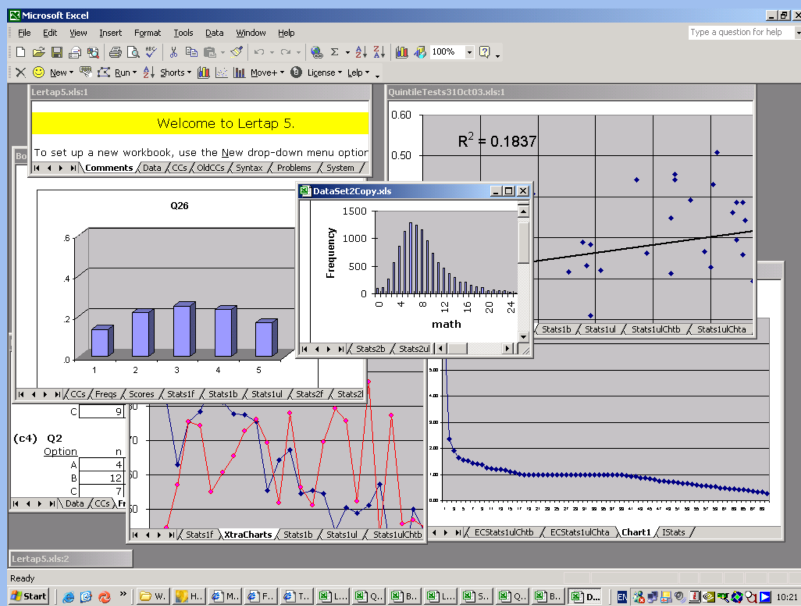


"Lelp"

Lertap 5 help

Interactive PDF version

Made in Western Australia

Table of Contents

1 Welcome

Welcome to Lertap5 help, "Lelp".

Lelp (this document) is designed to be used in conjunction with the Lertap5 item, test, and survey analysis system for Windows and Macintosh computers.

Lelp exists in a variety of formats: as a PDF file, as a CHM file, as an e-book for use with an iPad, and as a website. Paths to all of these are at Link (1) below.

The website is always the most up to date version of *Lelp*. To see the date of the version you're enjoying right now, be it PDF, CHM, or website, look at the bottom of this topic, below the blue line, next to 'Last update:'.

Please note: some of the screen snapshots found in this document show the old Lertap toolbar at the top instead of the Excel Lertap5 tab captured above.

Links to a variety of Lertap resources are listed below.

1 A PDF copy of this website's topics. A CHM copy (compiled help file for Windows). An epub copy, ready for reading on an iPad or an iPhone. A link to the website itself.

2 A small set of PowerPoint slides with a quick introduction to Lertap5. These are also available as a PDF file.

3 The main Lertap5 website. Has more examples and samples, with links to videos, the manual, and a variety of *riveting* technical papers (also known as "erudite epistles").

4 Sample datasets for downloading. Many of these showcase special features, and contain numerous examples of output.

5 Some tricks, tips, and examples of using Lertap5 and Excel in the context of an authentic national assessment of science study.

6 The e-store for Lertap5, the place which sells licenses for Lertap5 users when they have more than 50 cases to process.

7 The QUIA website. This site is our developmental and experimental site. It sometimes has special tidbits, and, at times, special options of interest to instructors in test and measurement classes.

Please direct questions or comments to: lertap5@gmail.com

Last update: 25 February 2023

1.1 What is Lertap?

Lertap, the Laboratory of Educational Research Test Analysis Package, is a computer program used to process and analyze results from tests and surveys. It has gone through several generations since its birth in the early 1970s. The fifth generation of Lertap, known as Lertap5, was first released in 2001, and is designed to work as an application running within Microsoft's Excel program.

This document provides some idea of what Lertap5 does, and how to go about getting it to do what it does. If you're new to Lertap5, you might find other resources to have a more introductory flavor.

For a quick but comprehensive look at Lertap5, try these web-based PowerPoint slides, or, if you prefer, browse the pdf version of the same slides. (Note: some of these slides show screen shots from previous versions of Excel.)

There's also the knock-your-socks-off manual, and, of course, the main website. There are some videos, too.

And yes, if that's not enough, there's even more: have a gander at our sample datasets website. It's a jim-dandy source for quickly getting an idea of what Lertap5 does, and it even has some samples which may be downloaded for test drives on your computer.

1.2 Requirements

Lertap5 is an Excel application. Your computer must be able to run Excel, the spreadsheet program from Microsoft, usually distributed as one of the applications in the software suite called "Microsoft Office". Other applications in this suite are Word, Access, and PowerPoint.

So? Excel. You need it to run Lertap5. Got it? Good. But note: your copy of Excel has got to be one which will "run macros". The version of Excel for the iPad is not capable of running macros. The various "mobile" versions of Excel, handy and powerful as they may be, are not capable of running macros either. (Macros are modules of computer code; Lertap 5 code is written in "VBA", Visual Basic for Applications.)

Do you have a Windows or Macintosh computer using Excel? If not, Lertap5 may not be for you (alas!).

1.3 Specifications

Frequently-asked questions (FAQs) often have to do with the number of items, and the number of data records which Lertap 5 will handle.

The answers depend on the version of Excel in use, on the number of columns and rows it supports.

Way back when, in the days of earlier versions of Excel, only 256 columns were allowed, with a maximum of 65,536 rows. In later years these dimensions increased to 16,384 columns and 1,048,576 rows (current at 2023); this effectively means (in theory) that Lertap 5 users may have over 16,000 items, and over a million students.

At a practical level, here at Lertap 5 headquarters we hear from users who sometimes run as many as 500 items, and over 50,000 student data records, well within the limits of recent versions of Excel.

As to the speed with which Lertap 5 can process results, well, the picture is a bit mixed, depending not only on computer speed, but also on the version of Excel in use. Please see the latest performance data here: time trial summaries.

As to some more specific Lertap 5 specifications, items and questions may use as many as 26 responses codes, or options. If letters are used as response codes, they may be upper- or lower-case. An analysis may have virtually any number of "subtests" or "scales" (also sometimes called "domains"). Any test item or question may belong to more than one subtest or scale, and may be scored differently in each.

Cognitive test items may have more than one keyed-correct answer (via *mws cards). Affective (or "rating scale") items may be reverse-scored quickly (by using *pol cards).

The number of scoring points awarded to any response is free to take on values between -9999.9999 and +9999.9999.

A brag page highlighting selected Lertap 5 features is available at this website.

1.4 How to get Lertap

Remember that Lertap5 runs with Microsoft Excel. Your computer must have a copy of Excel in order to be able to run Lertap5. Excel is part of the suite of programs which Microsoft refers to as "Office". Other apps in the Office suite include Word, Access, and PowerPoint.

The Lertap511.zip file with the latest version of Lertap5 may be downloaded using this link -- this version works on both Windows and Macintosh computers. It may also be downloaded from this OSF Link.

The ReadMe-First.pdf document is a must read after Lertap511.zip has been downloaded (or even before, for that matter). Some tips for Windows users are here, tips for Mac users here.

As mentioned in the ReadMe-First document, once the Lertap511.zip file has been downloaded, it must be unzipped, resulting in a folder with the three Lertap5 Excel files and supporting pdf files.

Readers unfamiliar with the use of zip files might search for "Using zip files with Windows", or "Using zip files with MacOS".

Related tidbits:

Users should make sure that their computer is running the latest version of Excel -- this may be especially important on a Mac as Excel for the Mac seems to be playing a bit of "catch up" with the Windows versions (the development of the Mac version has historically lagged a bit).

A document covering the on-going development of Lertap5, with version numbering, is available here.

1.5 Installing Lertap 5.10

Note 1: this topic is *only* for those using the former Windows Lertap5 installer program. The text below does not apply if users have download the Lertap511.zip file mentioned in the previous topic. The installer program is now outdated; its use is no longer recommended.

A recommended PDF document with a step-by-step summary of the installation and licensing process may be downloaded here.

Note 2: the installer program only works on Windows computers and will install a version, 5.10.9, which is now slightly outdated (as of June 2021). An updates summary is here.

When a Lertap5 installer program runs, it seats Lertap5's Excel files, such as Lertap5.xlsm, and a small collection of support documents, on your computer. The installer also sets up a Windows startup folder with "shortcuts" to Lertap 5's various files -- these shortcuts may then be found via the Windows start button, or, in the case of Windows 8 and 10, from what is called the "start screen".

On a Windows computer, the default folder for the installation will usually be the user's Documents folder. However, if the installation is made on a computer with multiple accounts, that is, on a computer with more than one user, each with a unique username and password, then someone with administrator rights on the computer should take charge of the installation, and see to it that the Public Documents folder is used. Once Lertap 5 is installed in Public Documents, each user may, if wanted, then copy it to his / her folder of choice -- the advantage to doing this is that it makes it possible for each user to have unique settings for Lertap 5 in the System worksheet.

Care must be taken to see that Lertap5 is not installed in one of the special Windows folders, such as Program Files. This is because Lertap 5 will often write temporary scratch files for its own internal needs; folders like Program Files generally have restricted access and do not take kindly to temporary scratch files.

Please refer to this document for more comments on installing Lertap5 on a multiple-user computer.

Once it's installed, Lertap 5 will be configured as what's called a "Mini" system. This is a full-blown version of Lertap5 when it comes to functionality: all of Lertap 5's features will be present and usable. However, the Mini version will process a maximum of 250 cases, or data records. You may have hundreds or thousands of data records ready to analyze, but, unless the Mini version is upgraded as described in the next paragraph, only the first 250 records will be looked at by Lertap5.

Upgrading the Mini version so that it will process more than 250 records requires a license. These are purchased from the Lertap e-store. Once a license has been obtained, Lertap 5 is activated, or "unlocked", by following steps provided in this document's License menu section.

1.5.1 Multiple installations

There are times when a single user may benefit from having more than one Lertap5 installation on his or her computer. The how and why of doing this is discussed towards the end of this document.

1.5.2 Uninstalling

What to do to remove Lertap5 from your computer, that is, how is Lertap5 uninstalled?

The answer will depend on how it was installed. The easiest answer is for Apple Macintosh users -- they will have downloaded a "zip" folder and copied its contents to a new folder on their computer, a folder which they (the users) will have named. Mac users can uninstall Lertap5 simply by deleting the folder.

Windows users will most likely have used the installer program to get Lertap5 onto their computers, as mentioned back on this page. The installer sets up Lertap5 as a true Windows "app"; this page has instructions for removing/deleting Windows apps (such as Lertap5) -- you'll see that there's more than one way, but, of the ways mentioned, for Lertap5 often the best way is to use the Control Panel's "Programs and Features" option as described in the instructions.

Need help? No worries, just send an email to: lertap5@gmail.com.

Related tidbit:

In some cases Windows users will have downloaded a zip folder containing Lertap5's files, and then copied the files to a new folder on their computer, a folder they (the users) will have named. They will uninstall Lertap5 simply by deleting the folder.

1.6 How to run it

A set of PowerPoint slides is a quick way to get started. Click here to see the slides saved as a PDF document.

For a more comprehensive look, see the "How to Lertap" topic at this website.

1.7 Versions and updates

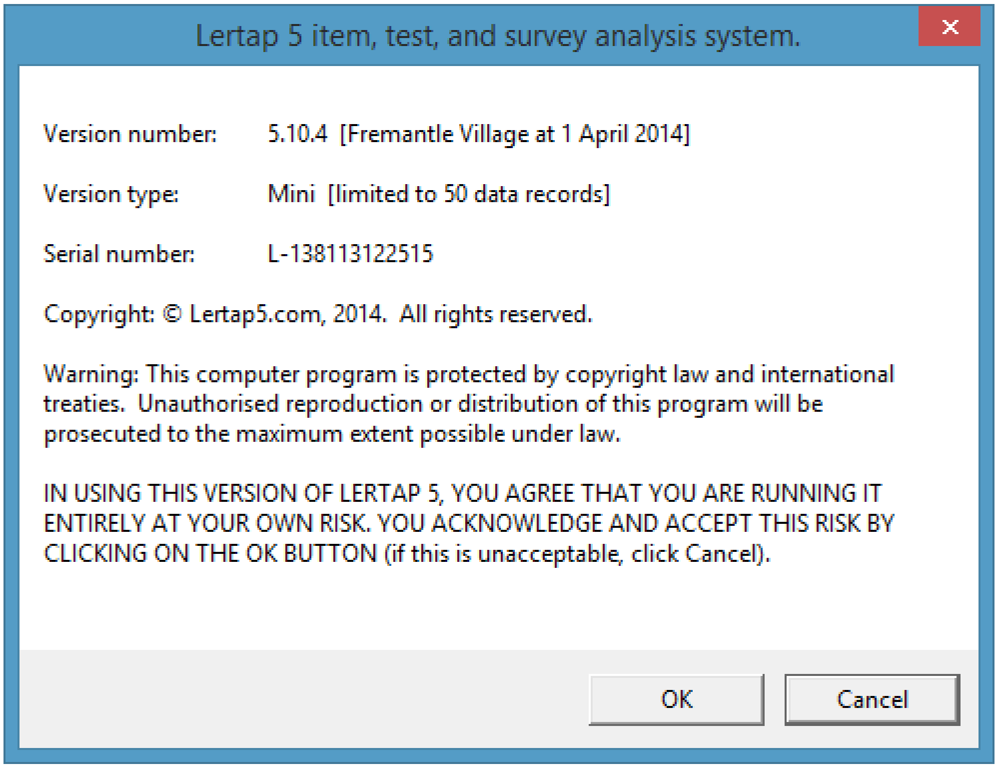
There are versions of Lertap5 for use with Excel on the Macintosh, and with Excel on computers running the Windows operating system. See the "How to get Lertap" topic for a bit more information.

Got a version of Lertap5 running on your computer?

To see which version it is, find the yellow smiley face and click on it. The screen snapshot below displays the Lertap5 "tab" on the Excel 2016 "ribbon". The yellow smiley face is seen in the "Basic options" collection of icons on the left. All versions of Lertap5 have the yellow smiley face available somewhere on the Excel ribbon.

These days, Lertap version numbers always start with a "5" which signifies an Excel-based version. After that will be a number; *this number is really the true version indicator*. Following it will be another number indicating sub-version, and then, at times, there may even a fourth number which usually denotes special bug fixes.

Here's an example:



In this case, the version number was 5.10.4, meaning the fourth edition of version 10 of Lertap5 for Excel. The date information indicates "build number" which, in this example, was 1 April 2014. This build was assembled in Fremantle Village, a small community in South Fremantle, Western Australia.

The example indicates that the "Mini" version was in use. A Mini version may be upgraded at any time by purchasing a license, a process addressed in the License menu topic. (The current Mini version is limited to processing no more than 250 data records. Upgrading eliminates this restriction. There is no limit on the number of items that can be processed.)

Updates

> There's another component to Lertap5: a separate Excel workbook with special-purpose code modules called "macros". It's called Lertap5MacroSetA.xlam. This workbook is a stand-alone, and is often updated on its own. Being a small file, it is easy to download and not at all difficult to install. Read more about it at this topic.
>
> A summary of system changes and updates, covering both Lertap5 itself and the special macros workbook, may be seen with a click here.

Write to lertap5@gmail.com for more information about versions, updates, and upgrades.

1.8 Known problems

There are a few problems which can arise when running Lertap.

Up until late 2018, the most common problem had to do with macro security. Lertap5, an Excel-based app, is programmed in a language called "VBA", Visual Basic for Applications. Computer code modules written using this language are referred to as "macros".

People who for some reason make it their business to cause havoc with our computers, for example by introducing a virus, have at times done so by nesting nasty, damage-causing computer code, in Excel macros.

To control for this real problem, when a user tries to open an Excel workbook which has macros (such as Lertap5), Microsoft commonly requires users to confirm that it's okay to open the workbook, and Excel may say something like "*Hey, the workbook you want to open contains macros, is it really okay to open it?*". The way a user indicates her/his approval varies, depending on the version of Excel, but it often involves setting "macro security" levels. Read about doing this by reading this topic.

Microsoft significantly enhanced macro security screening late in the year 2018 with the introduction of an "Anti-malware" screening system. When it comes to Lertap5, this new level of macro security can, and has, caused serious problems; the Lertap5.xlsm workbook has hundreds of macros, and the new security screener can get hung up trying to scan all of them. The result? Excel can appear to hang -- minutes will pass and nothing seems to happen. Or, Excel will just refuse to open the workbook. Or, Excel will agree to open Lertap5.xlsm, but will take minutes to do so (under ordinary conditions, Lertap5.xlsm will open in less than a minute).

There is a solution to this problem. Read about it here.

Minor problems: Excel version compatibility (click here), and an issue that can arise when the Windows Region and Language setting is set to a format where the decimal separator is a comma instead of a full stop. For example, if a test mean is formatted as 50,35 instead of 50.35, some of Lertap's options may fail. The "Res. Charts" option is one; it will produce a "Whoops" error message.

A more complete discussion of issues associated with the various versions of Lertap may be seen at the following URL:

http://www.lertap5.com/Documentation/UpdatesSummaryLertap57.pdf

Page forward to take in the macro security, malware-screening, and compatibility mode topics.

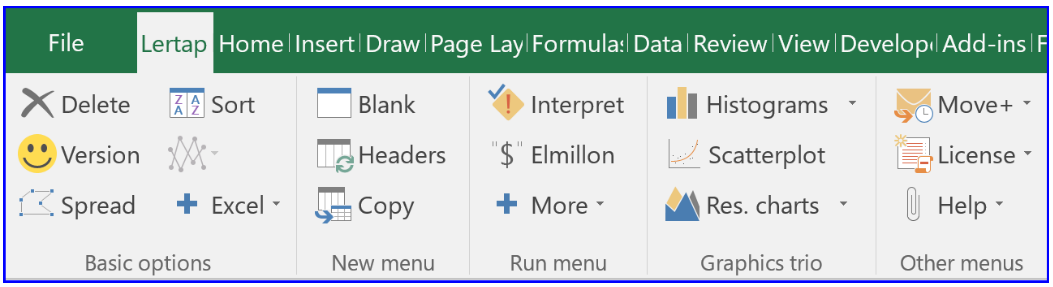
1.8.1 Macro security

Of the known problems, one which warrants special mention as it crops up so often: the matter of "Excel macro security". This problem is highlighted below -- it's usually quite easy to solve. (If you don't have the time right now to solve the macro problem, ask one of the kids to do it, or grandma if she's not busy making an apple pie (never *ever* disturb someone who's making an apple pie).)

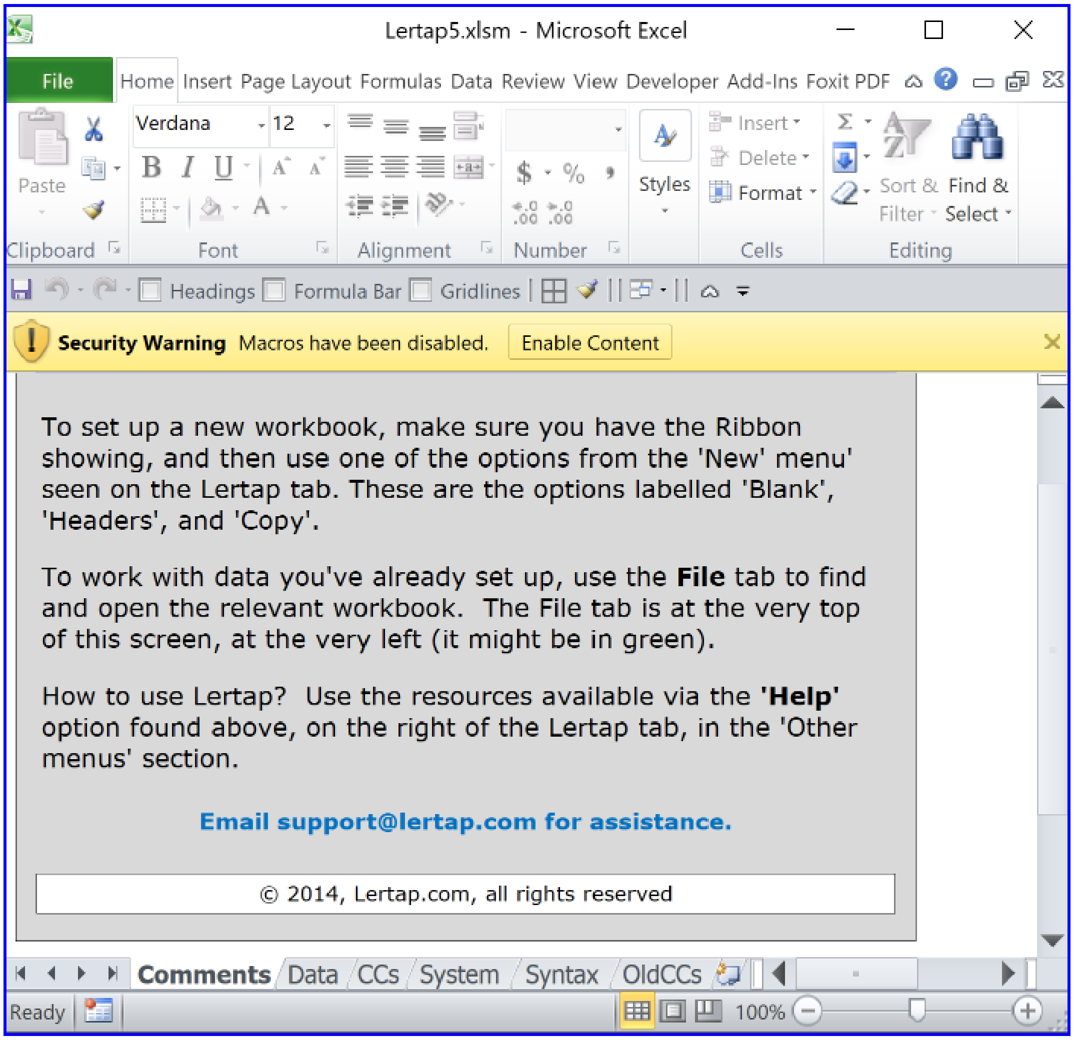
Excel macro security

Lertap's computer code, written in Visual Basic, is nested in a set of code modules referred to as 'macros' by Excel. In order to run Lertap, Excel has to be told that it's okay to 'enable' the macros found within Lertap. If this is not done, the Lertap tab for Excel will <u>not</u> be displayed, and without the tab you can't do any Lertapping.

The Lertap tab for current versions of Excel looks something like this:



Here's what <u>*may*</u> happen when you open the Lertap5.xlsm workbook without having enabled macros:



Note that the Lertap tab is not showing at the moment. Excel has displayed its security warning about macros. It has done so because macros have historically had something of a stained reputation; in the past they've been used by the bad guys/girls to sneak a virus through to your computer. Lertap's macros are safe and virus free -- it's okay to ignore the security warning in this case.

Click on Enable Content and you should be okay, the Lertap tab should load between the File tab and the Home tab on the ribbon.

What to do if there's no Enable Content message, and no Lertap tab?

Get a cup of coffee, a slice of grandma's apple pie, relax, and read this topic from Microsoft.

1.8.2 Compatibility mode

Excel 2007, and the versions that have come after it (365, 2010, 2013, 2016, 2019) like to run with workbooks having an extension of xlsx, xlsm, and/or xlam.

Earlier versions of Excel, such as the very popular Excel 2003, generally expected to create and run with workbooks having an extension of xls.

Excel 2007 and later versions usually have no problem at all with "old" xls workbooks.

When an xls workbook is opened with Excel 2007/2010/2013, Excel will say that it's running in "compatibility mode". The screen snapshot shown below exemplifies what's seen at the top of an Excel window when Excel 2010 is running in compatibility mode.

At times, Lertap will have a problem or two with compatibility mode. For example, when the "Item scores and correlations" option is taken, Lertap will sometimes fail. In such cases, a message such as the following may appear:

Sorry, there's been a problem inverting the correlation matrix ....

As far as Lertap goes, it is best to avoid running in compatibility mode. If you've opened an xls workbook with Excel 2007 or later, save a copy of it as an xlsx workbook. Then close Excel completely. When you return to Excel and open the saved xlsx copy, all should be well -- compatibility mode should not be running. Invisibly, Lertap will thank you.

For more assistance with this problem, please feel very free to write to lertap5@gmail.com.

1.8.3 Antimalware scan

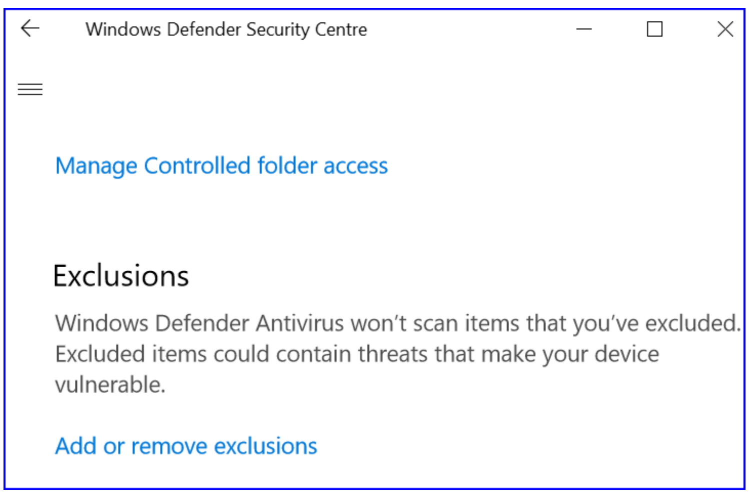
Microsoft implemented "AMSI", the antimalware scan interface, in September, 2018, as explained in this webpage.

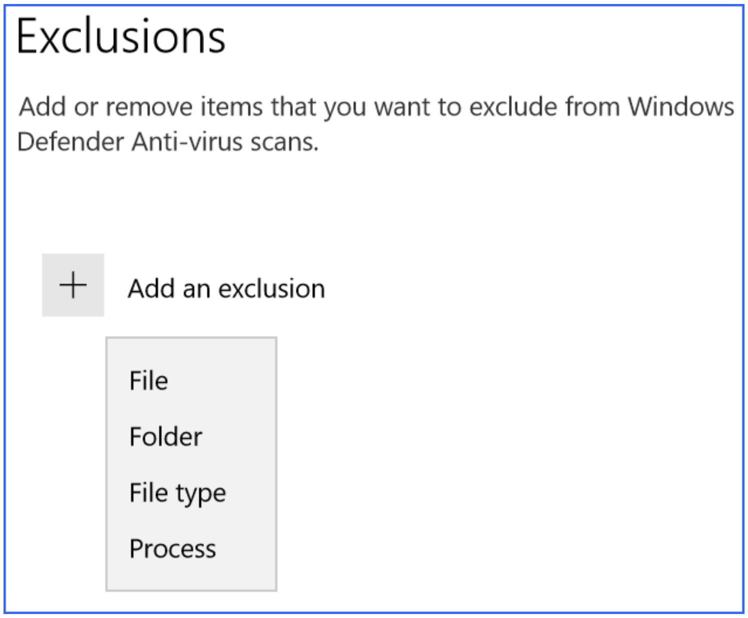
One of its tasks is to scan macros for malicious code (such as code which may introduce a computer virus).

