columns should probably not be greater than 9. This is because the little tables of statistics always have 9 columns, from "Option" on the left, over to the flags column to the right of "z". I think the best setting for this is 7 as that will centre the plot above these 9 columns. (All of the examples above have 7 for this setting. Here I might mention that setting this to 9, and the following setting in row 88 to 2, will give a plot which spans the frame – useful when the number of upper-lower groups is greater than 5.)

Row 88 column where plots start

If Stats1f or ECstats1f results have been selected (Yes in row 81 or row 83), then I suggest 3 or 4 for this number.

Lertap 5's quantile plots have always used the first column as a type of margin, a small empty "buffer" to the left of the charts (thought to make them easier to interpret), so this number should never be less than 2.

Row 89 width of each column

This is usually measured in "number of characters in a cell". I have experimented with a variety of settings from 5 to 7, and personally recommend 5.5; 6 is not a bad setting either.

Row 90 Add a frame around the plots?

This option applies only when Stats1f or ECstats1f results have been selected (Yes in row 81 or row 83 or both). The samples above all have these frames.

Row 91 Turn on the number of 'upper-lower' groups adjuster?

The number of "upper-lower" groups is set in row 13 of the System worksheet. In previous versions of Lertap 5, the maximum number of these groups was limited to five. Now it's gone up to a maximum of 10, and the groups have been renamed; they used to be named Upper, 2nd, 3rd, 4th, and Lower. Now they're Grp1 (instead of "Upper"), Grp2, and so on.

In this new version of Lertap 5, whenever the "Elmillon" option is taken, then, if this option is set to Yes, Lertap looks to see if the groups will have fewer than 20 people (given the present setting in row 13 of System). If so, then a little dialog box appears, asking if the number of groups should be temporarily reduced. Use this option to stop Lertap from asking this question, that is, set it to "No".

Row 92 Use Lertap's page break controller?

This can be set to Yes if the results are to be printed – the "controller" will strive to insure that an item's plot and statistics are not split over pages.

Note: *using this option slows things down*. With Excel 2010, I have found that it can more than double the time required to produce the quantiles. For example, it took about three minutes for Lertap and Excel 2010 to produce the quantiles shown in the "m65" example above.
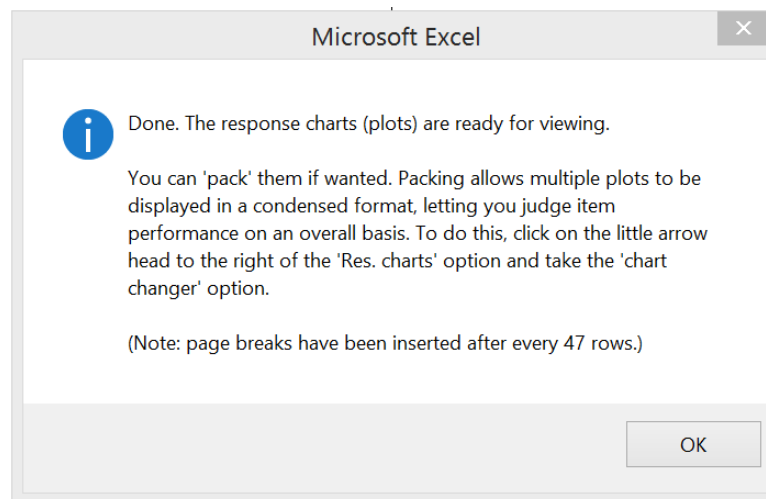
I have also found that Excel is not always obliging when it comes to adjusting its default page breaks, which is what the page break controller tries to do. We've put in a little error catch to let you know when Excel throws a page break fit; if the error message appears, you'll see that it suggests changing page margins in the corresponding "ul" sheet (more about this below, under "Printing").

Row 93 Use custom page break setting?

(This setting only applies when row 92 is set to "No".)

Lertap's page break controller can be *very* slow when a subtest has many items. Rows 93 and 94 provide a way to speed things up.

This message appears just before response plots are displayed:

In this example, the "Note" indicates that the controller has inserted a page break after every 47 rows. Armed with this information, I can then turn off Lertap's page break controller by putting "No" in row 92, "Yes" in row 93, and then 47 in row 94.

As long as I continue to use the same printer and paper size, it will be fairly safe to continue to use these settings, and things will run noticeably faster when a test has many items.

Row 94 Set a page break at row number:

This setting will be used only when row 92 is set to "No" and row 93 is set to "Yes".

A setting of 44 might be about right for printing to "Letter" page size and using Excel's default page margins, while 48 might be okay if using "A4" page size.

Row 95   Use page margin settings in rows 64 – 69 above?

Here are the rows referenced:

63	Page size measurement (inches or centimetres)	in	in / cm
64	Header margin	0.00	
65	Top margin (suggest 0.50 inches or 1.50 cm)	0.50	
66	Bottom margin (suggest 0.50 inches or 1.50 cm)	0.50	
67	Footer margin	0.00	
68	Left margin (suggest 0.50 inches or 1.50 cm)	0.50	
69	Right margin (suggest 0.50 inches or 1.50 cm)	0.50	

When this option is set to "Yes", Lertap and Excel look at what's in the yellow box, row 63. If "in" is found there, then, with the settings above, the top page margin will be set to half an inch (0.50). If "cm" is found in the yellow box, then the top page margin will be half a centimetre.

Printing

Training this new version of Lertap 5 to have regard for page breaks so that quantile plots will not be split in the middle, half on one page, half on the next, seems to be working well most of the time for me (but not always). Feedback from users will be useful.

The first thing to do, print-wise, is to set page margins if you're not happy with Excel's default margins. There are two ways this may be done:

Set row 95 = "Yes", and then put wanted margins in rows 64 – 69.

Set row 95 = "No" and the page margins will be those Excel finds in the Stats-ul or ECStats-ul sheet. Usually these will be the standard workbook-wide page margins applied to all the book's worksheets. However, you can get into the Stats-ul or ECStats-ul sheet and change them; whatever you change will then be used as the page margins for the quantiles.

Comments: Excel picks up default page margins from a special workbook called "Book.xltx", housed in your computer's "XLSTART" folder. The XLSTART folder is not always easy to find as, in some versions of Windows, it's generally hidden away as a system file. But if you can find it, open Book.xlxt and change the default page margins; from then on every *new* workbook will use them. Or, an alternative: in any or your own workbooks, right-click on any worksheet tab at the bottom of the screen (such as, perhaps, "Stats1ul"). Then click on "Select All Sheets". Next, change the page margins to whatever you want – your selections will then apply to all worksheets in that workbook, including any new ones Lertap might generate.

Okay, I've got my margins set, and I want to print those dandy quantiles. Before doing so, I'll make sure I have Lertap's page break controller activated, some-thing I do by putting "Yes" in row 92. Then I print as per usual.

Problem: the page break controller is slow, slow, running to very slow if there are many quantiles. So, at this stage I might take note of what the page controller adds to a message box just as the quantiles are about to appear – it's shown above under the row 93 discussion. I then might turn the page controller off by

putting "No" in row 92 and following those steps introduced above for rows 93 and 94.

Now, however I happen to be operating, margin-wise, more than half the time I'll find Excel and Lertap have only put one quantile on a page, and yet there's so much empty space below it, surely enough for two quantiles per page I say to myself. So, what I do is delete the quantiles page (which might be named Stats1ulChta, for example).

Then I go over to Stats1ul and change its page margins, trying to squeeze out more space. Okay. Back to the "Res.charts" option as per usual to get my new quantiles in Stats1ulChta. Now do I have two on a page? If not I repeat this process, with fingers crossed.

There's a bit more stuff on printing Lertap 5 reports at this webpage. And, of course you don't want to forget Excel's stock standard printing helpers: Page Layout view, and Page Break Preview. Page Layout view provides the opportunity to add page headers and footers, much as in Word. Naturally this assumes that you have not scrunched the top and bottom page margins so much as to pinch the space normally allocated for headers and footers.

A caveat: just when test runs were going well, I came across a problem related to the "m65" analysis featured above. When taking the "Res. charts" option while looking at a Stats1ul summary, and with the page break controller on, Excel spat the dummy (as they say in New Zealand and Australia): it threw up a fatal error. We *think* we know where the problem lies: if Excel's default page break for the last page aligns with the last row of the corresponding spreadsheet, up goes the error, down comes Excel.