While this is undoubtedly a good step, when it comes to Excel workbooks that have lots of macros, such as Lertap5.xlsm, AMSI can *dramatically* add to the time it takes for Excel to open the workbook, effectively slowing things down at the start, and, at times, even resulting in Excel seeming to be unable to open Lertap.

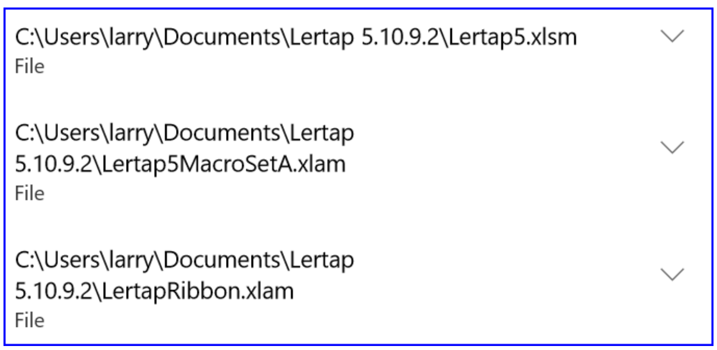
There's a way to control AMSI's behaviour via the "Windows Defender Security Centre", available from the Windows 10 Settings options.

The "Virus & threat protection" option leads to a menu of related settings. These may be used to define files that do not need to be scanned for viruses, files that can be excluded from scanning.





Use the File option to add Lertap5.xlsm, Lertap5MacroSetA.xlam, and LertapRibbon.xlam to the list of files to exclude from virus scanning:



Once this has been done there will be a noticeable decrease in the time it takes Excel to open Lertap5.xlsm.

1.9 About this document

This document, "Lelp", was made using Help&Manual 6, a hypertext authoring system from EC Software.

Note that this version of Lelp applies to the Excel 2007, 2010, 2013, 2016, 2019, 365 versions of Lertap5 for Windows (versions 10.9 and up). It's also appropriate for Macintosh users who have Excel 2016, 2019, 365.

H&M lets us compile Lelp in one of several formats, and we've done so. To date there are PDF and chm compilations. "chm" help files are now found in most Windows applications; they're easy to use. PDF files are now a world standard, usable on all sorts of computers; this is nice, but PDF files cannot display what are called "popup" topics. This document contains quite a number of popups -- they flash up quickly in the chm version, but don't show at all under PDF. If you're looking at the PDF version, you'll sometimes see spots which say something like "click here", or "note", but there's nothing clickable -- alas, you've come upon a popup that won't pop.

Links to all versions of Lelp are found back on the Welcome page, see the first link in the green box towards the bottom of that page.

Since all forms of this document are produced from the same source, they're identical. The chm, PDF, and website versions have exactly the same content. However, as noted, the PDF version cannot display popup topics.

Screen shots in manuals have a habit of dating quickly; the shots you see in this document will not have a 100% correspondence with the screens seen in the most recent version of the Lertap 5 system. This is particularly true since Microsoft released Excel 2007 for Windows, the first version to work with the so-called "Ribbon" of tabs and icons.

This version of Lelp, the one you're looking at right now, was made by modifying and enhancing the previous version, Lelp for Excel 2003 (Windows) and Lelp for Excel 2004 (Macintosh). Some of the screen snapshots found in this version of Lelp display information in the format found in former versions of Excel, versions which had "toolbars" with controls and icons instead of "ribbons".

1.9.1 Differs from manual.

Lelp, that is, this document, serves a variety of purposes.

Above all, Lelp effectively updates the manual. For a summary of the major changes made since the manual was printed, see the Revisions topic.

A number of supporting documents have been made available to users since the manual came to light, and Lelp has links to them, often as URLs found under the "Related tidbits" section at the end of some topics.

Most versions of Lelp, such as chm, PDF, and website, may be viewed online. The manual is also available in an electronic form, but, unlike Lelp, it's not cross-referenced, it has not been formatted as hypertext -- it's not as easy to jump from topic to topic in the manual as it is in Lelp.

Both the manual and Lelp have numerous examples. The ones in Lelp tend to be somewhat more advanced.

Lelp's explanation of CCs lines and syntax is somewhat more extensive than that found in the manual.

Users of the Windows version of Lertap 5 will find that some of Lertap's dialog boxes, and most of its menus, have automatic, context-sensitive links to Lelp. By and large, these links are denoted as "Lelp". Macintosh users do not yet have a similar resource -- they may refer to the website, or to the PDF version.

Finally, this document has an invaluable educational supplement: in numerous spots it introduces elements of Australian slang (*strewth!*); in other spots it has links to selected Western Australian cultural highlights, such as Emu Export, Mt Barker, and the Southwest Capes.

1.10 Contact us

Lertap.com is pleased to bring you Lertap 5. Larry Nelson is the project director. Marketing, workshops, and user records are matters handled by Angie Nelson (aka Khin Khin Than). Click here for a bit more staff information.

The project's main website is here:

www.lertap5.com/lertap/

Address:

Larry Nelson
Lertap.com
Fremantle Village Suite PH42a
25 Cockburn Road
South Fremantle, Western Australia 6162

Western Australia is in the same timezone as Singapore and Beijing.

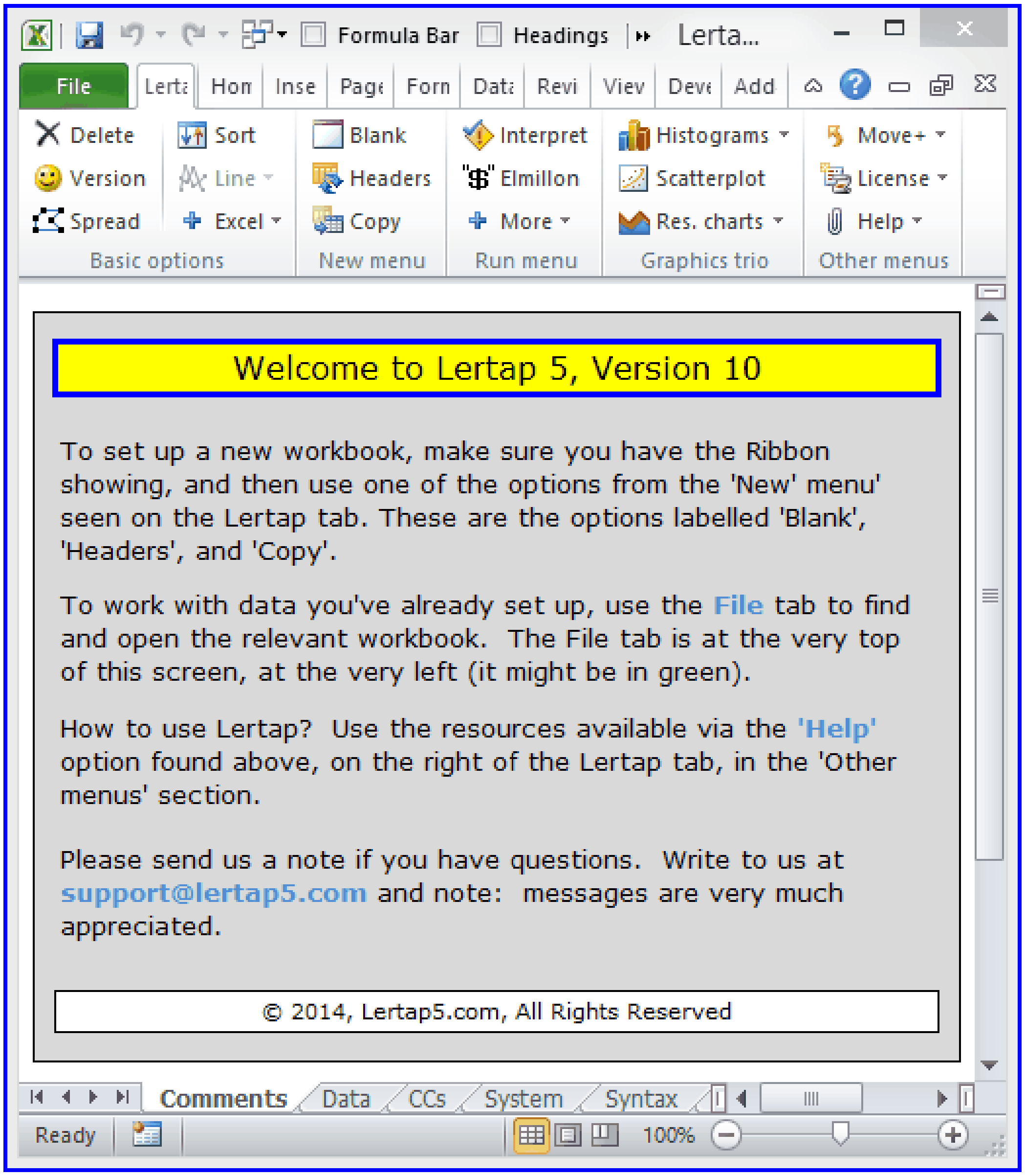
Phone +61 4 6836 2643
email (1) lertap5@gmail.com
email (2) angie@lertap.com
email (3) larry@lertap.com

2 Getting started

(Be sure to also visit the "What is Lertap topic" for getting-started suggestions.)

It's a fairly simple matter to get started with Lertap. Let us get you launched on the so-called "Cook's tour", and you can see for yourself.

When you start Lertap5 using Excel for Windows, or Excel for the Macintosh, your computer screen should bear quite a resemblance to the screen snapshot show below: (If it doesn't, refer back to HowToRunIt, and MacroSecurity.)

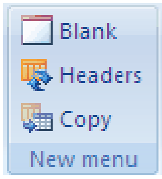


An important component of the screen seen above is the line of worksheet tabs showing towards the bottom of the screen. There are several tabs: Comments, Data, CCs, System, Syntax, and OldCCs.

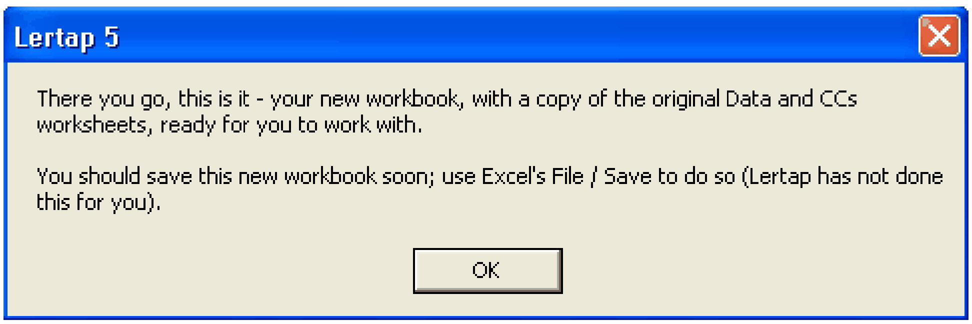
Each tab corresponds to an Excel worksheet. The Data worksheet has the responses of 60 people to two quizzes. The CCs worksheet has a series of lines with Lertap's control syntax. Each of the syntax lines gives Excel information on how to process the data found in the Data worksheet. The Syntax worksheet is a quick reminder for experienced Lertappers on the correct format for CCs lines. The System worksheet turns on/off assorted Lertap options.

Go ahead and look at each of the worksheets if you'd like. Then come back here.

Okay? The Cook's tour starts with the Copy option on the New menu. This is your first task: find the New menu; find its Copy option; click on the Copy option. Hint: see screen snapshot below.



The Copy option causes Lertap to do a couple of somersaults and back flips -- it makes a copy of the Data and CCs worksheets, and places them in a new Excel workbook. When it finishes this task, Lertap says something like this:



Don't worry about saving the new workbook for the moment. Click the OK button.

Your next move will be to click on the Interpret option on the Run menu.

This gets Lertap to read the lines in the CCs worksheet, checking to make sure they have the right syntax. If they do, Lertap creates some new worksheets, and displays one of them: "Freqs".

The Freqs worksheet is a simple one. It displays what are called "response frequencies" for each of the columns in the Data worksheet.

Q1

Option	n	/60
A	26	43.3%
B	25	41.7%
C	9	15.0%

The little boxes above indicate that 26 people answered "A" on Q1.

The main purpose of the Freqs worksheet is to see if there may have been any errors in the preparation of the data. Users generally scroll down the Freqs worksheet rather quickly, looking for unexpected results. For example, a response of "D" on Q1 would be strange as Q1 allowed for just three responses: A, B, C.

To continue the tour, return to the Run menu now, and click on the Elmillon option. This causes Lertap to do quite a number of things -- for example, it creates test scores, putting them in a new worksheet called, appropriately, Scores. It also creates a variety of statistical reports, giving these names such as Stats1f, Stats1b, Stats1ul, Stats2f, and Stats2b. The partial screen snapshot below shows what the worksheet tabs will look like after this step:

The various "Stats" worksheets are usually why people run Lertap -- they give item and test statistics. Stats1f provides a complete, or "full", item and test analysis report for the first quiz, or test, giving quite detailed information for each test item, and presenting several mini-reports with overall test statistics. Branch out to the Stats1f topic for much more information.

The Stats1b report is a briefer summary of the item statistics found in Stats1f; it's easier to read. Branch out to the Stats1b topic.

If the test being analyzed is a cognitive test, Lertap generally produces a third report, Stats1ul; the "ul" means upper-lower, referring to a method of item analysis preferred by some. Branch out to the Stats1ul topic.

Lertap produces more reports than many people want. For example, the "b" Stats sheets, such as Stats1b and Stats2b, are redundant in that their information is also to be found in the respective "f" sheets, such as Stats1f and Stats2f. Some users make little use of the "b" sheets, preferring to drink their cup of tea with the "f" sheets. On the other hand, some Lertap users find there to be too much information in the "f" sheets; their needs are met by the "b" sheets, where they find that just half a cup of tea, a "brief" cup, is all that's needed to peruse the output. (It's possible to get Lertap to reduce the number of reports it gives. This is done by setting options in rows 9 and 10 of the System worksheet.)

Why are the "f" sheets made so that their initial focus is well down the worksheet? Stats1b and Stats1ul, for example, display their top rows at the top of the screen, but

Stats1f and Stats2f do not display their upper-most rows at first. Why? Because experience has shown that the majority of users scroll to the reliability section of the "f" sheets before they look at anything else. A similar effect is seen in the Scores worksheet, where Lertap places the initial focus well down the worksheet, where the summary statistics begin.

For a thorough discussion of the statistics found in the Stats sheets, use the links above and/or click here to gain access to the manual.

2.1 About the Cook's tour

We might do well to highlight some of the things that happened when you took the "Cook's tour" covered in the previous topic. And, as you'll read below, we've made some changes to Lertap which result in the tour producing more reports.

The Cook's tour is based on the Lertap Quiz data set. This data set is fully described in Appendix A of the manual. Briefly, the Quiz consists of a "test" given to 60 people who had the good fortune to participate in one of the very first Lertap workshops, held in Dunedin, New Zealand, just a few years ago (well, it was actually 1973, more than 40 years ago).

At the end of the workshop, each participant was asked to answer 25 multiple-choice questions and 10 Likert-style affective questions.

The multiple-choice questions were meant to indicate how well the participants had mastered the content of the workshop -- the 25 questions were a test of their knowledge of the functioning of Lertap.

The 10 Likert questions, scored on a 5-point strongly-agree to strongly-disagree scale, asked the participants how they felt about Lertap -- did they judge it to be the dynamite test and survey analysis system we know it to be?

Each participant was also asked to respond to two open-ended questions. One of these asked them to state how long they had been using computers, while the other requested information on how long they had been using tests in their research or teaching.

The Lertap Quiz data set is included in the Lertap5.xlsm file. When you start Lertap, you'll see tabs at the bottom of Excel's screen, one for each of the worksheets contained in the Lertap5 workbook (refer to the previous topic for a picture). If you look at the Data worksheet, you'll see the actual question responses given by the 60 participants in the 1973 workshop.

When you took the tour, your first action was to get Lertap to make a new data set for you to fool around with. This you did by going to the New menu, clicking on the option which directed Lertap to make a new workbook by creating a "copy of the present one".

The "present one", in this case, referred to the data set which comes with the Lertap5.xlsm file, that is, to the Lertap Quiz. At the end of this step you had a new workbook which contained copies of the Data and CCs worksheets corresponding to the Lertap Quiz. All subsequent actions in the Cook's tour involved the use of these copied worksheets.

2.2 How it works

You've taken the Cook's tour? Goodonya (!).

We've got some examples coming up, but first how about a quick overview of how Lertap works?

Lertap is used to analyse the responses people have given to a test or survey. Lertap is designed to work with fixed-choice items (items = questions). For a test, fixed-choice usually means true/false or multiple choice. For a survey, fixed-choice means that the answers people can give to a question have been listed, and enumerated. For example, a survey might ask people if they think beer is a good thirst quencher on a hot summer's day; people can answer 1 (strongly disagree, a foolish response); 2 (disagree); 3 (undecided); 4 (agree); or 5 (strongly agree).

The answers people give are placed in the Data worksheet. Each row in the Data sheet corresponds to the answers of one person.

Having responses recorded in the Data worksheet is good and necessary before results can be obtained, but it's not all that's required. Some instructions are also needed. Instructions? For whom? For Lertap.

Lertap is not smart enough to be able to look at the Data sheet, and figure out what's what without your help. You've got to tell Lertap which columns in the Data worksheet have the item responses you want it to look at. If the items are from a test, you have to tell Lertap what the right answer to each item is. If the items are from a survey, on the other hand, you have to tell Lertap to forget about having to have a correct answer for each question -- there aren't any.

You tell Lertap these things by putting your instructions in another worksheet called the CCs sheet, using a special control language.

Once you've got the responses in the Data sheet, and your Lertap instructions in the CCs sheet, away you go ... you just follow the same steps seen in the Cook's tour. Could life be simpler? Page on ... have a look at some examples (there are more in the manual).

Related tidbit:

For more how-it-works insights, see: Lertap's Output.

2.3 Examples

There are three primary places to visit for those interested in browsing some samples of Lertap in action.

The first "place" is right here, in the very document you're looking at now. Simply page ahead and you'll see.

The second place repeats some of the samples found here, but adds more, presenting a greater variety of samples, largely based on "real-life" applications of Lertap collected over the years. It offers the chance to download actual Excel workbooks set up to work with Lertap 5, and suggests practical "exercises" to provide more extensive insights into how Lertap may be applied. Where is this "second place"? Here.

The third spot is a website which displays Lertap 5 in action, processing data from a 2012 national assessment of science achievement among primary-school children.

But wait, mention should also be made of yet a fourth place: the manual. Its Chapters 7 and 8 also have real-life examples; they may be a bit easier to follow for some readers as they're more introductory in flavour.

2.3.1 Cognitive example

Lookit, lookit, lookit ... here's a Data worksheet ...

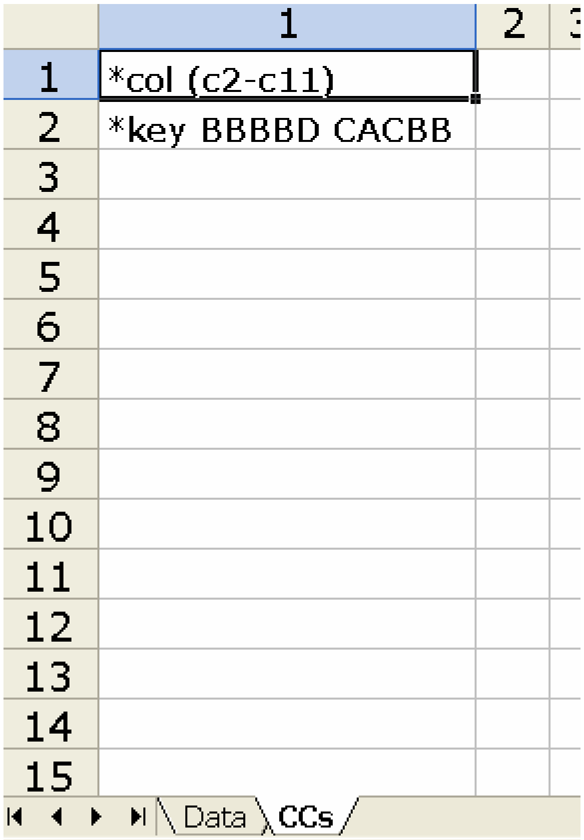
	1	2	3	4	5	6	7	8	9	10	11
1	Data from the ChemQuiz sample.										
2	**ID**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
3	Anderson	D	B	B	C	D	D	A	C	A	B
4	Baker	B	B	B	B	D	B	C	C	B	B
5	Camberwell	B	B	A	B	D	A	B	C	B	D
6	Donaldson	B	B	C	B	D	C	B	C	C	B
7	Eggmont	B	B	A	B	B	C	C	C	B	B
8	Fredricksson	B	A	B	B	D	C	A	C	B	A
9	Graphner	D	D	A	B	D	C	A	C	B	A
10	Humphrey	B	B	B	D	D	C	C	D	B	B
11	Invererity	B	B	A	D	B	C	B	C	A	A
12	Johnson	B	D	B	C	A	B	C	C	B	B
13	Klein	D		B	A	D	C		C	A	A
14	Lampton	B		A	A		C		A	B	A
15	Mecurio	B		B	D	A	C	C	A	B	B
16	Nesbit	A	C	A	D	B	C		A	D	A
17	Oldfelt	A	D	A	A	A	C	A	C	B	A

Data / CCs

What have we got here? Fifteen students answered 10 multiple-choice questions. On the first item, Q1, Anderson selected option D. On the fifth item, Invererity selected option B.

Anderson's data are found in row 3 of the worksheet. The answers to Q1 are found in column 2 of the worksheet. The answers to the last question are found in column 11. (You can't see row 18, but it's empty.)

And here's the corresponding CCs worksheet:



Yes, the CCs worksheet has just two rows of information. The first one tells Lertap that item responses are to be found in columns 2 through 11 of the Data worksheet. The second line has ten letters; each letter is the right answer, the "key", for an item. For example, the right answer to the first question is B. The right answer to the 6th item is C. The correct answer for the 10th item is B.

And that's it. Yes. The data set is ready for Lertap. At this point, one would go to Lertap's Run menu and click on "Interpret CCs lines". Then, after being Freq-ed out, we'd go back to the Run menu and click on "Elmillon item analysis". Beauty.

Related tidbit:

The example above is based on a cognitve data set which may be seen at this URL:

http://www.lertap5.com/Documentation/Samples/TenCogs/TenItemCognitive.doc

2.3.2 Affective example

Here's another example:

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Ed 503 class survey, 8 September.												
2	No.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
3	1	3	3	3	4	3	3	3	3	2	3	3	4
4	2	3	2	3	3	2	4	4	4	4	3	3	5
5	3	3	3	2	2	2	2	4	4	3	3	3	5
6	4	1	2	3	4	4	2	1	2	2	2	2	5
7	5	2	2	2	2	3	3	1	3	2	2	2	2
8	6	2	3	2	3	3	3	2	3	4	5	2	3
9	7	2	3	2	3	3	3	1	2	4	3	3	5
10	8	2	4	3	3	3	2	3	2	2	1	2	3
11	9	1	3	3	3	3	2	2	3	2	2	1	5
12	10	2	4	1	1	1	1	3	2	2	2	3	1
13	11	1	3	2	2	2	2	3	2	3	4	3	4
14	12	3	2	2	2	3	2	3	3	2	3	3	4
15	13	3	3	5	1	1	1	1	3	2	3	3	4
16	14	2	2	1	1	3	3	3	2	2	3	3	3
17	15	3	3	3	2	3	2	4	3	3	3	2	4

Here we've got another 15 students, and this time we have responses to 12 survey items. Some sort of number is found in the first column, with the responses to the first question, Q1, found in the 2nd column.

	1	2	3
1	*Col (C2-C13)		
2	*Sub Affective		
3			
4			

Data CCs

Once again there are only two CCs rows with information. The first row tells Lertap that item responses are found in columns 2 through 13 of the Data worksheet. The second row tells Lertap that these items are "affective", or survey, items. Having seen the word "affective" on a *sub CCs line, Lertap knows that the questions do not have a correct answer -- they're to be scored using the default survey scoring scheme where a response of 1 equals one point, a response of 2 = 2 points, and so on.

This example is complete. Nothing is missing. At this point, one would go to Lertap's Run menu and click on "Interpret CCs lines". Then, after being Freq-ed out, back to the Run menu and a click on "Elmillon item analysis". Beauty abounds, doesn't it?

Related tidbit:

Another example of a survey, with actual questions used, may be seen at this URL:

http://www.lertap5.com/Documentation/Samples/CEQ/CEQ1.DOC

2.3.3 Lertap data set

We've given you two really straightforward examples. In our experience, the two simple examples you've seen would be very much like what maybe a quarter of Lertap users regularly live and breathe. But Lertap was designed to handle more complex situations. Those little CCs lines can pack more punch than what's been on display in the previous two samples.

As an example, look at this CCs worksheet:

	1
1	These control "cards", or lines, set up two subtests.
2	Different background colors are used below, but they're not required.
3	**The first subtest has 25 cognitive items; responses start in column 3.**
4	*col (c3-c27)
5	*sub Res=(A,B,C,D,E,F), Name=(Knowledge of LERTAP2), Title=(Knwldge), Wt=0
6	*key AECAB BEBBD ADBAB BCCCB BABDC
7	*alt 35423 35464 54324 43344 45546
8	**The second subtest has 10 affective items; responses start in column 28.**
9	*col (c28-c37)
10	*sub Aff, Name=(Comfort with using LERTAP2), Title=(Comfort), Wt=0
11	*pol +---- ++--+

There are 11 lines in use in this CCs example. Four (4) of the lines are comments; these are the lines which do not begin with an asterisk. The use of comments is entirely optional, but they can be real helpful.

There are two *col lines above. Each of these defines a group of items which will be processed together, as a unit. Such units are generally referred to in Lertap as "subtests".

A subtest may be comprised of cognitive items, or it may be comprised of affective items.

You've already seen examples of the CCs lines used with these two types of subtests. We've pointed out that cognitive subtests will always have a *col line and a *key line, while affective subtests will always have a *col line and a *sub line.

In the example above, a *sub line has been used with the cognitive subtest for several reasons. In this example, some of the items had as many as six possible responses -- that is, some of the multiple-choice items used in this subtest had six choices, or responses.

The Res=(A,B,C,D,E,F) declaration tells Lertap this. Without an Res= declaration, Lertap assumes Res=(A,B,C,D), the default set of item response codes for cognitive items. (The default set for affective items is Res=(1,2,3,4,5).)

The *key line gives the right answer for each of the 25 items in the first subtest. The *alt line tells Lertap that the items used a different number of the six possible responses. For example, the first item used just 3 responses: they would be A, B, and C, the first three characters found in the Res= declaration. The second item used 5 responses: A, B, C, D, and E. Only two of the 25 items made use of all 6 possible responses.

The Name= declaration provides a brief description of each subtest, while the Title= declaration gives a short title. When Name is used, some of Lertap's reports display Name on their top line. Title, when used, appears in some reports as a label for subtest scores.

Subtest scores? What are they? Well, as an example, on a multiple-choice test people usually get one point for each right answer. On a 25-item test, or "subtest", it would be possible to get a score of 25, assuming one point for each correct answer (Lertap permits the right answer to have any number of points, and it even allows the wrong answers to have points too -- sometimes wrong answers are penalised by assigning them negative points).

Let's look now at the affective subtest defined above. The *col line points to 10 (ten) columns, c28 through c37. A *sub line is required for affective subtests, and it must have the "Aff", or "Affective", control word on it. This is seen above. In this case, the *sub line has also been used to assign a Name and a Title.

What's that *pol line doing? To answer this question, we have to return to the matter of scoring. Affective items do not have a correct answer. It's customary to give a certain number of points for each of an affective item's answers.

What were the possible response choices for this set of 10 affective items? Very good question. There is no Res= declaration on the *sub line, and, in this (common) case, Lertap assumes Res=(1,2,3,4,5). Each affective item had five possible responses. Unless you say otherwise, Lertap will give one point if someone selects 1 as their response; two points when someone selects 2; and so on. This is called "forward scoring". On a 10-item affective subtest with five response choices per item, the maximum possible score would be 50; the minimum possible would be 10.

The *pol line allows each affective item to be reverse-scored, if wanted. On a reverse-scored item, the first possible response will get 5 points, not 1. The last (fifth) response will get just 1 point, not 5. Reverse scoring for affective items is pretty common.

The *pol line above indicates that the first item is to be forward-scored, while the next four items are to be reverse-scored. This subtest has 10 items, so there are 10 + or - (minus) symbols shown on the *pol line. (In fact, for this subtest, four items are forward-scored, while six are reverse-scored.)

Lertap's forte is in the flexibility it provides for item scoring. Any response to any item can have any "weight", that is, any number of points. In the three examples we've presented thus far, including the one above, items are being scored in a conventional manner. Departures from normal are supported by the use of other CCs control lines, such as *wts and *mws.

Are we going to get away without talking about the Wt= declaration seen on the two *sub lines? No siree Bob; here goes: whenever multiple subtests are scored, Lertap will add up all the subtest scores to make a "Total", or "composite", score for each person. Each subtest ordinarily comes into the composite with a weight of one (1); to keep a subtest out of the total score, Wt=0 is used. In the example above, both subtests have been given a weight of zero, and Lertap will not make its Total score.

One final point which people often ask about ... there are spaces in the *key line above, in the *alt line, and also in the *pol line. There's a space after every five characters in each of these lines. Why? Simply to make the line a bit more legible. The spaces are not required.

Is the example above a common one? Yes and no. Yes in the sense that Lertap users frequently have more than one subtest to process, no in the sense of mixing subtest types -- this example has a cognitive subtest, and an affective subtest: a mix of subtest types -- that's quite uncommon. If you browse on into the following topics, you'll see a couple of other examples.

Related tidbit:

For a really bonza example of a job which worked Lertap's CCs lines close to the limit, have a look at "Using Lertap in a Parallel-Forms Reliability Study", a 16-page Word document available via the Internet: click here if you're connected.

The Total score, a composite formed by summing subtest scores, gets further mention here.

2.3.4 Multiple cognitives

Consider these CCs lines:

```
*col (c1-c10)
*sub name=(Addition), title=(Add)
*key DCCAB BCDDA
*col (c11-c20)
*sub name=(Subtraction), title=(Sub)
*key BBBCA DAACB
*col (c21-c30)
*sub name=(Multiplication), title=(Mult)
*key CDCAB AAACC
*col (c31-c40)
*sub name=(Division), title=(Div)
*key AADCC CBAAA
```

Someone's given a maths test with four subtests. Each subtest had ten items.

Lertap will create four subtest scores, and a total score. The total score will simply be the sum of the four subtest scores. The maximum possible score on each subtest is 10, hence the maximum possible total score is 40. (It is possible to change the number of points given for right answers by using *mws and *wts lines in the CCs worksheet.)

The *sub lines do not have Res= declarations, so Lertap will assume Res=(A,B,C,D) for each subtest.

2.3.5 Multiple affectives

Here's a common example of CCs lines for a survey with three subscales:

```
*col (c5-c20)
*sub aff, title=(Anxiety)
*pol ++--+ -++++ ++-++
*col (c21-c35)
*sub aff, title=(Friends)
*col (c36-c50)
*sub aff, title=(Homesick)
*pol ---++ +++-- -+-++
```

Three 15-item affective subtests are defined by these lines. Two of the three subtests, the first and the third, have a mixture of forward- and reverse-scored items. There is no *pol line for the second subtest, which means that all items for this subtest are forward-scored.

Lertap will make three subtest scores, and also a total score.

What about possible score ranges for this example? Each subtest has 15 items. There being no Res= declaration on the *sub lines, Lertap assumes Res=(1,2,3,4,5), that is, five possible responses per item.

Lertap will score each item on a one- to five-point basis. Why? Because there are five possible responses. The minimum score a person can get on an item is one; the maximum is five. There are 15 items in each subtest. Therefore, the score range for each subtest is 15 to 75, and, there being three subtests, the range for the total score will be 45 to 225.

What happens when a person doesn't answer an item? What sort of score do they get?

For cognitive tests, a non-answer gets a score of zilch (zero). However, for affective items, a non-answer will get a score equal to the mean of the item's response weights (note). It is possible to defeat this scoring system by using the MDO control word on a *sub line.

Note that it is possible to achieve almost any sort of scoring for affective items, or, for that matter, cognitive items. This is done by using *mws lines in the CCs worksheet.

3 CCs details

You've seen that Lertap analyses start with the creation of two Excel worksheets: Data and CCs.

The CCs worksheet contains the all-important lines of Lertap syntax which effectively control how Lertap and Excel analyse the data found in the Data worksheet.

CCs really stands for Control Cards. It used to be the case that data analysis was based on the use of punch cards. Years back, a typical data analysis job involved the use of a keypunch machine -- data were punched on cards, as were the instructions which told the computer how to analyse the data. The first versions of Lertap were based on the use of punch cards, and the term "control cards" has been carried into most subsequent versions.

You'll see that we sometimes refer to the rows in the CCs worksheet as "lines", and sometimes as "cards". We use these terms interchangeably; they mean the same thing.

There are a total of eight "cards" which may be used in a CCs worksheet. The number of cards used in any given job depends on two main factors: the type of subtests being processed, and the complexity of the item scoring desired.

We'll spell out the general nature of all of the cards below. The topics immediately following provide more exact details on the syntax of each card.

*col	The basic Lertap control card, used and required by all subtests, cognitive and

	affective. Each time Lertap sees a *col card in the CCs worksheet, it thinks "Ah-ha, here comes a new subtest", and it expects the user to then indicate the columns in the Data worksheet which are to be processed. (Click here to read about a problem which can arise with *very* long *col lines.)
*sub	This card is optional for cognitive subtests, but required for affective subtests. *sub cards are used to convey particular subtest characteristics to Lertap, such as the name and title of the subtest, and the number and nature of the response codes used by the items belonging to the subtest.
*key	Gives the right answer for the items of a cognitive subtest. This card is always required for cognitive subtests, but it's not used at all with affective subtests.
*pol	"pol" stands for polarity, that is, for plus (+) or minus (-). Not used by cognitive subtests, and optional for affective subtests. When used, it defines the type of scoring to be applied to affective items: plus (+) for forward, and minus (-) for reverse.
*alt	An optional card for both cognitive and affective subtests. When used, it indicates the last response code used by each item. If this card is not used, it is assumed each item uses the same number of response codes. (Note: the format of this card has now changed; please refer to Example C7 under the Cognitive CCs topic for details.)
*wts	An optional card for cognitive subtests; not used by affective subtests. This card makes it possible to quickly tell Lertap that the items of a cognitive subtest have different "weights", that is, the right answers to the items have differing point values (the first question might be worth one point, for example, while other questions might be worth more points). This card is also known as the *wgs card.
*mws	An optional card for both cognitive and affective subtests. "mws" stands for multiple-weights specification. This is the most powerful control card of all -- it allows any weight to be applied to any item response.
*exc	An optional card for both cognitive and affective subtests. "exc" stands for exclude. This card is used to quickly remove items from a subtest. (*exc is not mentioned in the manual.)
*tst	This is a very special control card. It can only be used once in any CCs worksheet, and, when used, it has to be the very first card. It's used to get Lertap to make a copy of the data set, with only certain data records to be copied to the new data set's Data worksheet. *tst is used to set up a new Lertap workbook containing a subset of the original Data records (for example, just the males, or only those in a specified School District).

This has been a quick introduction to Lertap's control "cards". We'll go on now to provide more specifics, and we'll do this by subtest type, cognitive first, then affective.

3.1 Cognitive CCs

Before getting into the syntax for CCs cards used to analyse cognitive items, let's come to terms with some terms.

Each cognitive item may use up to twenty-six (26) response codes. Response codes are also known as alternatives, or as options. A true/false item may use {T and F} as response codes, or {t, f}, or {1, 2}. A cognitive item with four possible responses may use codes of {A, B, C, D}; or {a, b, c, d}; or {1, 2, 3 , 4}.

Associated with each response code is a weight, the number of points a person gets for choosing the corresponding option. For example, if the right answer to an item is A, then people who select A will get a certain number of points; people who select one of the item's other responses will (usually) get no points.

Okay? Now then ....

Let's say we've given a 5-item cognitive test, with answers appearing in columns 2, 3, 4, 5, and 6 of the Data worksheet. We'd like Lertap to spin its magic, to analyse our data. In order to do this, we'll need to enter some lines in the CCs worksheet. Put on a fresh pot of coffee, pour yourself a cup, and have a look at the examples below.

Example C1:

This set of two CCs cards might be all that's required to get Lertap to analyse the data:

```
*col (c2-c6)
*key ACCDB
```

Anyone who selects A on item 1, C on items 2 and 3, D on item 4, and B on item 5 will get a score of 5 -- one point for each answer. Why? There are five items; the right answers, the "keyed-correct" answers, are shown on the *key card above. Unless you say otherwise, Lertap awards one point for each right answer.

Example C2:

We'll add a *sub card in order to have Lertap label some of its reports:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
```

The *sub card is usually optional for cognitive subtests. Here we're using one just to give a Name and a Title to the subtest. The Name will appear as a heading at the top of Lertap's item analysis reports, such as Stats1f and Stats1b. The Title will appear at the top of one of the Scores columns, making it a bit easier to interpret the Scores report. The Name can have any length, but Title should be no longer than 8 characters. If Name and/or Title are not given on a *sub card, Lertap will create default labels: Name=(Test 1), and Title=(Test1).

Example C3:

Next we'll use a *sub card in order to turn on certain scoring options:

```
*col (c2-c6)
*sub Title=(NewsQuiz), PER, SCALE
*key ACCDB
```

Now the *sub card has three control words, Title, PER, and SCALE. PER gets Lertap to create a percentage score for each test taker, being the student's score expressed as a percentage of the maximum possible score. For example, if the maximum score is 5, and a student got three items correct, PER=60%. The SCALE control word adds the student's z-score to the Scores report; on a test with a mean of 3, standard deviation of 1, a student test score of 4 would correspond to a z-score of +1.00.

Example C4:

To switch Lertap into its mastery scoring and report mode, include the word MASTERY on the *sub card, as shown here:

```
*col (c2-c6)
*sub Title=(NewsQuiz), Mastery
*key ACCDB
```

Using the MASTERY control word on *sub causes two things to happen. Each student will have her/his percentage score automatically included in the Scores report, just as happens when the PER control word is used. More importantly, the MASTERY control word gets Lertap to substantially alter one of its main statistical reports. The Stats1ul report will include a summary group statistics table, a variance components analysis, and two classification accuracy indices (please refer to Chapter 7 of the manual for details, and also take in a 2007 journal article dealing with the use of cut scores).

Lertap assumes the mastery cutoff percentage to be 70%. This can be reset quickly, as shown below:

```
*col (c2-c6)
*sub Mastery=80%, Title=(NewsQuiz)
*key ACCDB
```

The cutoff percentage has now been set to 80%. More generally, it is possible to have the default level of 70% set to any value by making a change in Lertap's System worksheet. It is also possible to set the mastery level at a raw test score. For example, Mastery=30 will set the cutoff at a test score of 30:

```
*col (c2-c6)
*sub Mastery=30, Title=(NewsQuiz)
*key ACCDB
```

(Click here to read more about the System worksheet, and click here to read a bit more about the use of the Mastery control word.)

Example C5:

This example reflects a common situation:

```
*col (c2-c6)
*sub Res=(1,2,3,4), Title=(NewsQuiz)
*key 13342
```

The RES control word is telling Lertap that the item response codes are digits, not letters. Unless you tell it otherwise, Lertap assumes that cognitive items have four options, with response codes of {A, B, C, D}. If this is not the case, you must use an Res=() declaration on a *sub card, as exemplified above. Note that the *key card has been changed -- if the response codes were digits, then the *key card will give the digit corresponding to the right answers. (Also note: RES= is the same as Res=, which is the same as res=, which can even be the same as Responses=;

Lertap really only looks at the first letter of the control words, and it doesn't care if letters are upper or lower case.)

Here are some other examples of valid Res=() declarations:

```
Res=(T,F)
```
(The subtest's *key card must contain Ts and Fs.)
```
Res=(A,B,C,D,E,F,G,H,I,J)
Res=(a,b,c,d)
```
(The subtest's *key card must contain lower-case letters.)
```
Res=(1,2,3,4,5,6)
```
(The subtest's *key card must have digits.)

~~Res=(A,B,C,D)~~

(Not required! This is the default setting for cognitive items.)

Critical note: the response codes seen in the Res= declaration tell Lertap what to look for when it reads the information in the Data worksheet's rows. If the response codes are upper-case letters, such as {A,B,C,D}, then Lertap will expect to find upper-case letters in the relevant columns of the Data worksheet. Nasty things can happen when, for example, the item responses seen in Data columns are lower-case letters, such as {a,b,c,d}, and the *sub card has Res=(A,B,C,D). This is a mis-match. Res=(A,B,C,D) tells Lertap to look for upper-case letters, but none will be found. Things will come a-crashing. (There's a bit more on this towards the end of the CCs sheet topic.)

Example C6:

Here's one more example of the *sub card in action:
```
*col (c2-c6)
*sub Title=(NewsQuiz), CFC, Wt=.5
*key ACCDB
```
CFC means "correction for chance", another scoring option entertained by Lertap. This control word isn't used all that often; it usually results in penalising students if they appear to be guessing (see Chapter 10 of the manual for more discussion). The Wt= declaration applies when the CCs worksheet defines more than one subtest, that is, when there are two or more *col cards. In this case, Lertap will usually generate a total test score by summing the subtest scores; the Wt= assignment controls how this is done. If Wt=0 then the subtest will not be included in the total test score.

Critical note: Wt=1 is the default action -- if there's more than one subtest, and no Wt= assignment is found on *sub cards, then Wt=1 is assumed. And note: when using Wt=, it's best to put it at the end of the *sub line.

Now we will exemplify the use of the other control cards for cognitive tests.

Example C7:

We'll add an *alt card:
```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*alt CDDDC
```

The *alt card is optional. Here it's telling Lertap that the last response code used by the first and last items is C, whereas the last response code used by all other items is D. Since there is no explicit Res= declaration, Lertap assumes Res=(A,B,C,D). (Note that this format of the *alt card differs from that shown in the manual. The format changed way back in the year 2005. In the old format this *alt card would have been *alt 34443.)

To read more about the practical effects of using *alt, please see the very end of this topic. A very practical real-life example may be enjoyed by having a look at this little paper.

Example C8:

We'll use a *wts card:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*wts 31121
```

The *wts card is optional. It indicates the number of points to be given for the correct answer, and it's only required when some of the items are worth more than one point. In this example, the correct answer to the first item, A, is worth 3 points, while the correct answer for the fourth item, D, is worth 2 points. All other items are worth one point.

If an item is to be worth more than 9 points, a *mws card has to be used. *mws cards are mentioned below.

Example C9:

Both *alt and *wts cards in use:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*alt CDDDC
*wts 31121
```

You understand this one, don't you? The right answer to the first item is A. It (the first item) uses three response codes, (A,B,C). A correct answer on the first item is worth 3 points.

Q: if I answer D on the fourth item, how many points do I get? Two.

What's the maximum score I can get over these five items? Eight.

If I answer C on the last item, how many points do I get? None; the right answer is B.

If I don't answer the third item, what happens? I get sent home early with instructions to have extra peanuts with my beer. (In truth: nothing. A non-answer to a cognitive item usually gets "scored" as a zero.)

Example C10:

Using Lertap's Big Gun, the *mws card:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*mws c2, 1, 0, 0, *
*mws c3, 0, 0, 1, 0
*mws c4, 0, 0, 1, 0
*mws c5, 0, 0, 0, 1
*mws c6, 0, 1, 0, *
```

This example is really the same as Example 9. We want to ease you into the idea of *mws cards by starting with an "easy" example.

Keep in mind that the default Res=(A,B,C,D) applies to this example, there being nothing to the contrary on the *sub card.

The *mws c2 card refers to the item whose responses are found in column 2 of the Data worksheet. This is, of course, the first item. Of the four potentially-possible responses to this item, (A,B,C,D), the *mws c2 card says that the first response is to get one point; the second and third responses are to get zero points, and the fourth response is in fact not used by this item -- hence the asterisk.

Look at the *mws cards above. They have the same format: the column number of the item in question, followed by the number of points corresponding to each of the item's response codes. If the item does not use one or more of the response codes, an asterisk is used.

Example C11:

More about the *mws card:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*mws c2, 1, 0, 0, *
*mws c6, 0, 1, 0, *
```

This example is the same as the last one.

We hear you saying "No it's not, come on now! The last example used five *mws cards; now there are only two".

Sure. You're right. What we should say is that this example accomplishes the same item scoring as the last example. Look at the three cards we've eliminated:

```
*mws c3, 0, 0, 1, 0
*mws c4, 0, 0, 1, 0
*mws c5, 0, 0, 0, 1
```

These cards say that the items found in columns 3, 4, and 5 of the Data sheet use all four response codes, have one correct answer, and award one point for the correct answer.

But this is the default. Lertap assumes all items will use all response codes, have one correct answer, and will award one point when the correct answer is selected. There's no need for *mws cards for these items -- their scoring is standard stuff.

So, what's special about the items in c2 and c6? They don't use one of the response codes. Now, this really isn't a big deal. Lertap would process the c2 and c6 items even if we didn't mention the fact that these items use just three response codes; Lertap's various reports would simply show that the fourth option, with a response code of "D" in this case, was not selected by anyone, and the Stats1b report would flag "D" as a poorly-performing distractor. Such things as test scores and coefficient alpha will not be not affected.

But why not do the job right? Lertap allows items to have a different number of options. The *alt card and the *mws card both allow you to set the record right, to inform Lertap that some items do not use all of the subtest's response codes. Use these cards and Lertap's reports will look a bit cleaner.

Example C12:

Still more about the *mws card:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*alt CDDDC
*mws c4, 0.5, 0, 0.5, 0
```

This example is quite typical. The item whose responses are coded in column 4 of the Data sheet is now being double-keyed. If someone selects the first response they get half a point. And, if someone selects the third response, they also get half a point. There are two "right" answers, each worth half a point. This example exemplifies one way in which partial credit may be addressed in Lertap.

```
*mws c4, 1, 0, 1, 0
```

Here again there are two right answers, but now they're each worth one point.

```
*mws c4, 0.50, -0.50, 2.00, -0.75
```

Things are getting real fancy now. The best answer is the third one, for which a whopping two points are awarded. The first answer is worth half a point. The second and fourth answers now have negative scoring weights; a person selecting the second option loses half a point, whereas someone going for the fourth option will lose three-quarters of a point.

Example C13:

Some buildings do not have a 13th floor, and we don't have a 13th example, either.

Example C14:

Re-scoring all items at once:

```
*col (c2-c6)
*sub Name=(Followup TV9 news quiz), Title=(NewsQuiz)
*key ACCDB
*mws call, 1, 0, 1, 0
```

The items found in all of the subtest's columns are to be scored with one point for the first and third responses, with no points for the second and fourth responses.

This sort of scoring is not at all common for cognitive items, not at all -- but if you want to do it, you can.

In the world of Lertap, *mws cards are the most potent cards going. They're dynamite. They completely override whatever information has come on preceding CCs cards.

A special form of the *mws card may be used when it's desired to quickly remove an item from a subtest. Click here to read about it.

In Lertap Version 5.25, another special form of the *mws card was introduced. It has this form:

```
*mws c12, 0, 1, 0, 1, other=1
```

To give credit to everyone for an item, even if they didn't answer the item, a card such as the following might be used:

```
*mws c12, 1, 1, 1, 1, other=1
```

The card above gives one point for each of the item's permitted answers, and it even gives people one point if they didn't answer the item.

Click here to read more about "other".

There are indeed times when, as in Example 10, a *mws card is used for each item. It may be only 2% of the Lertap-using world which will have an example of this sort, but it does happen -- we've seen it. In such a case, does the *key card make sense? No. But Lertap requires each and every cognitive subtest to have a *key card, so put one in (please).

Example C15:

Items which require multiple responses. Not covered above is the matter of scoring cognitive test items that require students to select more than one response in order to get an item correct. The example we have for this is not here, not in this topic. But we do have it covered, and invite you to click here to see our example.

Keep in mind that the manual has three chapters on CCs cards. Between what's written there, and what's appeared in this help topic, we hope you'll have an adequate to good grasp of CCs cards. But drop us a note if you've got questions: lertap5@gmail.com.

3.2 Affective CCs

This is the all-you-ever-wanted-to-know page about affective control "cards". First, some terms:

Let's say you had a couple of Likert-style items like these 'uns:

1) West Australian beaches are unsurpassed in the whole world.

    1 Strongly disagree
    2 Disagree
    3 Undecided
    4 Agree
    5 Strongly agree

2) The beaches of Maui are better than West Australia's.

    1 Strongly disagree
    2 Disagree
    3 Undecided
    4 Agree
    5 Strongly agree

Both of these items have five possible responses, or options, or alternatives, and use response codes of {1,2,3,4,5}. It's possible to have items with more options; Lertap allows up to 10 options per item. Items do not have to use the Likert style. The response codes used to not have to be digits (examples below).

Likert-style items are very common; another popular style is the semantic differential.

If we at Lertap central wanted to "score" these items, we'd be content to follow the conventional pattern of letting "strongly disagree" equal one point, "disagree" two points, ..., and "strongly agree" five points.

Respondents could end up with a low total "score" of two (2) points, and a top total "score" of ten (10) points. They'd get the low "score" if they answered "strongly disagree" on both items. They'd get the top "score" if they chose "strongly agree" on both items.

It's not really necessary to think about "scores" such as these, but some people find them useful. And, if they do, and if they work for WATC, the West Australian Tourist Commission, they'd likely want to reverse the scoring for the second item.

Say what? Reverse the scoring? You bet; it's a common happenstance. The WATC mob would want people to strongly disagree with the second item above. They'd give "strongly disagree" a "score", or weight, of 5, "disagree" a weight of 4, ..., and "strongly agree" a weight of 1. This way the top scores will come from people who not only love WA's beaches, but think they're better than those found around Maui.

Okay then, we've got some basic terms under the belt. Of course, if we're on one of those beaches, we may not have a belt to put them under, but let's proceed anyway.

Say we had five items of the sort shown above. Say we asked 200 WA-based people to respond to the items, and took the trouble to fly to Hawaii, paddle out to Maui, and ask another 200 folks to respond to the same five questions. We entered the 400 responses into a Lertap Data worksheet, with some sort of ID code in column 1, a location code of "W" or "M" in column 2, and the answers to the five questions in columns 3, 4, 5, 6, and 7. Were we to look down columns 3 through 7, we'd see 1s, and 2s, and 3s, and 4s, and 5s, corresponding to answers of "strongly disagree" through to "strongly agree".

Having entered the data, we need to go to work in Lertap's CCs worksheet.

Example A1:

We started with this set of two CCs cards:

```
*col (c3-c7)
*sub AFF
```

These two simple cards are all that some people might use to process the results. The *col card gets Lertap and Excel to read information from the Data worksheet, looking at five columns, 3 through 7. The *sub card has the AFF control word -- this is necessary in order to get Lertap to process the items as affective ones; without this card, and without the AFF control word on it, Lertap would have tried to process the items as cognitive ones, and would have wanted to find a *key card with the right answers to each item.

Lertap will "score" each item, giving a weight of 1 every time it encounters a 1 in the item's Data column, a weight of 2 for a response of 2, and so on. The minimum score on any item is 1, the maximum is 5.

Lertap will also make a "subtest score", or "scale score", for each respondent, placing these in its Scores worksheet. Such scores are just the sum of the item scores. Since there are five items, the minimum possible subtest score would be 5; the maximum would be what? Yes, 25.

Not everyone is interested in these scores. Many people are, but some aren't. Lertap makes them, and you'll just ignore them if you don't like to score affective items.

Example A2:

Example A1 was too simple if we're interested in scores. You've already heard that the second item was to have reverse scoring; you didn't know it, but the third item was also to be reversed. We have need for a *pol card:

```
*col (c3-c7)
*sub AFF
*pol +--++
```

The *pol card is usually optional for affective subtests. It's hauled out when some of the items are to be reverse-scored. The "pol" stands for "polarity".