We've inserted a catch for this error. Now it's not fatal. A work-around is to change the page margins in the corresponding "ul" sheet and then try again. This has worked for me. Another way to combat this problem would be to use System rows 93 and 94, experimenting with the page break row number until things turn out okay (remember to set row 92 to "No" first).
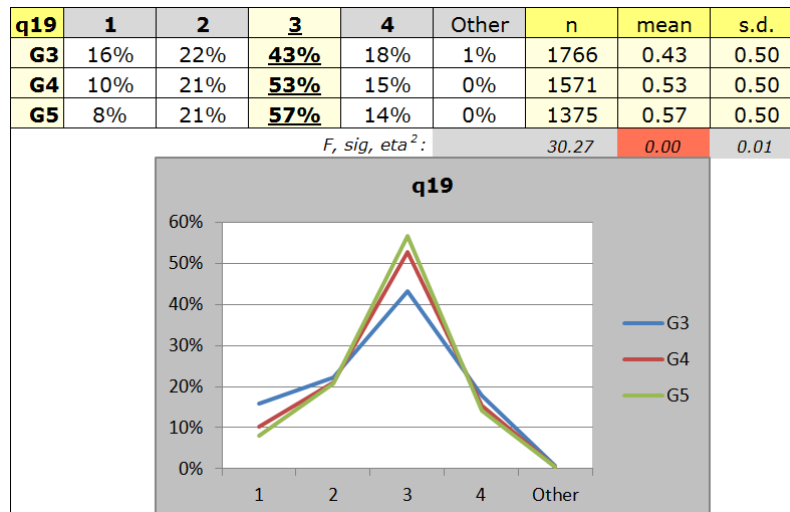
Stats1b charts have also changed

My examples have all had to do with quantiles. But remember: the "Res. charts" option also comes into action if you click on it while looking at a Stats-b summary. In fact I mentioned this back on the first page.

Some of the new options also apply in the Stats-b case. The settings in System rows 92, 93, and 94 will apply, and the page margins in the Stats-b worksheet will carry over to the charts sheet in the same manner that Stats-ul page margins control the margins on the quantile charts.

Ibreaks charts have not been left out

The Ibreaks option is a third Lertap 5 option capable of producing an item response chart for each test item. In this case, the response charts will show how groups of test takers have responded to an item's options. An example is shown below:

q19	1	2	**3**	4	Other	n	mean	s.d.
G3	16%	22%	**43%**	18%	1%	1766	0.43	0.50
G4	10%	21%	**53%**	15%	0%	1571	0.53	0.50
G5	8%	21%	**57%**	14%	0%	1375	0.57	0.50
				F, sig, eta² :		30.27	0.00	0.01

The table and plot above have to do with a four-option multiple choice item answered by children in three primary school grades. The correct answer, the "key", was option 3. It's obvious that Grade 3 students did not do as well as the other two grades on item q19.

Quite a few of the options discussed above will apply to the production of Ibreaks summaries. These are the options found in rows 64-69, 86, 89, and 92-95.

Note: the Ibreaks options presents a number of questions when it starts up. One of them asks if response charts should be created. *It's only when the answer to this question is "Yes" that these options come in to play.*

An Ibreaks run will always result in a new worksheet, a new summary, called Ibreaks1 (or Ibreaks2 … Ibreaks-x). When there are only two groups, Lertap 5's DIF analysis option may be activated, resulting in the output of Mantel-Haenszel statistics in another summary sheet called IbreaksMH1. None of the new options apply to IbreaksMH summaries.

Back to the future?

Use the settings seen below if you have no interest in the flashy new capabilities discussed above. These settings are the ones used for the quintile plots in former Lertap versions:

80	Additional settings for 'quantile' plots	Present setting:
81	Should the plots include Stats1f results?	no
82	include coefficient alpha figures too?	no
83	Should the plots include ECstats1f results (if they're available)?	no
84	Number of worksheet rows for each plot	9
85	Number of blank rows after each plot	1
86	height of each row (should be an integer, no decimal point)	24
87	Number of worksheet columns for each plot	5
88	column where plots start	2
89	width of each column	8.25
90	Add a frame around the plots?	no
91	Turn on the number of 'upper-lower' groups adjuster?	yes
92	Use Lertap's page break controller?	no
93	Use custom page break setting?	no
94	Set a page break at row number:	48