The *pol card above tells Lertap that the second and third items are to be reverse-scored -- that's what the minus (-) signs mean. The little plus (+) signs tell Lertap to score items 1, 4, and 5 in the usual, "forward", manner.

Example A3:

Now we'll put in some labels to grace Lertap's reports:

```
*col (c3-c7)
*sub AFF, Name=(Beach survey 1), Title=(Beachin)
*pol +--++
```

The Name shown above will appear at the top of some of Lertap's reports, such as Stats1f and Stats1b, while the Title will show up at the top of one of the Scores worksheet's columns. The Name can be any length; Title should be kept to 8 characters or less. If Name and Title are not given, Lertap defaults to Name=(Test 1), Title=(Test1).

Example A4:

Next we'll add some more control words to the *sub card, and then explain what they accomplish:

```
*col (c3-c7)
*sub AFF, Title=(Beachin), PER, SCALE
*pol +--++
```

The PER control word prompts Lertap to compute a percentage score for each respondent, being his or her score expressed as a percentage of the maximum possible score. For example, on our little test of five questions, the maximum possible score was 25; someone with a score of 15 would get a PER score of 60%.

SCALE gets Lertap to "normalise" the scores. It divides each person's score by the number of items. This is best used when all the items have the same number of options; it results in a score scaled back to the scores used at the item level.

For example, let's say someone scored 10 on our 5-item test. Divide this score, 10, by the number of items, 5, and SCALE=2.00 for this person. This might make us think that the person's "average" response to our items was "disagree", since "disagree" had a scoring weight of 2.

SCALE can be handy when processing an affective instrument with numerous subscales, each scale having a different number of items. If SCALE is used on each respective *sub card, then we can scan any person's SCALEd scores and quickly see their "average" positions on the 1-to-5 scale used to score each item. (The manual has a real-life example of just such a situation: see Chapter 8, where the MSLQ instrument is discussed.)

Example A5:

Now we will toss in two more control words for the *sub card:

```
*col (c3-c7)
*sub AFF, MDO, Title=(Beachin), Wt=0
*pol +--++
```

The MDO control word will get Lertap to turn off its missing data option. Lertap makes a standard response substitution when people don't answer an item: it gives them an item score equal to the mean of the item weights (note). In our example, the item weights range from 1 to 5; the mean is 3. A person not answering an item gets a score of 3; someone declining to answer all items would get a score of 15 on our 5-item scale.

To turn off this automatic substitution, use MDO on the *sub card, as exemplified above. To have more control over how missing data are processed, use *mws cards with an "other=" declaration, as mentioned below, under Example A11.

Revision note: the way the MDO option works has changed. Now MDO effectively means "missing data out"; when MDO is present, subtest statistics are adjusted so that they exclude people who haven't answered items. This is further discussed in the following topic.

The Wt=0 declaration shown above says to Lertap: "As you go about summing all the subtest scores to make a total score, give this subtest a weight of zilch (zero)."

What's that, you say?

Well, whenever a CCs worksheet has more than one *col card, we're into a situation where there are multiple subtests. We are? Sure: each *col card defines a new subtest.

Granted, there aren't multiple *col cards in this example, but pretend there were.

Lertap's standard modus operandi is to add all the subtest scores together, making a total score for each person. Usually this total is just the sum of the subtest scores, much as if Wt=1 had been used on each subtest's *sub card. The Wt= declaration gives you the ability to control the way a subtest adds to the total score. Since Wts can be negative, you can even get a subtest's result to be subtracted from the scene.

How you doing? This is all pretty straightforward, is it not? We'll step up the tempo a bit, and get into some more advanced matters.

Example A6:

We've used an Res= declaration below, have a look:

```
*col (c3-c7)
*sub AFF, Res=(A,B,C,D,E), Title=(Beachin)
*pol +--++
```

Lertap assumes that affective items have five options, with the response codes corresponding to the options being (1,2,3,4,5). The Res= declaration on a *sub card tells Lertap this, but, when Res=(1,2,3,4,5), there's no need to explicitly say so. Lertap assumes Res=(1,2,3,4,5) for affective items. (For cognitive items, Lertap assumes Res=(A,B,C,D).)

If this isn't the case, an Res= declaration is required. In this example, the five affective items have five options, with the response codes for the items being (A,B,C,D,E). It's as if the items had this sort of format:

4) A few Emu Exports a day keep the doctor away.

A Strongly disagree
B Disagree
C Undecided
D Agree
E Strongly agree

If you had no trouble with this example, try the next one. (Sooner or later we'll stump you.)

Example A7:

Here's a big example, getting into the real intricacies of Lertap and item weighting:

```
*col (c3-c7)
*sub AFF, Res=(5,4,3,2,1), Title=(WAuni)
*pol +--++
```

Here we're still in a situation where the items have five options, and the response codes are back to (1,2,3,4,5). However, now a response of 5 is to get a weight of 1, a response of 4 a weight of 2, ..., a response of 1 a weight of 5.

It's as if the items looked like the one below, where the Likert scale has been re-ordered so that "strongly agree" is the first option (very common):

5) West Australian universities are truly world class.

1 Strongly agree
2 Agree
3 Undecided
4 Disagree
5 Strongly disagree

Those of us based in West Australia would want this item to be scored in a manner which gives the highest score to the first option. Ordinarily, the first option gets the lowest weight. We could use a *pol card to reverse this, and most Lertappers probably would (including us). However, we said we'd step up the tempo; we've started to wade into deeper water.

An understanding of this example gets into the very basics of Lertap's affective item scoring. The default scoring weight applied to any particular response code corresponds to the response code's ordinal position in the Res= declaration. If 1, or A, or 5 is the first response code to appear in Res=, it gets a scoring weight of 1 (but let's say 1.00 to emphasize that we're talking about a real number).

If Res=(1,2,3,4,5), scoring weights are 1.00, 2.00, 3.00, 4.00, and 5.00.

If Res=(A,B,C,D,E), scoring weights are 1.00, 2.00, 3.00, 4.00, and 5.00.

If Res=(5,4,3,2,1), scoring weights are 1.00, 2.00, 3.00, 4.00, and 5.00.

The default scoring weights have <u>no</u> correspondence to the actual response codes. They're based entirely on the ordinal position of the response codes in the Res= string. The entries in the Res= declaration are never read as numbers. Never -- they're just characters.

If Res=(5,6,7,8,9), scoring weights are 1.00, 2.00, 3.00, 4.00, and 5.00.

If Res=(1,2,3), scoring weights are 1.00, 2.00, and 3.00.

If Res=(w,x,y,z), scoring weights are 1.00, 2.00, 3.00, and 4.00.

Example A8:

The stage is set; we now introduce the most powerful card in Lertap's mighty arsenal: the *mws card.

Example A2 above looked like this:

```
*col (c3-c7)
*sub AFF
*pol +--++
```

We'll knock out the *pol card, and instead go with these CCs cards:

```
*col (c3-c7)
*sub AFF
*mws c4, 5.00, 4.00, 3.00, 2.00, 1.00
*mws c5, 5.00, 4.00, 3.00, 2.00, 1.00
```

The *mws cards explicitly apply scoring weights for the two items whose responses are found in columns 4 and 5 of the Data worksheet. They're saying that the first response code in the Res= declaration is to get a weight of 5.00, while the last response code is to get a weight of 1.00.

There isn't an Res= declaration on the *sub card, is there? No. In the absence of one, Lertap assumes Res=(1,2,3,4,5).

The two *mws cards have reversed the scoring for the second and third items, respectfully located in column 4 (c4) and column 5 (c5) of the Data worksheet.

Why aren't there *mws cards for the items in c3, c6, and c7? There really are; if you could look deep inside the heart of Lertap, you'd see that it's effected these statements:

```
*mws c3, 1.00, 2.00, 3.00, 4.00, 5.00
*mws c6, 1.00, 2.00, 3.00, 4.00, 5.00
*mws c7, 1.00, 2.00, 3.00, 4.00, 5.00
```

The weighting pattern seen in these three cards is the default pattern for items with five response codes. By "default" is meant "not requiring mention; this is what I'll do unless you tell me otherwise". This being the case, we don't need to say anything. Let Lertap apply its default weights for each item unless we say different. Use *mws cards to "say something different".

Now, in this case, we have used two *mws cards in place of one *pol card. That's not real efficiency. The majority of Lertap users are happy with the *pol card, but there are some advantages in using *mws cards. One advantage is that there's no ambiguity with *mws cards -- they make it absolutely clear how item responses are to be scored, and explicitly indicate the items affected.

About that idea of looking "deep into the heart of Lertap" ... a Lertap workbook's Sub worksheets provide quite a detailed glimpse of exactly how item weights have been set up. To read more about Sub sheets, just click here.

Example A9:

Here's an example which shows off a special use of the *mws card:

```
*col (c3-c7)
*sub AFF
*mws call, 5.00, 4.00, 3.00, 2.00, 1.00
```

'*mws call' means we have a multiple-weights specification which is to apply to <u>all</u> the items mentioned on the preceding *col card. This use of the *mws card often comes into play when there are items such as the following:

5) Mt Barker wines are as fine as those of California's Napa Valley.

1 Strongly agree
2 Agree
3 Undecided
4 Disagree
5 Strongly disagree

The *mws call card will reverse the default scoring weights, equating the first response code, 1, with 5.00 points, and the last response code, 5, with 1.00 points.

Astute readers might note that the scoring accomplished by the three CCs cards of this example is the same scoring achieved in example A7. In A7 we used a special Res= declaration to do the job, whereas now we're using *mws call.

Examples A7 and A9 will result in the same item scoring, but Lertap's Stats1f and Stats1b will differ. One could see the differences by running with the following set of CCs cards:

```
*col (c3-c7)
*sub AFF, Res=(5,4,3,2,1), Title=(WAuni)
*pol +--++
*col (c3-c7)
*sub AFF
*mws call, 5.00, 4.00, 3.00, 2.00, 1.00
```

We've combined the two examples, A7 and A9, making two subtests from the same items. Now the two full-statistics reports, Stats1f and Stats2f can be compared; item and test stats will be the same, but the order in which item response codes are listed will differ. The same will hold for the Stats1b and Stats2b reports.

Example A10:

Item scoring weights can be any real number, positive or negative:

```
*col (c3-c7)
*sub AFF, Res=(1,2,3,4,5,6,7), Title=(CapesAtt)
*mws call, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
```

This type of weighting is sometimes seen when semantic differential items are used, as in the following example:

8) The weather in the southwest capes region of West Australia is:

rotten	__	__	__	__	__	__	__	wunderbar

There are seven blanks above. If someone clicks the fourth blank, the centre one, their response will be entered in the Data sheet as a 4, and the *mws card above will have their response scored as 0.00 points. Selecting the first blank will see a 1 entered on the Data sheet (because it's the first blank), and a weight of -3.00 points will apply.

How many points if someone ticks the sixth blank? 2.00.

Example A11:

Enter the *alt card:

```
*col (c3-c12)
*sub AFF, Res=(1,2,3,4,5,6,7), Title=(CapesAtt)
*alt 55555 77777
```

To try and explain this set of cards, we ask you to imagine that we've given a survey with ten items. The first five items are of the Likert style, and use response codes of (1,2,3,4,5). The last five items employ the semantic differential style, and use response codes of (1,2,3,4,5,6,7).

The *alt card tells Lertap this. It says that the first items use the first five response codes seen in the Res= declaration, while the last five items use seven of the response codes (which of course is all of them). It's still possible to add a *pol card in cases like this, as shown here:

```
*col (c3-c12)
*sub AFF, Res=(1,2,3,4,5,6,7), Title=(CapesAtt)
*alt 55555 77777
*pol +--++ +++++
```

But there's a probable problem here. Users of semantic differential items often like to have negative weights corresponding to the negative side of their questions, as seen above in Example A10. The set of four CCs cards above is not right; the five semantic differential items will have scoring weights of 1.00, 2.00, ..., 7.00. We need some *mws cards:

```
*col (c3-c12)
*sub AFF, Res=(1,2,3,4,5,6,7), Title=(CapesAtt)
*alt 55555 77777
*pol +--++ +++++
*mws c8,  -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c9,  -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c10, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c11, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c12, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
```

Of course, we could have used *mws cards for all items, and, in this case, we'd then have:

```
*col (c3-c12)
```

```
*sub AFF, Res=(1,2,3,4,5,6,7), Title=(CapesAtt)
*mws c3,   1.00,  2.00,  3.00, 4.00, 5.00, *, *
*mws c4,   5.00,  4.00,  3.00, 2.00, 1.00, *, *
*mws c5,   5.00,  4.00,  3.00, 2.00, 1.00, *, *
*mws c6,   1.00,  2.00,  3.00, 4.00, 5.00, *, *
*mws c7,   1.00,  2.00,  3.00, 4.00, 5.00, *, *
*mws c8,  -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c9,  -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c10, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c11, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
*mws c12, -3.00, -2.00, -1.00, 0.00, 1.00, 2.00, 3.00
```

We know you can just about understand these cards, but realise you might want to know what the asterisks are doing on the *mws cards for the items in c3 to c7. The asterisks tell Lertap that the respective items do not use the last two response codes seen in the Res= declaration.

It is possible to use special forms of the *alt and *mws cards when it's required to quickly remove an item from a subtest. These forms can be very handy in certain circumstances; please refer to the 'Remove an item' topic for examples.

In Lertap Version 5.25, it became possible to use the "other=" control word on a *mws card, as exemplified below:

```
*mws call, 1, 2, 3, 4, 5, other=9
```

The line above says that anyone not selecting one of an item's five permitted responses will get a score of 9 points. The use of other= presents a more flexible way to handle missing affective data than does the MDO option discussed above.

For more about using the other= control word, and item weighting, please click here.

Example A12:

Just one more example, and this one's simple:

```
*tst c2=(W)
*col (c3-c7)
*sub AFF, Name=(Beach survey 1), Title=(Beachin)
*pol +--++
```

This example displays use of the *tst card. The *tst card is telling Lertap to make a new Lertap workbook, one whose Data worksheet will contain copies of only those Data records which have a "W" in their second column. We'd probably then set up another Lertap analysis with these cards:

```
*tst c2=(M)
*col (c3-c7)
*sub AFF, Name=(Beach survey 1), Title=(Beachin)
*pol +--++
```

Now Lertap is being told to make a new workbook whose Data records will be only those with an "M" in column 2. We have now set up two new data sets, two new Lertap workbooks, one for the respondents from West Australia, and one for those from Maui.

(For a more thorough discussion of using the *tst card, and breaking out subsets of records, see the section titled "A survey with multiple groups" in Chapter 8 of the manual.)

Good going; you stuck it out and have come to the end of this topic / page. Lertap provides extensive support for different item scoring methods, as we've exemplified above. Of course there's more material in the manual, where three chapters are devoted to a discussion of CCs cards.

Still, a few years of experience have shown us that the matter of item scoring in Lertap at times creates questions among users, no matter how much documentation we might provide. Should you have questions, whiz off a note to us: larry@lertap.com.

3.3 Missing data

Update 30 June 2022: substantial improvements have been made in Lertap's missing data capabilities. These include having cells with missing data highlighted, and the use of "NA" missing data codes. See this document.

In Lertap, "missing data" is a term which usually means that a person has not answered, or has omitted, an item.

If you look at a Lertap Data worksheet and see empty cells, such cells usually indicate that there was no answer to the item corresponding to the cell. In the screen snapshot below, Virgo has not answered Item 2, nor Item 7; Westphal has left three items unanswered (2, 5, and 7); Xeno did not answer Item 2; Yalso has not answered Item 7.

	1	2	3	4	5	6	7	8	9	1(
1	An example from the Lertap 3 manual (1983, Table 3-1).									
2	No.	ID	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item
10	8	Smith, S	B	B	B	D	D	C	C	D
11	9	Terace, T	B	B	A	D	B	C	B	C
12	10	Uptown, U	B	D	B	C	A	B	C	C
13	11	Virgo, V	D		B	A	D	C		C
14	12	Westphal, W	B		A	A		C		A
15	13	Xeno, X	B		B	D	A	C	C	A
16	14	Yalso, Y	A	C	A	D	B	C		A
17	15	Zenu, Z	A	D	A	A	A	C	A	C

There are quite a number of users who prefer to use a special code for the case of unanswered items. Instead of leaving a cell empty, they might use a 9 to indicate missing data (there are historical reasons for this; some of the most popular data analysis programs have traditionally used 9s to represent missing data). When using scanners to process mark-sense answer sheets, the software driving the scanner may have its own missing data code, such as an asterisk (*).

To fully understand how Lertap processes missing data, it helps to have a good understanding of what Lertap calls "response codes".

The letters (or digits) seen under a Lertap Data worksheet's item columns represent response codes. In the example above, it seems that items have used response codes of A, B, C, and D; this set, {A,B,C,D}, is, in fact, Lertap's default response code set for cognitive items -- for affective items, the default response code set is {1,2,3,4,5}.

It is common for users to have items which use other response codes. Whenever a test does not use the default response codes, the response codes used by that test's (or survey's) items are specified by using an Res= declaration on a *sub card.

For example, Res=(A,B,C,D,E,F), Res=(1,2,3,4,5,6,7), and Res=(t,f) are all valid Res= declarations. In Lertap, a test may use up to 26 response codes.

Now, with this understanding of response codes in hand, the definition of missing data can be made a bit more precise: in Lertap, an item response is said to be missing whenever a cell in the Data worksheet has an entry which does not match the response codes used by the item. When this happens, Lertap says it has encountered an "other" response.

As Lertap goes about tallying item responses, it keeps track of the number of "other" responses in a special bin.

What's this bin called? The "other bin", naturally. Deep inside Lertap, each and every item is assigned a storage bin for "other" responses. The contents of an item's "other" bin are displayed in many of Lertap's reports.

As an example, look at these snapshots:

Q11

option	wt.	n	p
A	1.00	28	0.47
B	0.00	0	0.00
C	0.00	13	0.22
D	0.00	13	0.22
E	0.00	0	0.00
other	0.00	6	0.10

Stats1f / Stats1b / Stats1ul / Stats2f

Q35

option	wt.	n	%
1	1.00	2	3.3
2	2.00	13	21.7
3	3.00	12	20.0
4	4.00	17	28.3
5	5.00	7	11.7
other	3.00	9	15.0

Stats2f / Stats2b / Stats2ul / **Stats3f** / Stats3b

In these two examples, the "other" bin for Q11 has n=6, meaning that there were six people missing data on the item. For Q35, there were n=9 people missing data.

So. Are you now full bottle on Lertap's definition of missing data? And an expert on Lertap's "other" bin?

Goodonyamate. But wait, there's more ....

What this stuff under the "wt." column? The answer to this question is also very relevant to understanding how Lertap processes missing data, especially for affective tests.

The "wt.", for "weight", indicates how many points are associated with each item option.

On Q11, the 28 people who chose response code A (usually referred to as option A, or alternative A) will get 1.00 points for their answer. Choosing any of Q11's other options gets no points. You guessed it: Q11 is a cognitive item whose right answer is A.

Not answering Q11 gets no points. Not answering Q11 gets no points. We have repeated this as it's important: an item's other bin can have scoring points attached to it.

Look at Q35, an affective item. Each and every response code has associated scoring points, and, should someone not answer Q35, they will still get 3.00 points (!).

What we've displayed in these two examples represents Lertap's default handling for cognitive and affective missing data. A cognitive item's default missing data scoring action is 0.00 points; an affective item's default missing data score is equal to the mean of the other scoring wts.

With this in mind, take a break. When you return, have a look at the following topics to read more about Lertap and missing data.

3.3.1 Did-not-see option

Options have been added to the [System worksheet] which allow users to adjust cognitive item statistics so that they are based only on those cases (students) that had a chance to answer the item.

1	These are Lertap5 system settings. Change them only if you understand them.	System Settings		
2		**Present setting:**	**Allowed settings:**	**Usual setting:**
40	Use a **did-not-see** code?	yes	yes / no	no
41	Did-not-see code (single character; may be blank):		any char	
42	Create an adjusted percentage score?	no	yes / no	no

The did-not-see option is activated by placing "yes" in column 2 of the appropriate row in the System worksheet; in the example above, this is row 40.

Use of this option requires a corresponding did-not-see code, a single character recorded in the Data worksheet whenever a student did not have the chance to answer an item (for whatever reason). In the example above, the code is a blank. The code may be any single character, including a blank, an upper-case letter from the Latin alphabet (ABC...XYZ), a lower-case letter (abc...xyz), or a single Arabic numeral (0123456789).

How does Lertap process item responses with this new option? Well, for any item, Lertap first looks to see if a student's response corresponds to one of the response codes used by the item's options. If it does not, then Lertap looks to see if the "response" matches the did-not-see code, assuming the did-not-see option has been activated. If it does not, then the response is classed as "other", a response category often referred to in Lertap documentation as "missing data".

Lertap will automatically adjust all of its item statistics so that they exclude did-not-see cases (that is, of course, assuming the did-not-see option has been activated -- if the option has not been activated, then what would have been a did-not-see response will be included in the "other" response category, and processed as missing data).

Use the [Freqs worksheet] to check on missing data and did-not-see cases, and note: if the did-not-see code is a blank, Freqs will include did-not-see cases in its "other" line.

To see what the various "Stats" reports look like when the did-not-see option is on, just page forward to the following topics.

Scoring

Whenever the did-not-see option is in effect, users have the option to have Lertap calculate two scores for every person on each subtest defined in the CCs worksheet:

the usual subtest score, and a percentage score which is based on the maximum possible score which could have been obtained on the items actually presented.

The first score, the "usual" subtest score, will be the number of points earned on the number of items the person had the opportunity to answer. In the cognitive case, the most common scoring method is to give one point for each correct answer -- in this case, the "first score" will simply be the number of items right.

The second score is only produced when "yes" is found in the System worksheet in the second column of the line which says "Create an adjusted percentage score?". This score is formed by dividing the first score, the "usual" subtest score, by the maximum possible score which the person could have had if s/he had correctly answered every presented item; this figure is then multiplied by 100 to get a percentage index.

Coming back to the common cognitive case, with one point for each correct answer, the second score is the number of answers right divided by the number of items presented, multiplied by 100.

Note that both of these scores exclude items which the person did not see, that is, did not have the opportunity to answer.

It is not necessary to use the PER control word on the *sub card in order to get the second score: it's computed automatically, providing "yes" is found in the second column of the line which says "Create an adjusted percentage score?".

Adjusted?

When would you want to have the percentage scores "adjusted"? What does "adjusted" mean, anyway?

Well, it turns out that certain online test generators can create tests of variable length; some students might see a 48-item test, others a 45-item test, and still others a 50-item test. A student's test score will depend on the number of items s/he had the chance to answer. Two students with the same "usual" subtest score, say 35, will not have the same percentage score if they have been presented with a different number of items.

For example, if Jorge got a score of 35 on a test with 50 items, his percentage score would be 70%, assuming one point for each correct item. Suppose it turned out that Marisol also got a score of 35, but she was presented with only 45 items; her rounded percentage score will be 78%.

When you've asked Lertap to "Create an adjusted percentage score", the percentage test score will be adjusted according to the items a student actually saw.

If all students get the same number of items, there is no need to create an adjusted percentage score. Ah, but wait a minute! We should adjust this comment: if all students are presented with the same number of items, an adjustment is not required *if all items are scored the same way*. When items are scored in different ways, then the adjustment might again be useful. For example, if Item 26 is worth one point, but Item 27 is worth two points, then we might well want to make an appropriate adjustment to scores, depending not only on how many items a student saw, but also on the number of points the items are worth.

Lertap creates the adjusted score by dividing the "usual" subtest score by the actual maximum possible score a student could have achieved, based on the items presented to the student.

Sample Scores output

As you page forward to following topics, you'll come across an example from an actual online testing situation, one which used the Test Pilot system from McGraw-Hill.

Over 400 students at a large North American university took an online version of a test delivered by Test Pilot. Sampling from a pool of 80 cognitive multiple-choice items, Test Pilot served up tests whose lengths varied: some students were presented with 40 items, some with 43, some 45, and others 48. (Please read the second "tidbit" below -- the situation was actually a bit more complicated.)

The Lertap Data worksheet turned out to have 93 columns for item responses. Why 93 when there were only 80 items in the pool? Because a few items allowed for multiple responses (see second tidbit below).

In this example, blanks were used as the did-not-see code. Were you to look at the item responses for any student, scanning from left to right over the 93 columns allocated, you'd see actual responses, many blanks, and, for those items seen by a student but not answered, a 9.

We could, and will, entertain a variety of Lertap scoring "methods" which will demonstrate how the did-not-see options affect test scores.

To begin, say we have the did-not-see option off, that is, we have "no" in the System worksheet row which says "Use a did-not-see code?". Also, assume we have the PER control word on the *sub card.

The Scores output will look as follows:

Record No.	EE101	EE101%
1	38.00	40.9
2	30.00	32.3
3	27.00	29.0
4	29.00	31.2
5	21.00	22.6

The first student had a test score of 38. Lertap says that the corresponding percentage score is 40.9, a value found by dividing the score, 38, by 93, the maximum possible test score if each item is scored giving one point for a correct answer.

But this is not correct in this case; no student had the opportunity to answer 93 items -- Test Pilot gave each student a random sample of approximately 40 items.

So, we activate the did-not-see option by putting "yes" in the second column of the System worksheet row which says "Use a did-not-see code?". We have "no" in the "Create an adjusted percentage score?" row. And, we still have PER on the *sub card. Our Scores now look as follows:

Record No.	EE101	EE101%
1	38.00	79.2
2	30.00	62.5
3	27.00	56.3
4	29.00	60.4
5	21.00	43.8

See how the percentage scores have changed? Lertap is now basing its percentage score on the maximum possible score which could have been earned on the actual items presented to the student. For the first student, the maximum was 48.

Dividing 38 by 48 and multiplying by 100 gives the 79.2 seen as the EE101% score for the first student.

Now, for the second student, does 30 divided by 48 equal 62.5? Yes. The problem is that the second student was presented with 40 items, not 48. The percentage score is wrong.

Whenever students are presented with a different number of items, or whenever items in the pool have different scoring patterns, "yes" should be used in the "Create an adjusted percentage score?" row.

So, let's see what happens. We put in that "yes". Do we still have PER on the *sub card? No, now it's not needed -- there would be no harm in having it there, but it is not required. Here are the scores:

Record No.	EE101	EE101%
1	38.00	79.2
2	30.00	75.0
3	27.00	67.5
4	29.00	72.5
5	21.00	52.5

The percentage scores above have each been adjusted, according to the maximum possible score a student could have obtained on the set of items s/he was presented with.

Yes, Miss? (A young woman in the fortieth row of the third balcony has her hand up.)

Why wouldn't I just use the adjusted percentage score option all of the time? Why not put that yes where it's supposed to be, and just leave it there for always?

A good question, thank you. The answer: Lertap pinches extra memory from the computer whenever it has to keep track of the maximum possible score each student could have achieved given the items presented. You save memory space, and also a tiny bit of processing time, by not adjusting the percentages. If you know that each student was presented with the same number of items, and all items were scored the same way, then say "no" to the "Create an adjusted percentage score?" option. Now you know, no?

Related tidbits:

Several learning management systems support the development and delivery of online cognitive and affective tests. Some, such as Angel Learning, can be made to randomly sample items from a database, presenting different students with different versions of a test, each with the same number of items. In classical test theory, when certain conditions have been met, these versions might be termed parallel forms, or equivalent forms. The data file of item responses created by Angel may not indicate which test "form" a student took. Fortunately, the Angel output is padded with did-not-see codes so that each data record has the same length, a length equal to the total number of items in the database.

In the Test Pilot example discussed above, the test included a few items which used a "Check all of the following options which would be correct" format, effectively turning a single item into multiple true-false items. Test Pilot actually served up the same number of items to each student (40), but those items which used the "check all of the following" format did not consistently offer the same number of options. The practical end result was as described above: students received a variable number of test items.

Did-not-see data will affect the calculation of alpha, Lertap's reliability coefficient -- alpha is NOT corrected for did-not-see cases.

3.3.2 MDO cognitive, Statsf

There's a control word, "MDO", which may be used on the *sub card to control how Lertap processes missing data. MDO may be used with both types of test, cognitive and affective. The letters stand for "missing data out", meaning that records with missing data are to be excluded from Lertap's various calculations.

This topic discusses the effect of the MDO and did-not-see options on the cognitive test reports produced by Lertap; a following topic does likewise for affective tests.

Look at these CCs lines for a cognitive test:

1	*col (c3-c27)
2	*sub Res=(A,B,C,D,E,F), Name=(Knowledge of LERTAP2), Title=(Knwldge), Wt=0
3	*key AECAB BEBBD ADBAB BCCCB BABDC
4	*alt CEDBC CEDFD EDCBD DCCDD DEEDF
5	*col (c3-c27)
6	*sub MDO, Res=(A,B,C,D,E,F), Name=(Knowledge with MDO), Title=(MDOKnwl), Wt=0
7	*key AECAB BEBBD ADBAB BCCCB BABDC
8	*alt CEDBC CEDFD EDCBD DCCDD DEEDF

These eight "cards" involve the same 25 items. Being the avid reader you are, you know you've seen these items before -- they're from the Lertap quiz.

The cards define two subtests. The only difference between the two is that the second one has the MDO option on; you can see it on the 6th line.

Statsf reports

A squiz of the Lertap's Stats1f and Stats2f reports will serve to highlight the effect of using MDO:

Lertap5 full item stats for "Knowledge of LERTAP2", created: 20/03/2006.

Q15

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	1	0.02	0.03	0.08	14.00	0.20
B	1.00	30	0.50	0.53	0.66	16.67	0.58
C	0.00	13	0.22	-0.36	-0.50	7.92	-0.68
D	0.00	16	0.27	-0.33	-0.45	8.81	-0.55

Q16

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	12	0.20	-0.14	-0.19	10.75	-0.27
B	1.00	36	0.60	0.46	0.58	15.53	0.42
C	0.00	3	0.05	-0.26	-0.56	4.67	-1.15
D	0.00	6	0.10	-0.29	-0.50	6.50	-0.88
other	0.00	3	0.05	-0.23	-0.49	5.67	-1.00

The report above comes from Stats1f. It summarizes the performance of two items, Q15 and Q16, using a variety of statistics.

Both items use four options, employing response codes {A,B,C,D}.

A student selecting option B will get 1.00 points towards her/his test score, as seen under the wt. column -- we'd conclude that both items have just one correct answer as all the other wt. values are zilch (zero).

Thirty (30) students got Q15 right. Sixty (60) students were involved, so p for Q15's option B is 0.50, which is, of course, 30 divided by 60.

The pb(r) and b(r) columns are, respectively, point-biserial and biserial correlation coefficients, indexing the relationship between option selection and the criterion score. If the item forms part of the criterion, as it does in this case, Lertap applies a part-whole correction, taking out the inflation the correlation coefficients would otherwise have (see the manual for a more extensive discussion).

The avg. column shows the average criterion score for those students selecting each item option. On Q16, the 36 students who selected option B had average criterion scores of 15.53. As a z-score, 15.53 is 0.42 (the manual has more to say; you ought to read it some day, perhaps when next at the beach).

Note that Q16 has an "other" line. Three students had missing data for Q16. They were weak students; their avg. was low, as it was for those who chose distractors C and D.

Okay? Got it? Good; now have a peep at the stats for the same two items after the MDO option has been used:

Lertap5 full item stats for "Knowledge with MDO", created: 20/03/2006.

Q15

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	1	0.02	0.03	0.08	14.00	0.20
B	1.00	30	0.50	0.53	0.66	16.67	0.58
C	0.00	13	0.22	-0.36	-0.50	7.92	-0.68
D	0.00	16	0.27	-0.33	-0.45	8.81	-0.55

Q16

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	12	0.21	-0.17	-0.24	10.75	-0.32
B	1.00	36	0.63	0.42	0.54	15.53	0.36
C	0.00	3	0.05	-0.28	-0.59	4.67	-1.20
D	0.00	6	0.11	-0.32	-0.54	6.50	-0.94
other	0.00	3	0.05	-0.25	-0.52	5.67	-1.06

There's no change in the stats for Q15; all 60 students answered that item. But Q16 has changed; many of Q16's stats above, from p through z, differ from the Q16 stats seen earlier. Why? Because the three students missing an answer to Q16 have been excluded from the calculation of the stats.

Look at the p column, for example. For Q16's option B, p is now 0.63, corresponding to 36 divided by 57, not 36 divided by 60.

The avg. values are now computed using just the criterion scores for the 57 students who answered Q16. This applies to the other line too: 5.67, as a z-score in the distribution of 57 scores, would be -1.06.

Now, take a few seconds and sum down the p column for Q16. In the first report above the sum is 1.00 (100%). But in the second report the sum comes to 1.05 (105%). What's up?

The 0.05 p value for Q16's "other" row is just indicating the proportion of respondents who did not answer the question. The other four p values for Q16 sum to 1.00, which is what is expected when the MDO option is in use.

Statsf reports and the did-not-see option

When you've got the did-not-see option going, the Statsf report will be quite similar to what you've just seen immediately above. Look:

Lertap5 full item stats for "Knowledge with DNSI", created: 24/03/2006.

Q15

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	1	0.02	0.03	0.08	14.00	0.20
B	1.00	30	0.50	0.53	0.66	16.67	0.58
C	0.00	13	0.22	-0.36	-0.50	7.92	-0.68
D	0.00	16	0.27	-0.33	-0.45	8.81	-0.55

Q16

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	12	0.21	-0.17	-0.24	10.75	-0.32
B	1.00	36	0.63	0.42	0.54	15.53	0.36
C	0.00	3	0.05	-0.28	-0.59	4.67	-1.20
D	0.00	6	0.11	-0.32	-0.54	6.50	-0.94
other			0.05				

There's just a wee difference in these results. Can you spot it? The other line is now missing most of its stats.

What about the joint operation of MDO and did-not-see. What happens then? Got time to look at results for two items delivered to 421 students over the internet, using the Test Pilot system from McGraw-Hill?

(c10) Q50

Option	n	/421
1	114	27.1%
2	94	22.3%
3	62	14.7%
4	42	10.0%
9	3	0.7%
?	106	25.2%

(c11) Q80

Option	n	/421
1	124	29.5%
2	196	46.6%
?	101	24.0%

In this example, a blank was used as the did-not-see code, and 9 as the code for missing data.

On Q50, three students failed to provide an answer, while 106 students did not see it (Test Pilot presented Q50 to 315 of the 421 students). Q80 was presented to 320 students, and they all answered it. Look at the corresponding Statsf summaries:

Lertap5 full item stats for "EE 101 T3 exam", created: 20/03/2006.

Q50

option	wt.	n	p	pb(r)	b(r)	avg.	z
1	0.00	114	0.37	-0.15	-0.20	27.37	-0.20
2	1.00	94	0.30	0.25	0.33	30.74	0.52
3	0.00	62	0.20	-0.15	-0.22	26.87	-0.31
4	0.00	42	0.13	-0.06	-0.10	27.57	-0.16
other	0.00	3	0.26	-0.10	-0.37	23.67	-0.99

Q80

option	wt.	n	p	pb(r)	b(r)	avg.	z
1	0.00	124	0.39	-0.27	-0.35	26.71	-0.34
2	1.00	196	0.61	0.18	0.23	29.47	0.22
other			0.24				

The statistics for both Q50 and Q80 have been computed by excluding the students who did not see the items, and by also excluding those with missing data.

Were you to sum the p values for Q50's four options, you'd have 0.37 + 0.30 + 0.20 + 0.13, or 1.00 (100%). The 0.26 on Q50's other line is the proportion of students who did not have an answer for the item, either because they did not see the item, or saw the item but did not answer it.

Let your mouse hover over the 0.26 value, and behold:

Q50

option	wt.	n	p	pb(r)	b(r)	avg.	z
1	0.00	114	0.37	-0.15	-0.20	27.37	-0.20
2	1.00	94	0.30	0.25	0.33	30.74	0.52
3	0.00	62	0.20	-0.15	-0.22	26.87	-0.31
4	0.00	42	0.13	[illegible]	[illegible]	27.57	-0.16
other	0.00	3	0.26	[illegible]	[illegible]	23.67	-0.99

Proportion without an item answer. 3 case(s) missing a response. 106 case(s) did not see this item.

Q80

option	wt.	n	p	[illegible]	[illegible]	avg.	z
1	0.00	124	0.39	[illegible]	[illegible]	26.71	-0.34
2	1.00	196	0.61	0.18	0.23	29.47	0.22
other			0.24				

Q80							
option	wt.	n	p	pb(r)	b(r)	avg.	z
1	0.00	124	0.39	-0.27	-0.35	26.71	-0.34
2	1.00	196	0.61	[illegible]	[illegible]	29.47	0.22
other			0.24				
Q100							
option	wt.	n	p	[illegible]	[illegible]	avg.	z
1	1.00	115	0.35	[illegible]	[illegible]	29.62	0.31
2	0.00	63	0.19	-0.10	-0.15	27.22	-0.21

Proportion without an item answer. 101 case(s) did not see this item.

See how it works? When Lertap is running with both options, did-not-see and MDO, then the other line will have stats only when there were some students who did not answer the item. For Q50, the avg. criterion score for the three students who were missing data was 23.67; when this avg. score is inserted into the distribution of criterion scores for those 312 students who did answer the item, the corresponding z-score is -0.99.

Related tidbit:

As discussed in the manual, when more than one option to a cognitive item has a non-zero "wt." value, the pb(r) and b(r) statistics are corrected for part-whole inflation only for the option having the greatest wt.

3.3.3 MDO cognitive, Statsb

The previous topic discussed how the MDO and did-not-see options affect the information reported in Statsf reports.

Now you're set to see the corresponding Statsb reports:

Lertap5 brief item stats for "Knowledge of LERTAP2", created: 20/03/2006.

Res =	A	B	C	D	E	F	other	diff.	disc.	?
Q15	2%	50%	22%	27%				0.50	0.53	A
Q16	20%	60%	5%	10%			5%	0.60	0.46	
Q17	12%	25%	57%				7%	0.57	0.56	
Q18	5%	43%	47%				5%	0.47	0.63	A
Q19	27%	10%	48%	8%			7%	0.48	0.76	
Q20	35%	40%	8%	7%			10%	0.40	0.70	D

Lertap5 brief item stats for "Knowledge with MDO", created: 20/03/2006.										
Res =	A	B	C	D	E	F	n	diff.	disc.	?
Q15	2%	50%	22%	27%			60	0.50	0.53	A
Q16	21%	63%	5%	11%			57	0.63	0.42	
Q17	13%	27%	61%				56	0.61	0.52	
Q18	5%	46%	49%				57	0.49	0.64	
Q19	29%	11%	52%	9%			56	0.52	0.74	
Q20	39%	44%	9%	7%			54	0.44	0.67	

The first report has an "other" column which indicates the number of people with missing data on each item, expressed as a percentage figure. The statistics in the diff. and disc. columns are based on calculations which include the people with missing data.

In the second report, the "other" column has been replaced by the "n" column. The entries in this column indicate how many people answered each item, and the various percentage figures, plus the diff. and disc. values, are based on n, that is, they exclude missing data. Lertap uses what is termed a "pairwise" exclusion rule to calculate the diff. and disc. values: if a student is missing data for the item, or did not see it, s/he is excluded from the calcs. for that item.

Lertap5 brief item stats for "EE 101 T3", created: 20/03/2006.								
Res =	1	2	3	4	n	diff.	disc.	?
Q50	37%	30%	20%	13%	312	0.30	0.25	
Q80	39%	61%			320	0.61	0.18	
Q100	35%	19%	34%	11%	326	0.35	0.13	
Q110	24%	39%	33%	5%	323	0.33	- 0.07	24
Q120	7%	43%	3%	47%	321	0.43	- 0.12	4
Q130	3%	6%	3%	88%	320	0.88	0.12	

The report above corresponds to the Test Pilot results mentioned in the previous topic. The numbers seen in the "n" column exclude the number of cases with missing data (if any), as well as the number of cases who did not see the item (if any). To see how many cases were missing data, or did not see the item, refer to the Statsf report which corresponds, or to the Freqs report.

Related tidbit:

A reminder from Chapter 10 of the manual: the Statsb reports are computed on an item level. The disc value they display is a conventional product-moment correlation between the item and the criterion, corrected for part-whole inflation.

3.3.4 MDO cognitive, Statsul

The two preceding topics have discussed how the MDO and did-not-see options affect Statsf and Statsb reports. Now: Lertap's third report for cognitive items, Statsul -- what happens when the MDO and did-not-see options are used?

Lertap5 U-L stats for "EE 101 T3 Fall 2005", created: 27/03/2006.

Res =	1	2	3	4	other	U-L diff.	U-L disc.
Q50 upper	0.23	0.38	0.11	0.01	0.27	**0.23**	**0.30**
2nd	0.19	0.29	0.12	0.11	0.30		
3rd	0.29	0.31	0.12	0.13	0.15		
4th	0.33	0.06	0.17	0.18	0.26		
lower	0.31	0.08	0.23	0.07	0.31		

First up, above, the standard format for Statsul, what's seen before the MDO and did-not-see options are put to use.

Item Q50 is from the Test Pilot system mentioned in the previous topics. Over 400 students took the "EE 101 T3" exam over the internet; 25.2% of the students did not see Q50 as Test Pilot did not present it to them, while just under 1% of the students who saw Q50 didn't answer it.

Lertap5 U-L stats for "EE 101 T3 Fall 2005 with MDO", created: 27/03/2006.

Res =	1	2	3	4	other	U-L diff.	U-L disc.
Q50 upper	0.31	0.52	0.15	0.02	0.00	**0.33**	**0.40**
2nd	0.27	0.41	0.17	0.15	0.00		
3rd	0.35	0.36	0.14	0.15	0.00		
4th	0.45	0.08	0.23	0.24	0.00		
lower	0.45	0.12	0.33	0.10	0.00		

The table above gives Q50 results after the MDO option has been turned on. Its statistics are based only on those students who answered the item. This means that all the did-not-see people, plus the we-saw-it-but-did-not-answer-it people, have been excluded from the proportions and from the calculations underlying U-L diff. and disc. The did-not-sees are excluded at this point as the did-not-see option has not yet been used; the did-not-see code has been processed as missing data.

Lertap5 U-L stats for "EE 101 T3 Fall 2005" (DNSI on), created: 27/03/2006.

Res =	1	2	3	4	other	U-L diff.	U-L disc.
Q50 upper	0.31	0.52	0.15	0.02	0.01	**0.32**	**0.40**
2nd	0.27	0.41	0.17	0.15	0.00		
3rd	0.35	0.36	0.14	0.15	0.00		
4th	0.44	0.08	0.22	0.24	0.01		
lower	0.44	0.12	0.32	0.10	0.01		

In this table (above), the did-not-see option has been activated, but MDO has not. The "other" column is now indicating the proportion of students in each group, from upper down to lower, who were presented with Q50 by Test Pilot, but did not answer it. Those not presented with Q50, the did-not-sees, have been excluded.

For a refresher on how the U-L diff. and U-L disc. values are calculated, pay a visit to your local ice cream shoppe, then have a look at Chapter 10 of the manual.

3.3.5 MDO affective

The "MDO" control word is used on the *sub card to get Lertap to exclude cases with missing data from its calculations. MDO may be used with both types of test, cognitive and affective.

The discussion found in this topic assumes some familiarity with material found in the topics immediately preceding. If you haven't been through them, take a few minutes to read the "Missing data" topic, followed by the topic dealing with the "Did-not-see option". Then report back here.

Ready, set, *go?* Have a look at the following CCs lines:

```
*col (c28-c37)
*sub Aff, Name=(Comfort with using LERTAP2), Title=(Comfort), Wt=0
*pol +---- ++--+
*col (c28-c37)
*sub Aff, MDO, Name=(Comfort items with MDO), Title=(MDOCmfrt), Wt=0
*pol +---- ++--+
```

The lines above set out two subtests. Both are affective as the "Aff" control word has been used on each of the *sub cards.

Both subtests involve the same ten items; no doubt you recognize the subtest? Right -- it's the set of Likert-style "Comfort" questions found on the Lertap quiz. And, no doubt you also recall that the items themselves may be see in Appendix A of that best-seller, the Lertap manual? Very good.

The only differences between the two subtests are found in the *sub lines. The second subtest uses the "MDO" control word.

Okay; rig yourself up with a refreshment of some sort, polish your glasses, and have a gander at Lertap's reports for these two subtests.

The Statsf reports

Lertap5 full item stats for "Comfort with using LERTAP2", created: 2

Q27

option	wt.	n	%	pb(r)	avg.	z
1	5.00	3	5.0	0.36	41.7	1.56
2	4.00	14	23.3	0.40	37.9	0.73
3	3.00	22	36.7	0.05	34.8	0.06
4	2.00	21	35.0	-0.57	30.9	-0.78
5	1.00	0	0.0	0.00	0.0	0.00

Q28

option	wt.	n	%	pb(r)	avg.	z
1	5.00	13	21.7	-0.24	32.4	-0.46
2	4.00	27	45.0	0.46	36.8	0.51
3	3.00	10	16.7	-0.16	32.8	-0.37
4	2.00	8	13.3	-0.08	33.5	-0.21
5	1.00	0	0.0	0.00	0.0	0.00
other	3.00	2	3.3	-0.22	29.0	-1.19

Lertap5 full item stats for "Comfort items with MDO", created: 29/03/

Q27

option	wt.	n	%	pb(r)	avg.	z
1	5.00	3	5.0	0.36	41.7	1.57
2	4.00	14	23.3	0.33	36.8	0.60
3	3.00	22	36.7	0.05	34.1	0.07
4	2.00	21	35.0	-0.51	30.2	-0.70
5	1.00	0	0.0	0.00	0.0	0.00

Q28

option	wt.	n	%	pb(r)	avg.	z
1	5.00	13	22.4	-0.20	32.2	-0.37
2	4.00	27	46.6	0.43	36.3	0.46
3	3.00	10	17.2	-0.22	31.6	-0.49
4	2.00	8	13.8	-0.13	32.4	-0.33
5	1.00	0	0.0	0.00	0.0	0.00
other	0.00	2	3.3	-0.31	26.0	-1.62

The first table above shows item stats for Q27 and Q28 without MDO, while the second table reflects the results of using MDO.

Q27's stats are the same in both tables, are they not? No-one omitted this item, so the statistics are unchanged, aren't they?

No. In fact, they're not unchanged (fooled you, eh?). Everything's the same until we get to the pb(r), avg., and z columns, wherein some changes enter.

To understand why Q27's results differ, look at the "other" row for Q28.

In the first subtest, without MDO, Lertap has wt.=3.00, giving 3.00 points to the two (2) people who did not answer Q28. Not so in the second subtest, where those two people have been stripped of scoring points. There are two different scoring methods in operation here: without MDO, people missing an answer to an item are given points equal to the average value of the wt. figures for the item's options. When MDO is active, as in the second subtest, no points are given when someone omits an item.

The result? The subtest scores will differ. Scores on the first subtest will be higher as people who miss out items are still getting points. The mean (average) of the subtest scores on the first subtest will be higher than that for the second subtest; the point-biserial correlation values, pb(r), between an option and the criterion score, the subtest score, are likely to differ, as are the avg. and z values. The more missing data, the greater these differences are likely to be.

Even though everyone answered item Q27, the criterion measure used to calculate item option statistics, pb(r), avg., and z, differs from the first subtest to the second, generally resulting in different values for item Q27's output.

Now, about Q28. As noted, two people did not answer this question. Compare the values found in the % column for Q28: they're greater in the second subtest. The % figures for Q28 in the second subtest, the one using MDO, have been calculated with n=58, the number of people who actually answered the item. In the first subtest, the % values were calculated with n=60, the total number of people taking the test (survey).

In addition, the pb(r) values for the item options seen in the second subtest have been calculated on a pairwise basis -- they are based only on the people who actually answered the item.

To read a bit more about how Lertap computes the Statsf figures, go back for a look at the "MDO cognitive, Statsf" topic.

There are two main differences between the Statsf reports for cognitive and affective subtests: it is rare for people omitting a cognitive item to get scoring points, so the statistics for cognitive item options may not be noticeably different going from no MDO to MDO. Lertap will apply a correction for inflation to the pb(r) and b(r) values corresponding to the right answer to a cognitive question, but this correction is not applied in the Statsf report for affective items (but it is for Statsb: see below).

The Statsb reports

Lertap5 brief item stats for "Comfort with using LERTAP2", created: 29/03/2

Res =	1	2	3	4	5	other	pol.	mean	s.d.	cor.
Q27	5%	23%	37%	35%			-	2.98	0.88	0.55
Q28	22%	45%	17%	13%		3%	-	3.75	0.94	- 0.14

Lertap5 brief item stats for "Comfort items with MDO", created: 29/03/20

Res =	1	2	3	4	5	n	pol.	mean	s.d.	cor.
Q27	5%	23%	37%	35%		60	-	2.98	0.88	0.48
Q28	22%	47%	17%	14%		58	-	3.78	0.95	- 0.10

In the normal case, without the MDO option, the Statsb report for cognitive items has an "other" column which indicates the percentage of non responses to an item. This column changes to "n" when MDO is in operation, as may be seen above.

Q27 has no missing data; its mean and s.d. values are the same in both tables. The Q27 cor. figures differ for the reason found earlier in the Statsf reports: the values of the criterion measure, the subtest score, change as we go from no MDO to MDO.

Q28's figures differ almost everywhere. The percentages, mean, s.d., and cor. statistics for the second table, where MDO is having its impact, are all computed using only the responses from the 58 folks who actually answered this item. In some other data analysis systems, such as SPSS, the correlation (cor.) between Q28 and the criterion would be said to done on a pairwise basis: only when a person has data for both variables are that person's results used in the calulations.

Another note about the cor. values found in the Statsb reports: they are corrected for part-whole inflation.

Using the did-not-see option

Suppose the did-not-see option has been turned on, with X used as the did-not-see code. The respective lines in the System worksheet would look like this:

These are Lertap5 system settings. Change them only if you understand them.	System Settings		
	Present setting:	**Allowed settings:**	**Usual setting:**
Use a **did-not-see** code?	yes	yes / no	no
Did-not-see code (single character; may be blank):	X	any char	

Comments / Data / CCs / **System** / Syntax / OldCCs / Problems

Next, have a look at a snippet of Freqs output:

(c37) Q35		
Option	n	/60
X	6	10.0%
1	2	3.3%
2	13	21.7%
3	12	20.0%
4	17	28.3%
5	7	11.7%
?	3	5.0%

Six people did not see Q35; three did not answer it.

Okay? Now, suppose MDO is not operating. The Stats1f output for Q35 will look like this:

Lertap5 full item stats for "Comfort with using LERTAP2", created: 30/03/2006.

Q35

option	wt.	n	%	pb(r)	avg.	z
1	1.00	2	3.7	-0.05	30.5	-0.24
2	2.00	13	24.1	-0.16	30.1	-0.28
3	3.00	12	22.2	0.02	33.1	0.04
4	4.00	17	31.5	0.35	37.4	0.51
5	5.00	7	13.0	0.30	40.0	0.79
other	3.00	3	10.0	-0.02	32.0	-0.07

Data / CCs / Freqs / Scores / **Stats1f** / Stats1b / Stats2f / Stats2b /

If the MDO option is turned on, the output will change:

Lertap5 full item stats for "Comfort items with MDO", created: 30/03/2006.

Q35

option	wt.	n	%	pb(r)	avg.	z
1	1.00	2	3.9	-0.15	30.5	-0.75
2	2.00	13	25.5	-0.51	29.8	-0.88
3	3.00	12	23.5	-0.19	32.6	-0.35
4	4.00	17	33.3	0.39	37.2	0.56
5	5.00	7	13.7	0.44	40.0	1.09
other	0.00	3	15.0	-0.26	29.0	-1.04

Data / CCs / Freqs / Scores / Stats1f / Stats1b / **Stats2f** / Stats2b /

To grasp what Lertap has done, look down the % column for these two reports.

In the top report, the % values (and the columns to the right, from pb(r) to z) are based on n=54; the six people who did not see Q35 have been excluded from the calculations.

In the next report we've got MDO operating, and now we'll have n= what? Fifty-one (51). In this case, the report excludes the six did-not-sees, and the three did-not-answers.

Right. What about the corresponding Statsb reports? Thought you wouldn't ask. Here they be:

Lertap5 brief item stats for "Comfort with using LERTAP2", created: 30/0

Res =	1	2	3	4	5	n	pol.	mean	s.d.	cor.
Q35	4%	24%	22%	31%	13%	54	+	3.26	1.07	0.94

Data / CCs / Freqs / Scores / Stats1f / **Stats1b** / Stats2f / Stats2b /

Lertap5 brief item stats for "Comfort items with MDO", created: 30/03/20

Res =	1	2	3	4	5	n	pol.	mean	s.d.	cor.
Q35	4%	25%	24%	33%	14%	51	+	3.27	1.10	0.56

Data / CCs / Freqs / Scores / Stats1f / Stats1b / Stats2f / Stats2b

Where do you stand now? You see what happens, or are you in the did-not-see group?