Packed plots implications
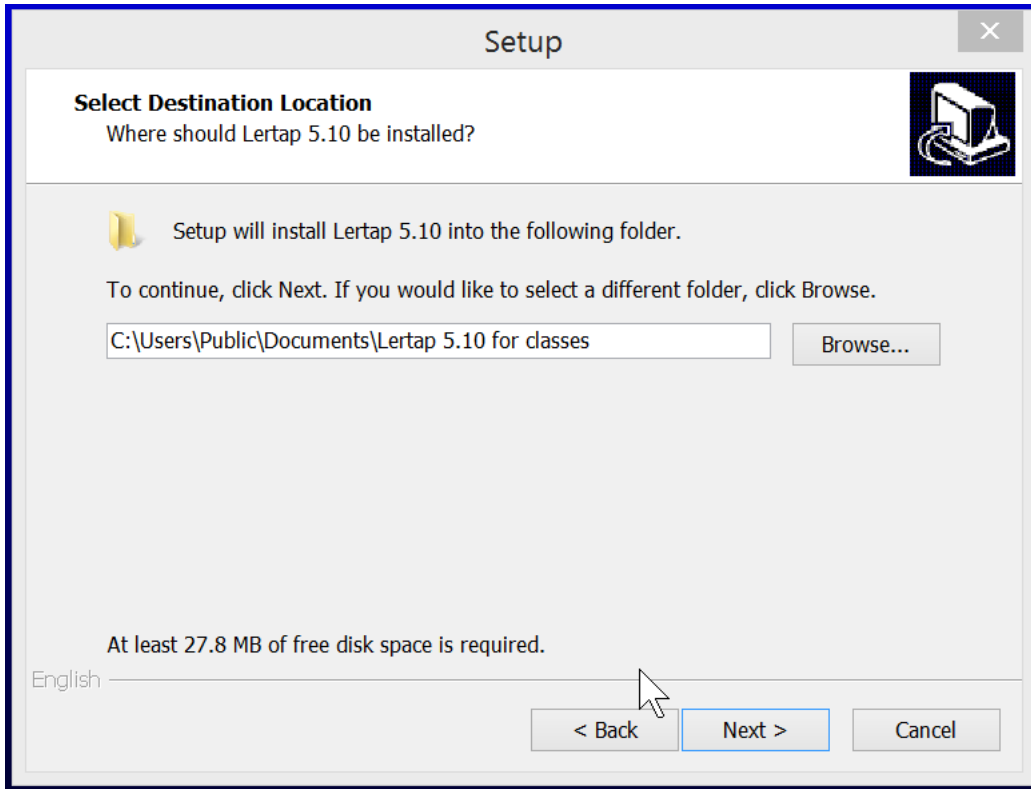
If the data table is displayed in the plot, the packed plots will have the tables too. This is not to my liking; I like "clean" packed plots, no data tables, thank you -- so I've seen to it that there's now a toggle which will turn the data tables on or off at any time. I might start up with them on, using a "Yes" setting in System row 15, but then toggle them off before going for the packed plots.
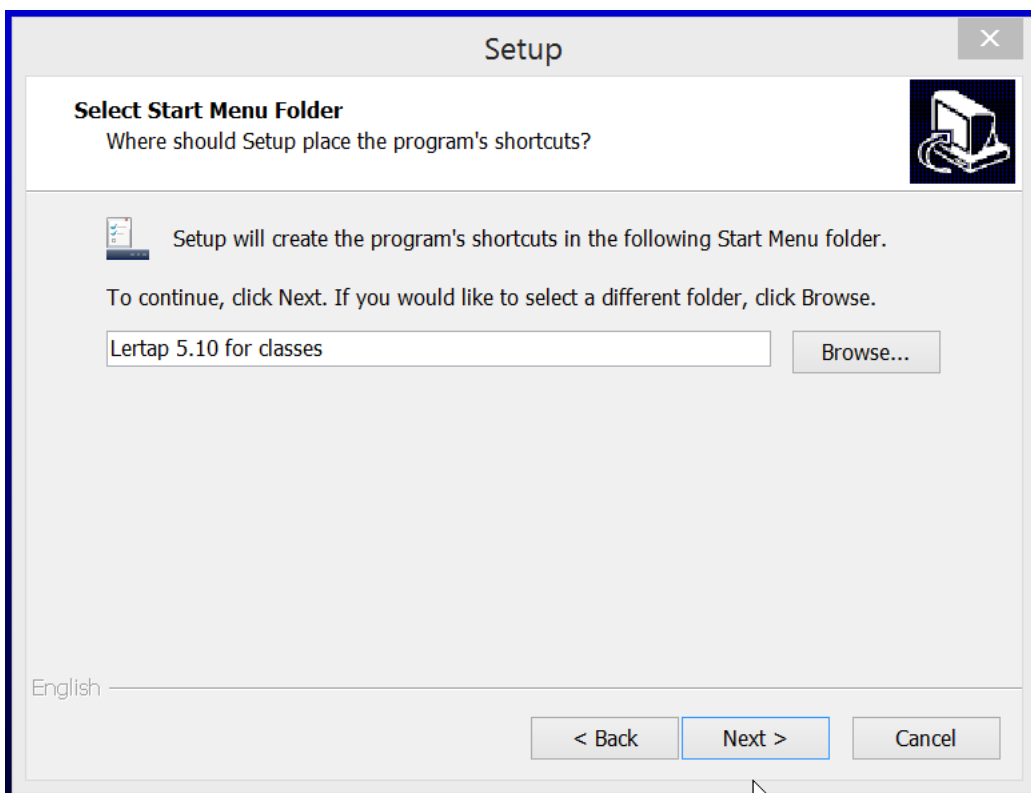
Multiple configurations

It's possible, and, in my opinion, not excessively difficult, to have multiple installations of Lertap 5 on the same computer. The advantage in this is that each installation may have its own System settings. I'll show how it's done.