Questions? Crank up your email program, and point it at: larry@lertap.com.

3.4 Pre-scored items

It is usually the case that the entries found in the columns of the Data worksheet correspond to the response codes selected by each person.

For example, have a squiz at this snippet from rows 2 and 3 of a Data worksheet:

Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30
A		E	E	B	F	2	3	1	2	2

On Q20, this person selected the response which had been coded as A. Apparently s/he did not answer Q21. On Q22, the person selected the response coded as E.

The column entries change to digits from Q26 on, but the meaning is probably the same: on Q26 the person selected the response coded as 1 (one), while on the next item, Q27, s/he chose the response which had been coded as 3.

What we're looking at here is part of the Lertap Quiz data set, described in Appendix A of the manual. This quiz consisted of 25 cognitive items, Q1-Q25, followed by 10 affective items, Q26-Q35. The cognitive items used letters as response codes, while the affective items were of the Likert style, with 1 the code for "strongly agree", and 5 the code for "strongly disagree".

How many points did this person get for her/his answer of A on Q20? We don't know; we can't tell just by looking at the data above. And, even though there are digits in some of the columns, we can't assume that a "1" for Q28 means that the person got one point for his/her answer.

These item responses have yet to be scored.

Okay? Consider now another case. Suppose a cognitive test included the following question:

33) Read the five sentences below, and place a tick next to those sentences which use the pluperfect tense.

Student answers to a question like this one have to be scored by hand. How? Well, if there were three pluperfect sentences, and the student found and ticked each, then the student would probably get 3 points. If the student found two of the three, s/he'd get 2 points. A student might get 0, 1, 2, or 3 points on this item.

Look now at a snippet from another Data worksheet, would you?:

L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
1	1	0	1	1	1	2	1	1	1	1	1	2	0	1	1	1	1	1	1

In this case, language students listened to an audio tape with 10 short sentences, and they also read ten brief paragraphs. The digits in the boxes are now real numbers -- they're not response codes -- the digits represent the number of points the student earned on each of the ten listening items, L1-L10, and the number of points earned on each of ten reading items, R1-R10. The items in this test are said to have been "pre-scored", or "pre-coded".

The following CCs lines were used to process results for the ten Listening items:

```
*col (c2-c11)
*sub res=(0,1,2,3), name=(Listening AARP test), Title=(L-AARP)
*mws  c2, 0, 1
*mws  c3, 0, 1
*mws  c4, 0, 1, 2
*mws  c5, 0, 1, 2
*mws  c6, 0, 1, 2
*mws  c7, 0, 1, 2, 3
*mws  c8, 0, 1, 2, 3
*mws  c9, 0, 1, 2
*mws c10, 0, 1
*mws c11, 0, 1, 2, 3
```

The *mws lines indicate that the number of item points possible varied from item to item. For the first two items, scores were limited to 0 or 1; for the items found in columns 4, 5, 6, and 9 possible scores ranged from 0 to 2. Three items, those in columns 7, 8, and 11, had possible scores ranging from 0 to 3.

Fewer lines would have been required had it been possible to get the same number of points on each item:

```
*col (c2-c11)
*sub res=(0,1,2,3), name=(Listening AARP quiz), Title=(L-AARP)
*mws call, 0, 1, 2, 3
```

Here, the "call" form of the *mws line has been used -- "call" means "columns all". According to these CCs lines, it's possible for a student to get up to 3 points on each item.

3.5 Open-ended items

It is possible to have Lertap score open-ended items, such as short-answer questions, constructed-response items, supply items, and essay questions.

Please refer to the following URL for more information and examples:

http://www.lertap5.com/Documentation/ScoringNonMCItems1.pdf

3.6 Remove an item

It sometimes happens that users will be Lertapping along, humming their favourite tune, and suddenly wish that they could omit a single item or two from a subtest, just to see how that might change things (such as the value of coefficient alpha). There are a variety of ways in which this may be accomplished.

The most obvious way is to make a new *col card.

For example, consider these CCs cards:

```
*col (c28-c37)
*sub aff, title=(Comfort)
*pol +---- ++--+
```

Now, suppose it was found that the item in column 36 was not correlating well with the others, was serving to bring down the subtest's reliability figure (coefficient alpha), or for some other reason had to be removed from the subtest. The following *col card will do the job:

```
*col (c28-c35, c37)
```

But this isn't the only change required. The *pol card must be changed too -- we've taken an item out of the subtest, and must remove the corresponding plus (+) or minus (-) sign from the *pol card:

```
*col (c28-c37)
*sub aff, title=(Comfort)
*pol +---- ++-+
```

An easier way to take the item out is to make use of a special form of the *mws card:

```
*col (c28-c37)
*sub aff, title=(Comfort)
*pol +---- ++--+
*mws c36, *
```

The *mws card above has a single asterisk after the column number. This is a special form of the *mws card, used to remove an item. This special use of the *mws card eases the task of taking items out of a subtest -- there's no requirement to make corresponding changes to other cards, such as the *pol card.

There's another way to remove items from affective subtests: use asterisks on the *alt card, as shown in the example below:

```
*col (c28-c37)
*sub aff, title=(Comfort)
*pol +---- ++--+
*alt 55555 555*5
```

The *alt card above tells Lertap that the penultimate item (second to last) is to be excluded from the subtest.

The examples above are based on an affective subtest, but the special uses of the *mws and *alt cards shown here also apply to cognitive subtests. Consider this example:

```
With all items:
*col (c3-c27)
*sub res=(A,B,C,D,E,F), Title=(Knwldge)
*key AECAB BEBBD ADBAB BCCCB BABDC
*alt 35423 35464 54324 43344 45546
Using *mws to remove the sixth item:
*col (c3-c27)
*sub res=(A,B,C,D,E,F), Title=(Knwldge)
*key AECAB BEBBD ADBAB BCCCB BABDC
*alt 35423 35464 54324 43344 45546
*mws c8, *
Using *alt to remove the sixth item:
*col (c3-c27)
*sub res=(A,B,C,D,E,F), Title=(Knwldge)
*key AECAB BEBBD ADBAB BCCCB BABDC
*alt 35423 *5464 54324 43344 45546
```

SAQ: would the example immediately above actually work? If I copied the 16 lines and pasted them into a CCs worksheet, would they actually work? Yes. This example is just a straightforward job with three subtests. The lines which do not begin with an asterisk are comments, and are not processed by Lertap.

Finally, our examples here have discussed removing a single item from a subtest. To remove more than one item, follow the same pattern. The examples below will remove two items from their respective subtests:

```
*col (c28-c37)
*sub aff, title=(Comfort)
*pol +---- ++--+
*mws c29, *
*mws c36, *

*col (c3-c27)
*sub res=(A,B,C,D,E,F), Title=(Knwldge)
```

```
*key AECAB BEBBD ADBAB BCCCB BABDC
*alt 35423 *5464 54324 4*344 45546
```

Note: the *exc card makes it somewhat easier to remove, or "exc"lude, items. Please to see the following topic.

3.7 Excluding items

The matter of removing one or more items from a subtest is discussed in the previous topic. As mentioned there, the quickest way to exclude a single item from a subtest is probably to use a single *mws card. For example, the following CCs line will see that the item resident in column 37 of the Data sheet is eliminated from its subtest:

    *mws c37, *

A new CCs card, or line, has been added to ease the task of excluding multiple items. Its format is identical to that of the *col card.

The following CCs line will remove, or exclude, the item in column 37:

    *exc (c37)

Other examples of the use of this card:

    *exc (c12-c14, c42)

Removes the items found in columns 12, 13, 14, and 42.

    *exc (c12, c13, c14, c42)

Also removes the items found in columns 12, 13, 14, and 42.

    *exc (c17, c21-c25, c27, c35-c40)

Will exclude the items in columns 17, 21, 22, 23, 24, 25, 27, 35, 36, 37, 38, 39, and 40.

Both the *mws and *exc lines may be used to remove items from a subtest, as shown in the example below:

    *mws c17, *
    *mws c27, *
    *exc (c21-c25, c35-c40)

The three lines above will see that the items in columns 17, 21, 22, 23, 24, 25, 27, 35, 36, 37, 38, 39, and 40 are excluded from the subtest.

To see how the *exc card can ease the process of removing items from a subtest, look at the following example:

Scale 1 uses 28 of the items.
*col (c1-c20 c31-c50)
*sub Res=(A,B,C,D,E), Name=(Critical Skills), Title=(Critical), Wt=0
*key CCACB CABAB BDBAD BBAAD ECDCE ABDDA CBCBB BDDCA
*mws c1, *
*mws c3, *
*mws c6, *
*mws c10, *
*mws c15, *
*mws c17, *
*mws c31, *
*mws c34, *
*mws c38, *
*mws c40, *
*mws c46, *
*mws c49, *
Scale 1 uses 28 of the items.
*col (c1-c20 c31-c50)
*sub Res=(A,B,C,D,E), Name=(Crit. Skills A), Title=(CrtSklA), Wt=0
*key CCACB CABAB BDBAD BBAAD ECDCE ABDDA CBCBB BDDCA
*exc (c1,c3,c6,c10,c15,c17,c31,c34,c38,c40,c46,c49)

The example above shows two ways to have 12 items excluded from a 40-item subtest.

The first way involves the use of multiple *mws cards; the second way uses a single *exc card. (There are other ways to exclude items: refer to the previous topic for more details.)

3.8 Pilot items

Sometimes a test, or subtest, will include items which are not to be scored.

These are often referred to as "pilot" items, "pretest" items, or "trial" items. They're riding in the test, usually embedded amongst the normal items, just to see how good they are. If they qualify as okay, they might appear in a subsequent version of the test, and be scored (that is, they will move from being unscored items to normal items).

To see how to handle these in Lertap, please refer to this topic.

3.9 Split-half reliability

Lertap uses Cronbach's coefficient alpha as its reliability estimate.

At times split-half reliability estimates may be useful, such as that found by correlating two half-tests.

For example, the half-tests may be created by splitting out the odd-numbered items as one half, and the even-numbered items as another half.

To see how to split a test into halves with Lertap, and correlate the halves, please refer to this topic.

A related topic having to do with assessing reliability has to do with "parallel forms", a situation where two tests designed to measure the same thing are developed simultaneously. At times these tests will have some common test items, at times they will not. They're called "parallel tests" as they can be used, ideally, interchangeably.

An interesting example of a parallel-forms development project may be seen here. Of added interest in this project was the use of a mixture of test items, multiple-choice and constructed-response.

3.10 Omega reliability

Lertap 5's main measure of test and scale reliability is coefficient alpha, also known as Cronbach's alpha. There are many other measures of reliability; the preceding topic discusses some of them. When a test with a cut-score is used, measures of decision consistency will often become primary indicators of test quality -- these measures are discussed in another topic.

The text by Meyer (2010) is a highly-recommended reference in all of these areas.

Coefficient omega, also known as McDonald's omega, is yet another measure of reliability, not discussed in Meyer's text. It has been shown that omega is a superior measure of reliability as it makes fewer assumptions; this is now widely acknowledged, but practical studies, such as this one, have found that alpha and omega will be generally not be noticeably different.

The "Item scores and correlations" option in Lertap will produce the "IStats" report. An estimate of omega computed using the "closed-form" method will be found towards the end of the report. An example is found in Appendix D of this working paper.

Those interested in omega might also want to make use of a special Lertap macro, "Omega1". This macro will create a special 'csv' data file and a small text file with several lines of R code designed for use with the Psych toolbox from CRAN.

3.11 Filtering records

By "filtering records" is meant the process of getting Lertap to produce its various results using *only some* of the records in the Data worksheet.

There are three ways it can be done.

The Move+ Menu has the "Recode a Data column" option which provides a flexible way to filter records. The word "Delete" can be attached to certain data records, as can the word "Exclude". In many cases this will be a very effective way to quickly remove or exclude records. Click here.

The NumericFilter2 macro will breakout, or select, only those Data records meeting a specified criterion. It's accessed via the Macs menu.

The special *tst control "card" may also be used to have Lertap filter records. Three examples of its use are:

A) *tst c6=(CS001)

B) *tst c6=(UCS001, CS001)

C) *tst c6=(DL5, DL6, DXL7), c7=(WI)

In example A), all data records with CS001 in column 6 of the Data worksheet will be filtered out *into a new data set (that is, a new Excel workbook).*

In Example B), all records with either UCS001 or CS001 will be filtered out *into a new data set (a new Excel workbook)*.

In Example C), all records with either DL5, DL6, or DXL7 in column 6, and WI in column 7, will be filtered out *into a new data set (a new Excel workbook)*.

When a *tst card is used, it has to be placed at the top of the CCs worksheet, it must be the first "card" (or line) in the CCs sheet. See this topic for a bit more information on *tst.

3.12 How CCs cards work

Each subtest requires a minimum of two CCs cards.

Cognitive subtests

Cognitive subtests must have a *col card, and they must also have a *key card with the right answers.

The number of columns mentioned on the *col card tell Lertap the number of items in the subtest. The default response code set for cognitive items is Res=(A,B,C,D).

The CCs card order for cognitive subtests is:

```
  *col  (a required card)
  *sub
  *key  (a required card)
  *alt
  *wts
  *mws
```

The *key, *alt, and *wts cards require a character for each item in the subtest. If there are five items, there must be five entries on the *key card; the *alt and *wts cards, if used, must also contain five entries. There may be spaces between the entries. We like to put a space after every five entries as that way the card is easier to read; we also like to use a fixed-pitch font with CCs cards, such as Courier New, so that when we use *key, *alt, and *wts cards the entries on all cards line up.

*mws cards are unique in that they correspond to just a single item.

Affective subtests

Affective subtests must have a *col card, and they must also have a *sub card with the AFF control word on it.

The number of columns mentioned on the *col card tell Lertap the number of items in the subtest. The default response code set for affective items is Res=(1,2,3,4,5).

The CCs card order for affective subtests is:

```
  *col  (a required card)
  *sub  (a required card; must have AFF on it)
  *pol
  *alt
  *mws
```

If used, the *pol and *alt cards require a character for each item in the subtest. If there are five items, for example, there must be five entries on these cards (if they're used).

*mws cards are unique in that they correspond to just a single item.

Control words

The *sub card may have a number of control words on it; these are summarised in the tables below:

*sub control words for cognitive subtests:

CFC	Optional. Means "correction for chance". Adjusts subtest scores for the (estimated) effects of guessing.
Master y	Optional. Gets Lertap to produce its mastery test analysis and report. Also acts as if the PER control word had been used, causing the PER score to appear on the Scores report. (Note: this control word is the same as using Mastery=70; a default mastery level of 70 is programmed into the System worksheet.)

Mastery=	Optional. Lertap will produce a mastery test analysis and report, using the cutoff figure found after the equals sign. Also causes a PER score to be created. Example: Mastery=65
MDO	Optional. Gets Lertap to turn off its missing data item scoring. Causes a non-response to be scored as zero points. (A click here will whisk you away to more about MDO.)
Name=()	Optional. Whatever is found between the parentheses is used as a header on some of Lertap's reports, such as Stats1f and Stats1b. While the header can have any length, something less than 40 characters is best. Example: Name=(SOC 505 FINAL, November 2003)
PER	Optional. Causes a percentage of maximum possible score to be created; this will appear as a column in the Scores report.
Res=()	Required if the default Res=(A,B,C,D) setting is not appropriate.
SCALE	Optional. Causes a z-score to be computed and added as a column in the Scores report.
Title=()	Optional. Provides a short label for the subtest score. This will appear as a header at the top of a Scores column. Should be no longer than 8 characters. Example: Name=(SocFinal)
Wt=	Optional. Applies only when there are multiple subtests. Determines how the subtest's score comes into the total test score. Example: Wt=0.5

*sub control words for affective subtests:

AFF	Required. The appearance of the AFF control word on a *sub card is the only way Lertap knows a subtest is of the affective type.
MDO	Optional. Gets Lertap to turn off its missing data item scoring. Causes a non-response to be scored as zero points. (A click here will whisk you away to more about MDO.)
Name=()	Optional. Whatever is found between the parentheses is used as a header on some of Lertap's reports, such as Stats1f and Stats1b. While the header can have any length, something less than 40 characters is best. Example: Name=(WA/Maui beaches survey, January 2004)
PER	Optional. Causes a percentage of maximum possible score to be created; this will appear as a column in the Scores report.
Res=()	Required if the default Res=(1,2,3,4,5) setting is not appropriate.

SCALE	Optional. Causes a new score to be computed and added as a column in the Scores report. The new score is the original score divided by the number of items in the subtest.
Title=()	Optional. Provides a short label for the subtest score. This will appear as a header at the top of a Scores column. Should be no longer than 8 characters. Example: Name=(BeachSur)
Wt=	Optional. Applies only when there are multiple subtests. Determines how the subtest's score comes into the total test score. Example: Wt=0.75

The order of the control words is not important. For example, the following two cards accomplish the same thing:

```
*sub Title=(NewsQuiz), PER, SCALE
*sub SCALE, Title=(NewsQuiz), PER
```

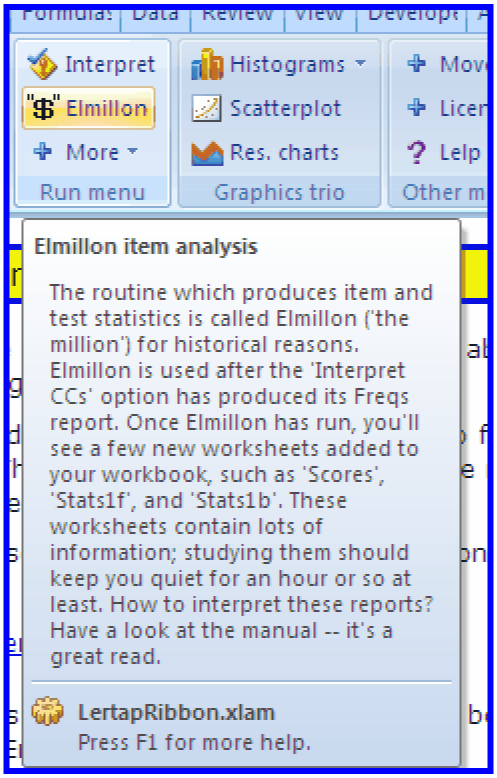
The control words may be abbreviated, or expanded. The following cards result in the same actions:

```
*sub AFF, Name=(Beach survey 1), Title=(Beachin)
*sub Affective, T=(Beachin), N=(Beach survey 1)
```

4 Toolbar and tab

The tab makes it possible to view on-screen hover-help for those running under Windows -- this feature will hopefully also be available in a future revision of Excel for the Mac.

Hover-help? Sure. Look:



Just let your little mouse cursor hover above the options on the Lertap tab, and a bit of help comes into view. *Hover help.* (The "view" is better on Windows versions of Excel -- the Mac version's hover help is a bit less helpful at the moment).

Page ahead to see what the Lertap tab's options do.

4.1 The tab in detail

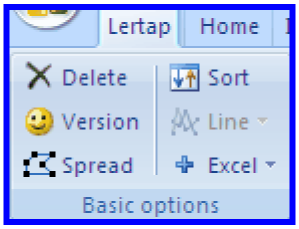
The Lertap tab for Excel has five groupings: "Basic options", "New menu", "Run menu", "Graphics trio", and "Other menus".

The three sections in the middle of the tab are used most often. The New menu is used to set up a new Lertap workbook; with a Lertap workbook in hand, the Run menu is then used to produce results: statistical summaries of item and test results, and student scores; with results in hand, the Graphics trio's three options may then be used to create pictures (charts, or graphs) of selected results.

Continue to page forward for a section-by-section discussion of what the tab's options do, or click on one of the sections as displayed above to jump directly to a particular section.

4.2 Basic options

The 'Basic options' section of the Lertap tab has six components, or options.



There isn't really a theme to these options -- they do quite different things. Click on them and you'll see. Or, page forward to browse each of these options, one by one.

4.2.1 Delete

 This icon deletes certain worksheets from a Lertap 5 workbook.

To understand what it does, and why it is used, consider a standard Lertap 5 job: a user puts his/her item responses, and perhaps other data, into a workbook's Data worksheet. S/he puts control statements, or control "cards", into the CCs worksheet.

Users then click on the Run drop-down menu, and selects the "Interpret CCs lines" option. What happens? Lertap looks at the rows in the CCs worksheet to find out which columns in the Data worksheet the user wants to analyse, and then it goes on to read data, and add new worksheets to the workbook.

The most obvious worksheet added by this step is the one called "Freqs".

Users pause to scan the information found in the Freqs worksheet. If all is in order, the user usually returns to the Run menu, and clicks on "Elmillon item analysis".

What happens? Lertap adds more new worksheets to the workbook, such as "Scores", "Stats1f", "Stats1b", and so on.

These worksheets which Lertap adds are called secondary worksheets. The original Data and CCs worksheets are referred to as primary worksheets. At any time, the secondary worksheets may be regenerated if the primary worksheets are still on hand.

What the little  does is allow the user to quickly delete the secondary worksheets. When this icon is clicked on, a dialog box appears which informs users that "all worksheets whose names begin with *freqs*, *scores*, *sub*, *stats*, and *histo* will be deleted" if the user clicks on the OK button.

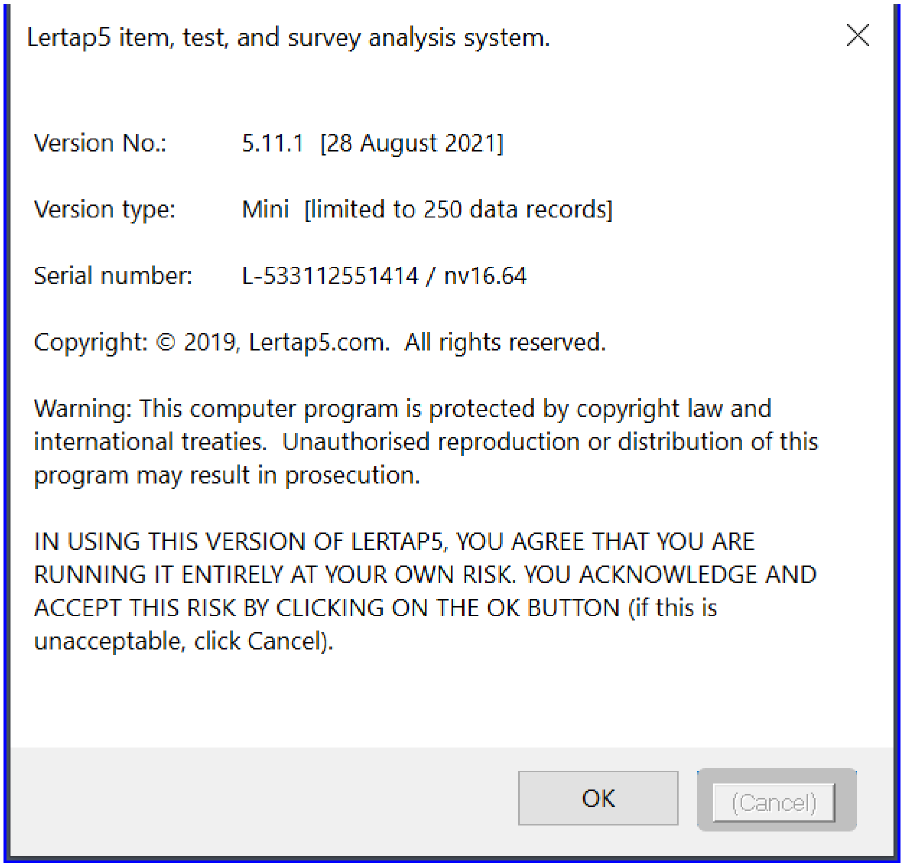
Why would a user want to delete the secondary worksheets? Probably to save disk space. The secondary worksheets can be quite large if there are many items in the data set, and/or if there are many records in the data set. Users may also want to delete the secondary worksheets if they want to send a copy of their workbook to someone else.

Ordinarily, no harm is done by using . Going through the process of "Interpret CCs lines" and "Elmillon item analysis" will restore the secondary worksheets at any time. Users who wish to make sure their secondary worksheets will not be deleted by this icon should rename the worksheets, perhaps by putting a prefix before their names. For example, *Freqs* might be renamed as *OrigFreqs*. Note that renaming *Freqs* as *FreqsOrig* would do no good as the worksheet's name still begins with *Freqs* -- a prefix is suggested, not a suffix.

How to rename a worksheet? The fastest way is to right-click on the worksheet's tab at the bottom of the screen. To read about the various ways which may be used to rename worksheets, please refer to Excel Help.

4.2.2 Yellow smiley face

Use this icon to get the Lertap5 version type, and date of generation (*it may not be in yellow*).



The screen snapshot showing above indicates that the "Mini" version of Lertap 5 was in use. At the time the snap was taken, Version 5.11.1 was in use, last revised on 28 August 2021.

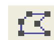
Note (1): The Cancel button leads to an exit from Lertap -- it will stop running.

Note (2): more about the "Yellow" Smiley and version numbering may be found here.

4.2.3 Data entry aid

this icon, when clicked on, activates "The Spreader", a powerful data entry aid. To understand how it works, look at the screen capture below:

56	60	A	E	C	B	B	B	E	B	B	D		B	B	A
57	56	A	E	C	B	C	B	E	B	F	D	A	D	B	A
58	15	B	B	D	B	B	A	C	C	A	C	D	B	C	A
59	40	A	E	C	A	C	A	E	B	A	D	A	D	B	B
60	46	A	E	C	A	A	B	A	B		D	C	D	B	A
61	22	BECDDADEDCBDEABCDEDAABBED2312454321													

A string of 35 characters has been entered into a single cell, directly on the Data worksheet. When you employ The Spreader, it will almost always be the case that row after row will have its item responses entered in this manner. After the last row of data has been entered, you scroll up to the first row, click on the cell containing the string, and then click on The Spreader's icon, .

The Spreader will dissect the string, spreading the characters, one by one, over the cells to the right. Having done this, it then looks at the next row in the worksheet. If there's another long string to be dissected, it does so. The Spreader continues to work down the rows, stopping when it finds one without a string.

The Spreader may be stopped at any time by pressing the Esc key on the computer's keyboard.

What about missing data? Say someone has not answered one of the items -- in this case, leave a blank in the string (pressing the space bar on the computer keyboard will produce a blank).

An important case arises when the string of responses begins with a digit. Excel will think the entry is going to be a number, and strange things can happen. When the string begins with a digit, it should be preceded by an apostrophe, as seen here:

'2344133124AADDB

The apostrophe tells Excel to consider the entry as "text", not a number. Another way to define an entry as text is to use Excel's Format menu / Cells, then select Text. The cells in an entire column may be formatted as text by selecting the column before using the Format menu.

The Spreader is a favourite tool at Lertap HQ. It's powerful indeed. It is often useful when importing data from other applications, including scanners.

The data entry aids are discussed in Chapter 3 of the manual, under the section titled "Entering item responses".

In the printed manual, the reference is page 61.

Update to the manual:

The Spreader gained more smarts after the manual was printed way back when. Its standard method of operation now involves two passes down the cells whose contents are to be spread.

In the first pass, the length of the cell's string is compared to the length of the string in the cell above. If the lengths are not the same, The Spreader sounds an alarm. You get the chance to stop The Spreader, or to continue. If you choose to stop, you can then edit the cell.

Once The Spreader has worked its way down all the relevant cells the first time, it will then ask if you're ready to truly spread cell contents. You can stop at this point without anything having changed. If you elect to continue, The Spreader goes back up to where you last started it from, and spreads things to the right

You can change the way The Spreader operates by changing the appropriate row in the System Worksheet. It's possible to tell The Spreader to forget about the first pass, the one where it checks string lengths. The Spreader is not slow, but, as you'd expect, it runs even faster if it doesn't have to make two passes.

At Lertap HQ we favour the two-pass method of operation as we feel it's reassuring to check string lengths. However, even the two-pass Spreader will not control for a nasty problem: when the string has been created by importing data from a scanner-made text file, blanks at the start of the string may be lost. Such blanks usually correspond to unanswered questions. It's rare for a respondent to leave the initial questions unanswered, but it does happen, and when it does real care is required to make sure that the blanks remain at the start of the string. If they do, The Spreader will spread them. If they don't, woe! -- the string will be shifted to the left, and Lertap will be unaware of what's happened. If a test is being scored, the score will be wrong.

There's a bit more about this nasty in the Import & Export topic.

4.2.4 Sort A to Z

This icon is used in conjunction with the Scores worksheet. It permits the information in the Scores sheet to be sorted according to criteria entered by you, the user.

When a sort is requested, Lertap makes a copy of the Scores worksheet, and adds it to the workbook as a new worksheet called Sorted. Then Excel's standard sort criteria box appears, and the stage is set -- *sorts are made using this new worksheet*.

After a sort has been made, may another sort be requested? Yes. There are a couple of ways to make an additional sort. First, the Sorted worksheet could be further sorted by going directly to Excel. It has a icon, found on Excel 2007's Home tab. You might want to use Excel Help if you're unsure about using Excel sort.

Another way to sort a second time is to delete the Sorted worksheet, and then use Lertap's icon again. Or, instead of deleting the Sorted worksheet, it could be renamed, after which Lertap's icon will be happy to once more do its job.

At all times care should be taken to see that the Scores worksheet itself is *never* sorted. Lertap needs to believe that there's a one-to-one correspondence between the records in the Data worksheet, and the records in the Scores worksheet. This will not be the case if the Scores worksheet is sorted, and it's precisely because of this restriction that Lertap makes the Sorted worksheet for users to work their sorts on.

4.2.5 Line grapher

This option is a little gem.

Here at Lertap HQ there's often a real urge to make line graphs from Lertap reports. For example, we will commonly get into a "Stats1b" report, and graph item difficulties, and also item discrimination coefficients. We'll use the Run menu to "Output an item scores matrix", from which we'll plot such things as item means and variances. And we often go for a scree test by plotting eigenvalues.

The little icon option makes it possible to get such plots with ease.

How to use this option? Couldn't be easier: select the cells you want to plot, and then just click on the icon. That's it? Yes; assuming you have selected cells which have numeric data in them, Excel 2007 will present its line graphing options, and in another click you'll have that graph.

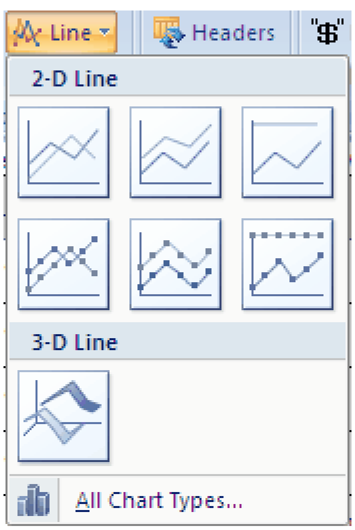
Here's an example -- we wanted a plot of item means as found in a "Stats2b" report.

We selected the cells of interest by running our mouse over them (there are 10 means values selected in the screen snapshot seen below, and note that we started selecting in the 'mean' cell):

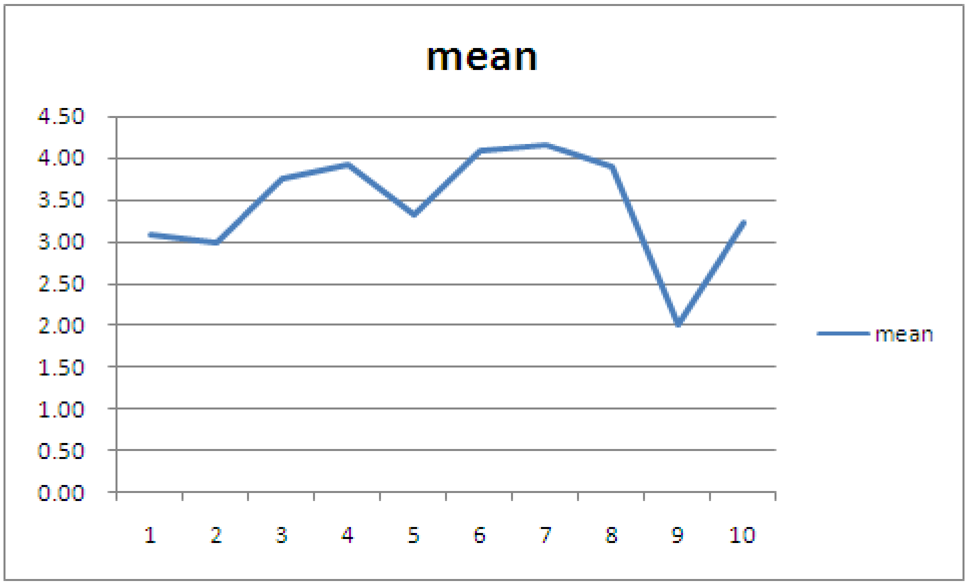
Lertap5 brief item stats for "Comfort with using LERTAP2", created: 17/07/2

Res =	1	2	3	4	5	other	pol.	mean	s.d.	cor.
Q26	13%	22%	25%	23%	17%		+	3.08	1.28	0.76
Q27	5%	23%	37%	35%			-	2.98	0.88	0.55
Q28	22%	45%	17%	13%		3%	-	3.75	0.94	- 0.14
Q29	32%	35%	25%	5%		3%	-	3.93	0.89	0.44
Q30	15%	33%	28%	13%	8%	2%	-	3.33	1.14	0.49
Q31		3%	18%	43%	35%		+	4.10	0.81	- 0.05
Q32			13%	53%	32%	2%	+	4.17	0.66	0.22
Q33	40%	23%	23%	13%			-	3.90	1.08	0.65
Q34	2%		17%	60%	22%		-	2.00	0.73	- 0.56
Q35	3%	22%	20%	28%	12%	15%	+	3.23	1.02	0.57

Then we clicked on .



Excel 2007 popped up its select-a-chart-type box, seen above. We clicked on the first chart type, the one in the upper left. This gave us what we wanted:



You can modify the resultant Excel chart, the line graph, by using Excel's standard chart options, of which there are many. You can add titles, legends, and change lots off colours ... why, you could probably spend two or three hours enhancing Lertap's initial plot, ending up with a graph which, when pasted into your final report, is bound to bring you great kudos.

One thing that's really neat about these line graphs is that you can see the x and y values associated with any of the line graph's points by just letting your mouse hover right above one of the points. Try it -- you don't need to hold down a mouse button -- just position the mouse pointer on top of a point, and the corresponding x and y values will jump out at you. (Note that this won't work with the graph above as it's just a picture of Excel output, not the real thing.)

What's that you're saying? You'd like to know more about Excel charts? *Goodonyou.* Use Excel's Help system -- it's got heaps of info. Heaps.

Line graph problems

Users just starting to use Lertap's line grapher shortcut may find that it won't work as advertised here. In the tests we've done to date, the problems which arise have to do with selecting cells.

Cells are selected in normal fashion: by highlighting them with the mouse, or by holding down the Shift key and using the arrows on the keyboard. The cells are expected to have numbers in them -- however, the first cell selected can have text information, such as a row or column header -- if Excel finds the first selected cell to contain text, it'll use the text as a title for the line graph (this can be real handy).

Lertap will fail to make a line graph if: no cells are selected; only one cell is selected; or if the selected cells do not contain numbers (except for what we've just mentioned: the first selected cell may contain text).

4.2.6 Excel

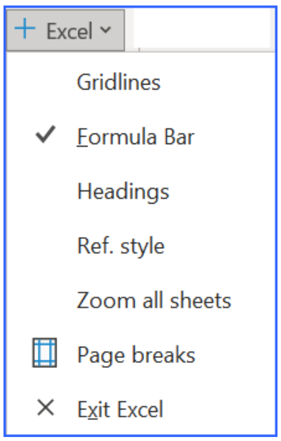
A variety of Excel options are available on this little drop-down menu. They're meant to be handy shortcuts.

The first three, Gridlines, Formula Bar, and Headings have their real home on Excel's View tab. The drop-down menu makes it easier to access them. (In the snapshot below, only the Formula Bar has been selected.)

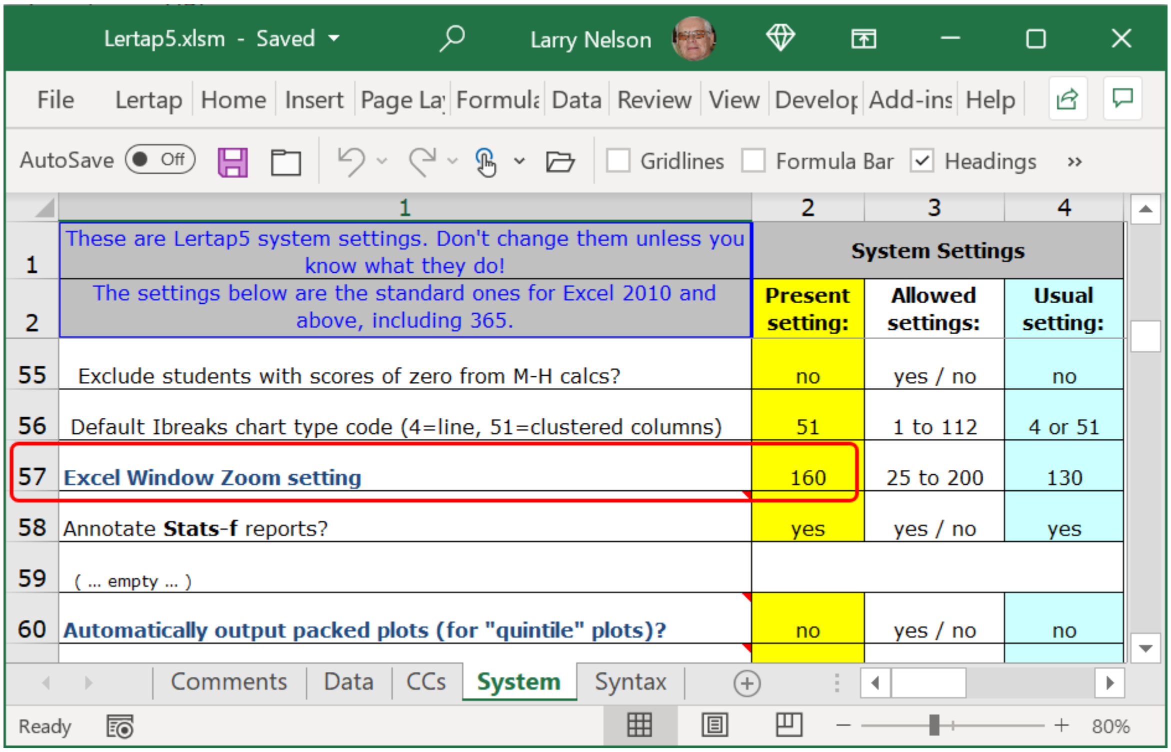
"Ref. style" relates to Excel having two ways of referring the cells found in a spreadsheet's rows and columns. They're referred to as "styles".

There are two styles: "R1C1" and "A1".

The cell in the first row and first column of a worksheet is called R1C1 (row 1, column 1) or A1 (row A, column 1), depending on the style selected. The styles just refer to labels -- changing the style has absolutely no effect at all on the content of worksheet cells.



The "Zoom all sheets" option will zoom <u>all</u> of a workbook's worksheets in, or out, depending on a setting found in row 57 of Lertap5's System worksheet. The example below shows the zoom factor set at 160.



Another way to zoom all sheets: right-click on the name of any sheet, then "Select all sheets", then adjust the slider on the bottom right (in the example above, the zoom is at 80%).

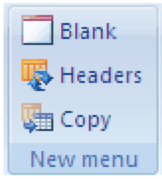
The "Page breaks" option is used to have Excel indicate where pages will break when a worksheet is printed. A more useful function might be Excel's built-in "Page Break Preview" option.

Additional information on most of these Excel options is available at an older version of this page -- click here to go there.

4.3 New menu

The New menu is the best way to create a new Excel workbook ready to work with Lertap. It's not the only way, but it's the best. Why? Because it results in an Excel workbook set up to display Lertap's preferred fonts, and a workbook with the two core Lertap worksheets, Data and CCs.

To find out what the New menu's options do, click on the options in this little snapshot:



It is possible to create a Lertap workbook without using the New menu options. All that's required by Lertap is an Excel workbook with data records in a worksheet named Data, and control lines (or "cards") in a worksheet named CCs.

If you presently have an Excel workbook with data records in it, you can try to rename the worksheet with the data records to Data, and insert a new worksheet with a name of CCs. That will work. However, font problems may develop when Lertap's Run menu options are taken -- Lertap's reports, such as that seen in the Stats1f worksheet, may turn out to be poorly formatted.

If this happens, a suggestion is to use the New menu to "Make a new blank Lertap 5 workbook" (the Blank option). Then, from the original workbook, copy all data records to the blank workbook's Data worksheet.

Click here for a definition of a Lertap workbook.

The New Menu is discussed in Chapter 3 of the manual, under the section titled "Setting up a new workbook". (Note: the examples in the manual are all based on older versions of Excel, prior to Excel 2007 and later.)

In the printed manual, the reference is page 60.

4.3.1 Blank

This option opens a new Excel workbook with two blank worksheets. One of the worksheets is named Data, while the other is named CCs.

The Data worksheet is split after Row 2 (the worksheet is fixed so that the top two rows always display header information).

Note that the new workbook should be saved as soon as possible. Excel's File menu options are used for this.

Click here for a definition of a Lertap workbook.

4.3.2 Headers

This option creates a new Excel workbook with Data and CCs worksheets which are empty of data records, but otherwise identical to the original workbook.

The Data worksheet's first two rows, the header rows, will have information copied from the original workbook's Data worksheet.

The lines in the new workbook's CCs worksheet will be identical to those in the original workbook's CCs worksheet.

Note that the new workbook should be saved as soon as possible.

Click here for a definition of a Lertap workbook.

4.3.3 Copy

This option creates a new Excel workbook which has Data and CCs worksheets copied from the original workbook. If the original workbook has more worksheets, they are not copied by this option -- only the Data and the CCs sheets get copied.

Given that this option only copies Data and CCs worksheets, how is a complete copy of a Lertap workbook made? With the original worksheet open, a complete copy may be made by using the Save As... option, found under Excel 2007's Office Button (see previous topic). Or, when the workbook's name is seen in a folder listing, right-click on the name, and then select Copy.

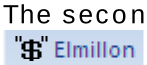
4.4 Run Menu

The Run menu provides paths to Lertap's data analysis routines. It's used after all data have been recorded in the Data worksheet, and after control lines, or "cards", have been typed into the CCs worksheet.

A typical data analysis procedure consists of two steps. First, users click on the


option. This gets Lertap to check the CCs lines for syntax errors, to "interpret" them. If no errors are found, Lertap goes on to produce the "Freqs" report, that is, a worksheet which summarises the response frequencies found in the columns of the Data worksheet. Which columns? The ones referenced by the *col lines in the CCs worksheet.

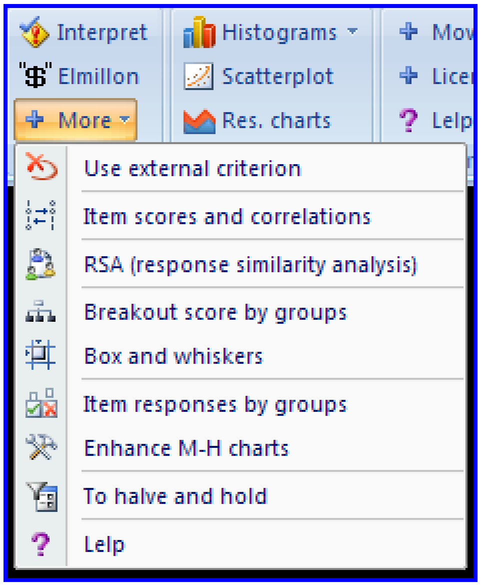
As part of the interpret CCs lines process, Lertap also writes some intermediate worksheets with data required by the Elmillon program. These are the "Sub" worksheets -- users with a keen eye can see the Sub worksheets being formed as Lertap goes about its business, but then, just before focus shifts to the Freqs worksheet, the Sub worksheets are hidden from view. Users usually have little need to see them; however, they're not secret -- they can be unhidden. There will be one Sub worksheet for each of the CCs worksheet's *col lines.

The second of Lertap's usual two-step data analysis process involves clicking on the


option. This is the option which creates the various statistical reports which are Lertap's reason for being. Each of these reports is an Excel worksheet; they have names such as Stats1f, Stats1b, and Statsul.

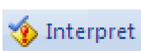
Elmillon also produces scores, one for each subtest, or scale.

For more information on what the Run menu's options do, click on the topics shown in the boxes below:



Note: the Rasch Analysis option is not shown above. It is described here.

4.4.1 Interpret CCs lines



Before this option is selected, it must be the case that (1) data records have been created in the Data worksheet, and (2) control lines, or "cards", have been placed in the CCs worksheet.

The way the Interpret option works depends on a setting in row 7 of Lertap's System worksheet where the "user run mode" is controlled.

By far the most common setting is "normal mode". The discussion below, in the rest of this topic, assumes that this is the setting. A following topic discusses the "Elmillon direct" user run mode.

When Lertap's "user run mode" is set to the normal mode, then, the first time the Interpret option is clicked on, Lertap gets Excel to read the responses found in the Data worksheet, looking in the columns specified in the *col lines of the CCs worksheet.

If there are no errors in the CCs worksheet, Lertap and Excel will produce new worksheets. One of them is called "Freqs", for frequencies.

Usually, the "Elmillon item analysis" option will be taken next. This results in even more new worksheets, with names such as Scores, Stats1f, Stats1b, and so on.

It is possible to get Lertap to automatically roll through from "Interpret CCs lines" to "Elmillon item analysis" without stopping, as detailed in the following topic: Production mode.

If the "Interpret CCs lines" is later selected again, a warning message will appear, saying that there's a possibility of losing data. This is purely precautionary -- Lertap is about to delete Freqs, Scores, Stats1f, and so forth -- but this is usually only natural; new versions of these same worksheets will be generated, which is almost always what users want. (Lertap is being overly cautious in sounding this warning; the primary worksheets, Data and CCs, are never affected by this action. Nor are any other special user-created worksheets, such as, perhaps, the coding guide seen in this example.)

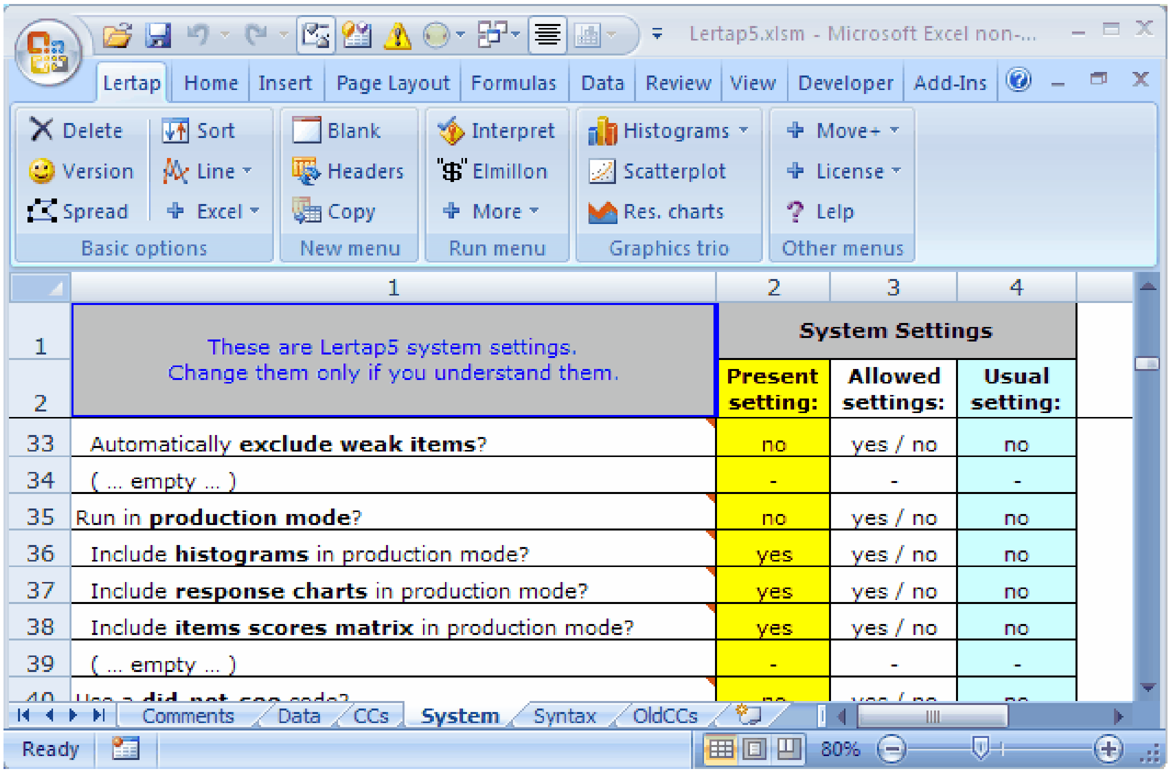
For a related topic, see "deleting secondary worksheets".

Chapter 9 in the printed manual also discusses these matters.

4.4.1.1 Production mode

A "production mode" capability may be activated by changing a setting in Lertap's System worksheet. It allows for an almost totally automated execution of Lertap; when the Interpret option is taken and production mode is in use, then the Elmillon item analysis takes place automatically after the Interpret option has finished its tasks. In fact, there are more settings which extend what Lertap will do without any user intervention. Please read on.

Have a look at the following snapshot of lines 35 through 38 of the System Worksheet, taken back in June, 2009:



If the production mode setting is set to "yes", then Lertap will not stop after it has been requested to "Interpret CCs lines" -- it'll power ahead, automatically activating the next option on the Run menu, "Elmillon item analysis".

There are three other "yes" settings which may be made in these rows -- you see them above. You can get Lertap to automatically output histograms for each subtest, response charts (bar charts for affective subtests, quintile plots for cognitive subtests), and an item scores matrix (IStats). This gives you time to sit back with your cup of coffee, and watch the screen flash before you as Lertap goes about its tasks.

Of course, this is not recommended. Lertap ships with all the production mode options set to "no". Why? Because sound data processing practice is always supposed to involve a data integrity check or two. You want to make sure the data you're feeding into Lertap has been subjected to some quality control. In Lertap 5, the main means of doing this is by having a careful look at the Freqs report, the worksheet produced by taking the "Interpret CCs lines" option from the Run menu.

The Freqs report readily indicates the characters found in each of the data columns. If you're running with a cognitive test whose options use the letters A, B, C, and D, then you'll want to check down the Freqs report to make sure that no other characters have crept into the scene, such as, perhaps, lower-case letters (a, b, c, or d).

If your scan of Freqs reveals strange characters, your task is to find them, and to fix 'em. It's only after you've done this that you'll go on to get results by activating Elmillon.

But there are times when users do not concern themselves so much with Freqs. For example, many users have their data prepared by using a scanner. Scanners can readily be trained to hoot and holler when they encounter bad data. Oft times scanners will output a special character, such as an asterisk, to signal a questionable result.

For more on the System worksheet, click here. The production mode options are all dynamic ones -- they take effect immediately, without requiring that Lertap first be closed and then reopened.

Note that the yes / no entry in row 26 controls the following three lines. If row 26 is set to "no", then that's it -- the following three rows are ignored, even if they have "yes" entries.

Finally, there will be users who'd like to have two versions of Lertap: one as shipped from our manufacturing plant, with production mode off, and another version with production mode on.

Can do, emu? Yes. A single computer, Windows or Macintosh, can have multiple copies of the Lertap5.xlsm file. Of course, each copy has to be in a distinct folder (subdirectory), but that's easy. If you're going to do this, and you're running under Windows, then you might like to copy the LRTP5HHelp.chm to each folder too -- it's the Lertap help file, "Lelp". If you have Lertap5.xlsm in a folder, without its corresponding Lelp file, then Lertap help will not be available whenever you try to access it from within Lertap (this comment applies only to Windows users).

Related tidbit:

For more about this topic, get productive: see "ScannerEjemplo1.doc", a marvelous, captivating Word document with lots of hints for what to do on rainy days, available via the Internet: click here if you've got connections.

4.4.1.2 User run mode

There's something a bit inconvenient about the way Lertap's Interpret option works. It will very often be the case that users will change the CCs lines, or add new ones, after Interpret and Elmillon have already been run. Once this is done, the Interpret option is again taken and, as it starts up, it announces that it's going to delete the worksheets created the first time.

This can be unnecessary and inconvenient at times. A special user run mode setting called "Elmillon direct" provides a way to stop Interpret from deleting worksheets. Read on to find out what this special setting does.

Let's say, for example, that a user has a 50-item multiple-choice cognitive test and has set up three CCs lines: *col, * sub, and *key.