When the Lertap installer runs it poses a number of questions. The first crucial one has to do with where Lertap is to be installed.

In the screen snapshot below, I'm shown setting up Lertap in a new folder which will be called "Lertap 5.10 for classes". In this case, I'm going to place this folder within C:\Users\Public\Documents; most users would probably put it in something like C:\Users\Documents.

The installer then asks where it is to place Lertap's shortcuts. I choose "Lertap 5.10 for classes".
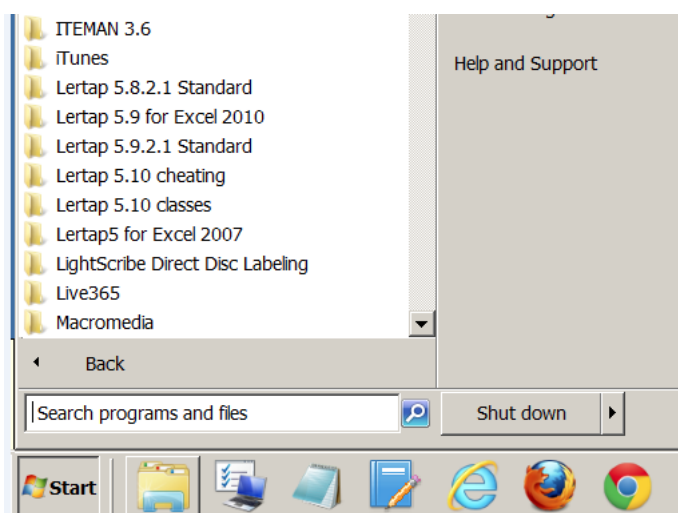

Once I've set up Lertap in this manner, I'll open it and change the settings in the System worksheet so as to meet the features I generally use with classes, such as "*Use experimental features*", "*IStats to include item intercorrelations, and*

eigenvalues", and "*Should Xcalibre files be created*". I will also turn off the page controller so that Lertap runs faster when it comes to creating response plots.

Having set up a copy of Lertap for use with classes, I will then run the installer again and set up another folder for Lertap, which could be, for example, "Lertap 5.10 for cheat detection", with shortcuts going into "Lertap 5.10 for cheating detection". (I don't have to use the same folder names, but prefer to do so as, to me, it makes things a bit clearer.)

Then I'll open this copy of Lertap and change System rows 25 through 33 to suit the cheat checker's settings.

The next screenshot shows how multiple Lertap installations look in a Windows 7 All-Programs list (my Windows 7 computer is set up to run many configurations):

Each Lertap 5 installation requires its own unlock code. These may be obtained by writing to support@lertap.com – when you write, be sure to mention that you're setting up multiple installations on the same computer, as described in this document.

How to get it

The new version[12] may be downloaded by visiting this website. If you already have Lertap 5 installed, you have the choice of uninstalling the former version before installing this new version, or you might install the new version and keep the old one using the just-discussed "Multiple Configurations" idea.

Here's a good idea: read the installation suggestions. Every new install seats the so-called "Mini" version with its in-built limit of 50 data records. Upgrading from Mini to a full version, capable of processing thousands of records, requires a license. Licenses may be purchased at The Store[13]. Users already holding a license are invited to look here for a special upgrade deal.

Note: "unlocking" Lertap 5 refers to the process of applying the "unlock code" that's e-mailed to license purchasers. Once a license has been obtained, Lertap 5 is activated ("unlocked") by following these steps.

---

[12] This version requires Excel 2010 or 2013. It will not run on a Macintosh computer.
[13] January 2015 saw the introduction of new license prices. Most are now below $100.

References

Haladyna, T.M. & **Rodriguez**, M.C.(2013). *Developing and validating test items.* New York: Routledge.

Lazarsfeld, P.F. (1950). The logical and mathematical foundation of latent structure analysis. In SA Stouffer et al. (Eds.), *Measurement and Prediction*. Princeton, NJ: Princeton University Press, pp. 362-412.

Larry Nelson (support@lertap.com)