After running the Interpret and Elmillon options, Lertap will have created several new worksheets, or reports. Among them will be sheets called "Scores", "Stats1f", "Stats1b", and "Stats1ul". The "1" in these names refers to what Lertap calls the first "subtest" -- there will always be one Lertap "subtest" for each *col line found in the CCs lines.

Behind the scenes, hidden from view, Lertap will also have created another worksheet with very detailed subtest information. It will be called "Sub1". This is a vital worksheet generally meaningful only to Lertap itself -- were you to look at Sub1 you might not sense its value. It might appear rather mysterious to you, but to Lertap it's truly vital. Sub1 is created as the Interpret option goes about its business. Elmillon will not work if Sub1 cannot be found.

Read more about these worksheets and Lertap's internal goings-on at the very readable and recommended "Lertap output" topic.

Now, suppose something arises to cause the user to add another set of *col, *sub, and *key lines. There may have been a mis-keyed item in the first *key line, for example, and it's desired to see the effect of correcting the error without erasing the initial results. Or, perhaps a user wishes to run the same subtest through Interpret and Elmillon but with a mastery test setting added to the *sub line; the mastery analysis is wanted in addition to the initial analysis.

In both of these scenarios users will end up with two sets of *col, *sub, and *key lines. In each set, the *col lines may be the same, but there will be differences in the *sub lines, where new name=() and title=() settings are likely to be used, and, in the second scenario, a mastery= setting will be inserted in the *sub line. The *key line will change in the first scenario as one item will be re-keyed.

Okay? What have we got? In each of these two examples, as mentioned, there will now be two *col lines. Ready, set, go: the user clicks on Interpret -- in "normal" user run mode, Interpret erases all results previous obtained. In the "Elmillon direct" mode it does not. Instead it points to each of the *col lines and asks the user if the corresponding subtest is to be re-processed. If it's the first *col line, the first subtest, and the user says "*No, I don't want you to process this again*", then Interpret leaves the Stats1f, Stats1b, Stats1ul, and Sub1 sheets as they were; it doesn't overwrite them.

What's the big deal here, what's the advantage? Time. In large datasets, with thousands of Data records and many test items, it takes time for Interpret and Elmillon to do their jobs -- sometimes more than two or three minutes (one of the "tidbits" below has time-trial data). Skipping over the first subtest will save time. Additional time will be saved by users who have created quantile plots by using the "Res. charts" options. Such plots take extra time to make, and users may have invested additional work by adjusting page margins so that the plots look better when printed (refer to this document for related comments).

This work will be lost if the Interpret option is taken again. Unless, that is, the "user run mode" option is set to "Elmillon direct mode". Page to the following topic to read about how to set the Elmillon direct option, and how to use it.

Related tidbits:

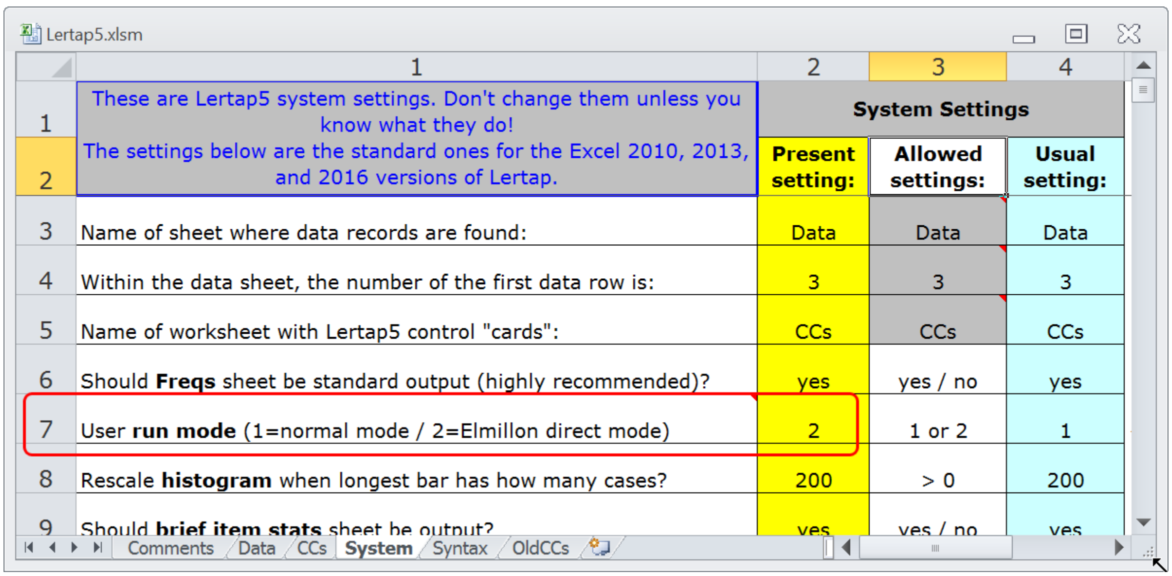
The user run mode setting is found in Lertap versions 5.10.7.2 and later. Look here to see how to determine version numbers. Users with a version of Lertap purchased from Assessment Systems Corporation may be eligible for a special upgrade.

Click here to branch out to a document with time trials.

The "Elmillon direct" run mode was a feature in much earlier versions of Lertap 5 where it was know as the "Liberty Bell" option. It has been resurrected after requests from users.

4.4.1.2.1 Elmillon direct

The Elmillon direct option is activated by making a change in row 7 of the System worksheet:


Lertap5.xlsm

	1	2	3	4
1	These are Lertap5 system settings. Don't change them unless you know what they do!	System Settings		
2	The settings below are the standard ones for the Excel 2010, 2013, and 2016 versions of Lertap.	**Present setting:**	**Allowed settings:**	**Usual setting:**
3	Name of sheet where data records are found:	Data	Data	Data
4	Within the data sheet, the number of the first data row is:	3	3	3
5	Name of worksheet with Lertap5 control "cards":	CCs	CCs	CCs
6	Should **Freqs** sheet be standard output (highly recommended)?	yes	yes / no	yes
7	User **run mode** (1=normal mode / 2=Elmillon direct mode)	2	1 or 2	1
8	Rescale **histogram** when longest bar has how many cases?	200	> 0	200
9	Should **brief item stats** sheet be output?	yes	yes / no	yes

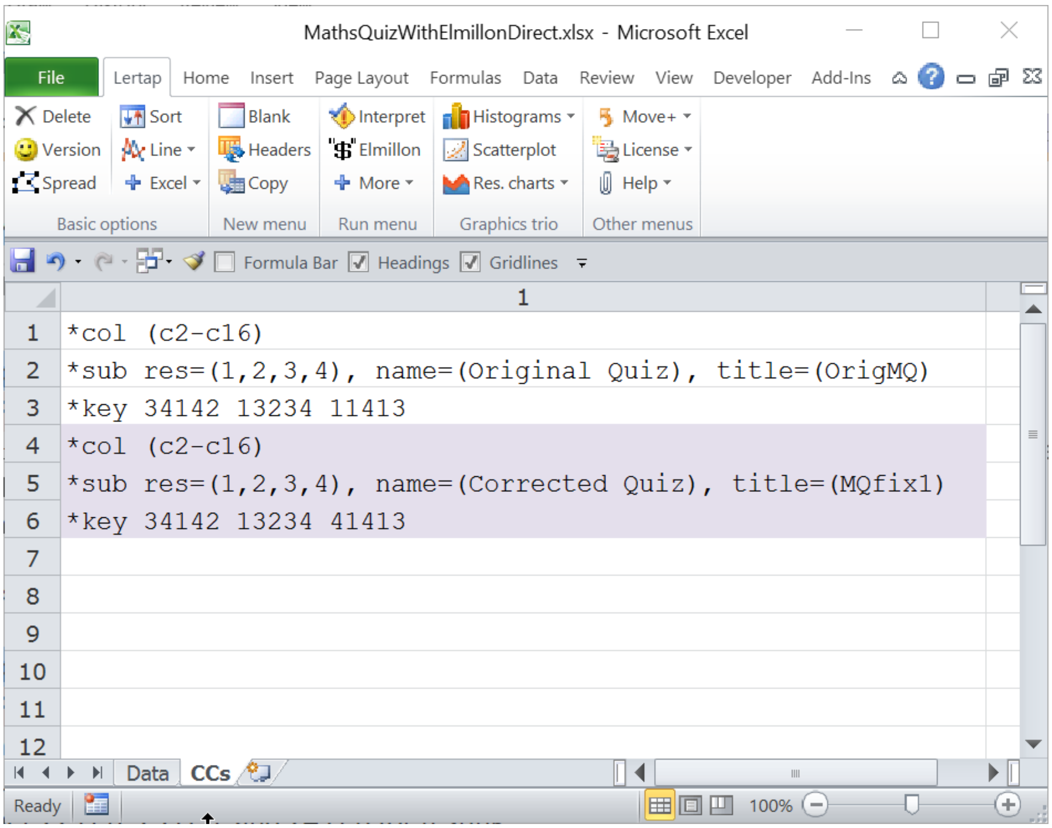
The screen snapshot above shows how to set the Elmillon direct option: put 2 in the second column of the 7th row of the System worksheet. Remember to save the Lertap5 workbook, Lertap5.xlsm, after you've done this -- put that 2 in the second column and then go up to Excel's File menu and click Save.

Now, before we explain how Elmillon direct works we should remind you of the discussion of the "Sub1" worksheet mentioned in the previous topic. It gets created when the Interpret option is taken, and it's always hidden from view when the user run mode is set to normal mode.

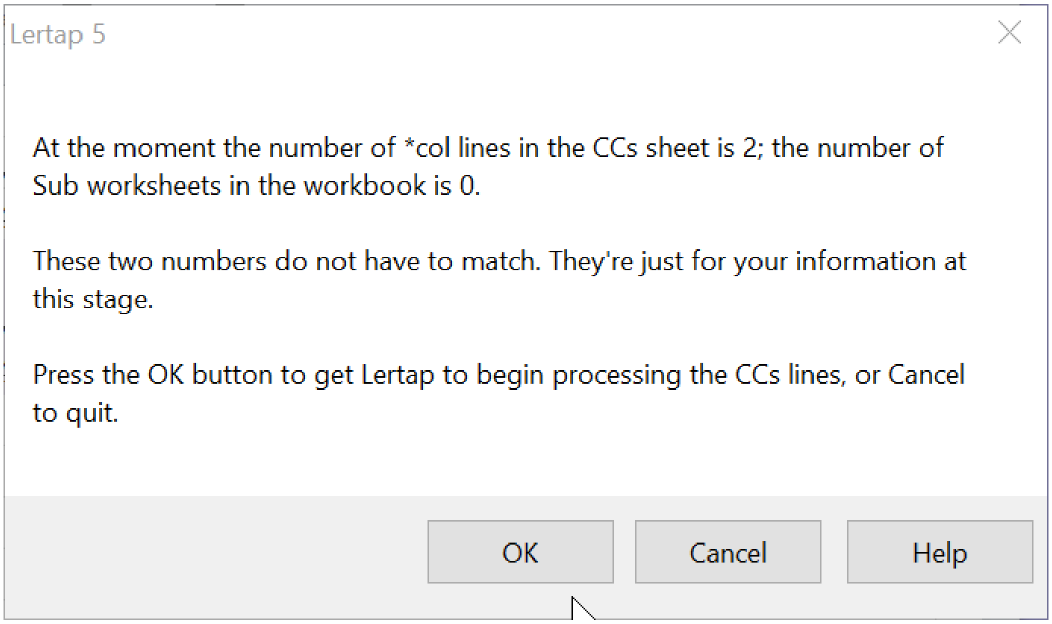
If there happens to be more than one *col line in the CCs worksheet, and if Lertap is running in normal mode, then, when Interpret runs, it will create one Sub worksheet each time a *col line is interpreted and processed by the Interpret option. If there are two *col lines in the CCs sheet, Interpret will create "Sub1" and "Sub2" worksheets. And they will be hidden from view, ready for subsequent use when the Elmillon option is used.

Things operate in a markedly-different manner when the Elmillon direct setting is on, that is, when there's a 2 in the second column of row 7 of the System worksheet.

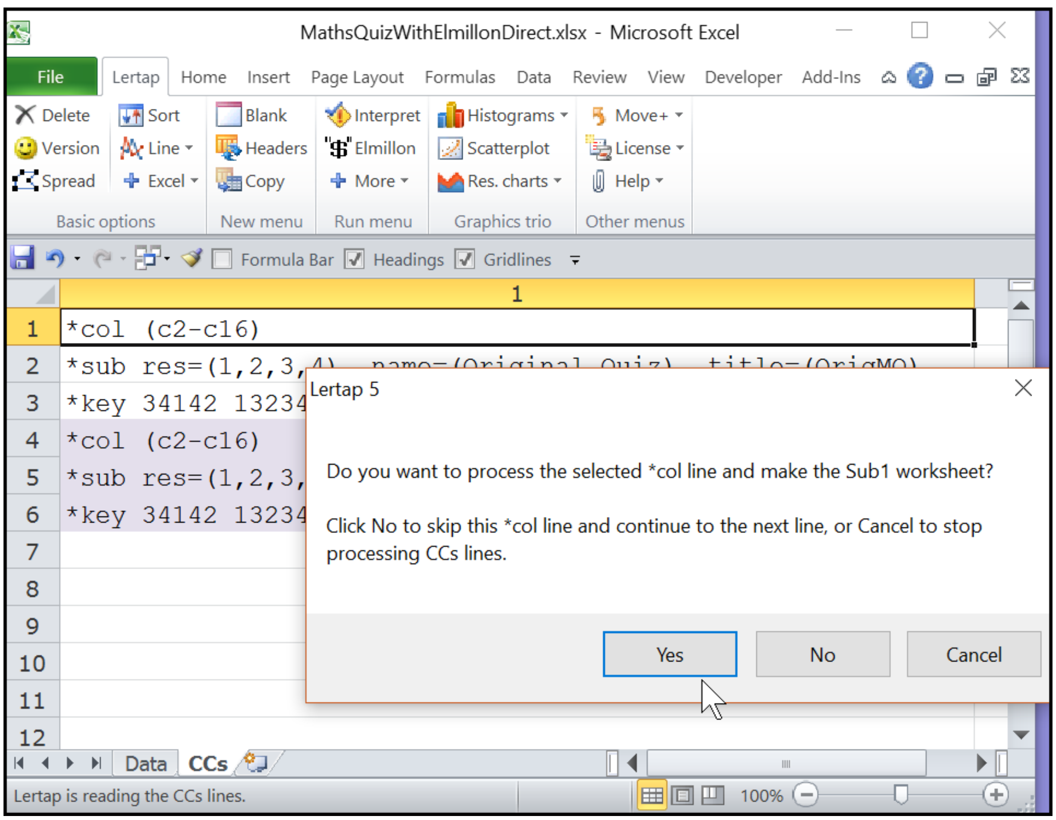
Consider the following CCs worksheet:



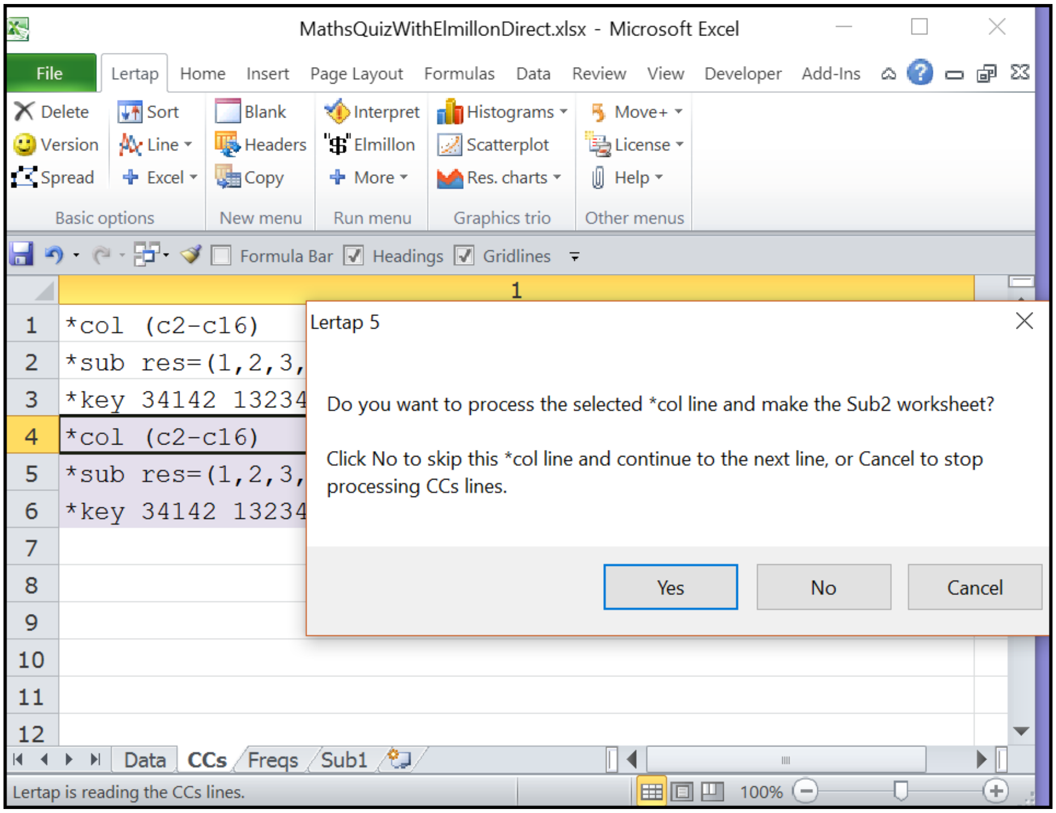
When the Elmillon direct mode is active, a click on the Interpret option produces this message:



Click OK and the following snap shows what happens; note that the first *col line has been selected by Excel:

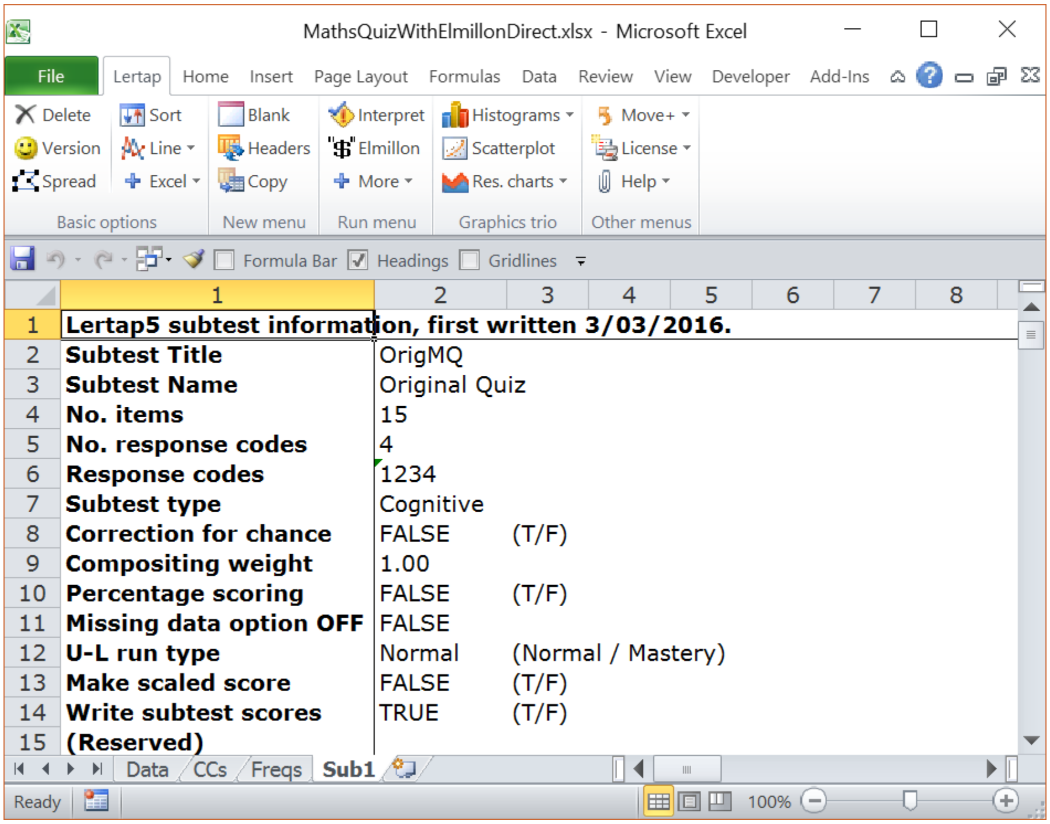


Click Yes. Lertap creates Freqs and the Sub1 worksheets. Excel selects the next *col line and presents another question:

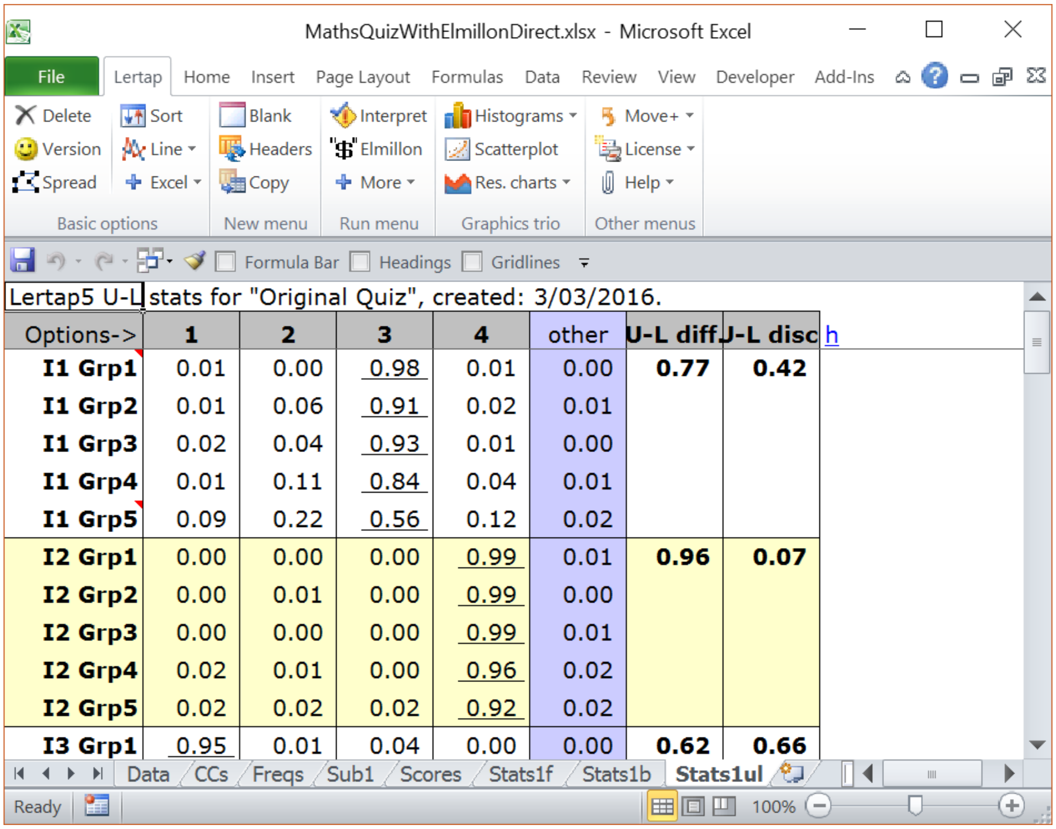


Click No and Lertap stops processing the CCs lines.

To get results for the first subtest, click on the Sub1 tab at the bottom of the screen. The top rows of the Sub1 sheet will display; note the title and name recorded in rows 2 and 3 -- these have been picked up from the first *sub line in the CCs worksheet.

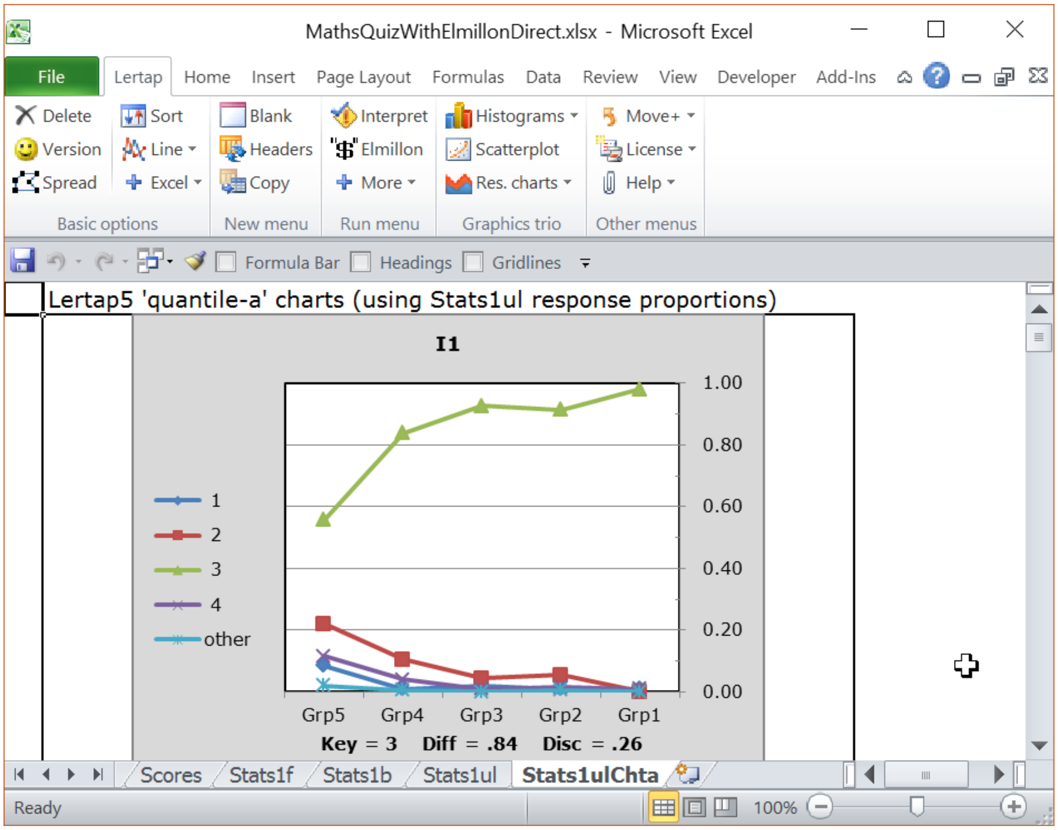


Now click on the Elmillon option. It starts up immediately and gets right down to work, no questions asked.

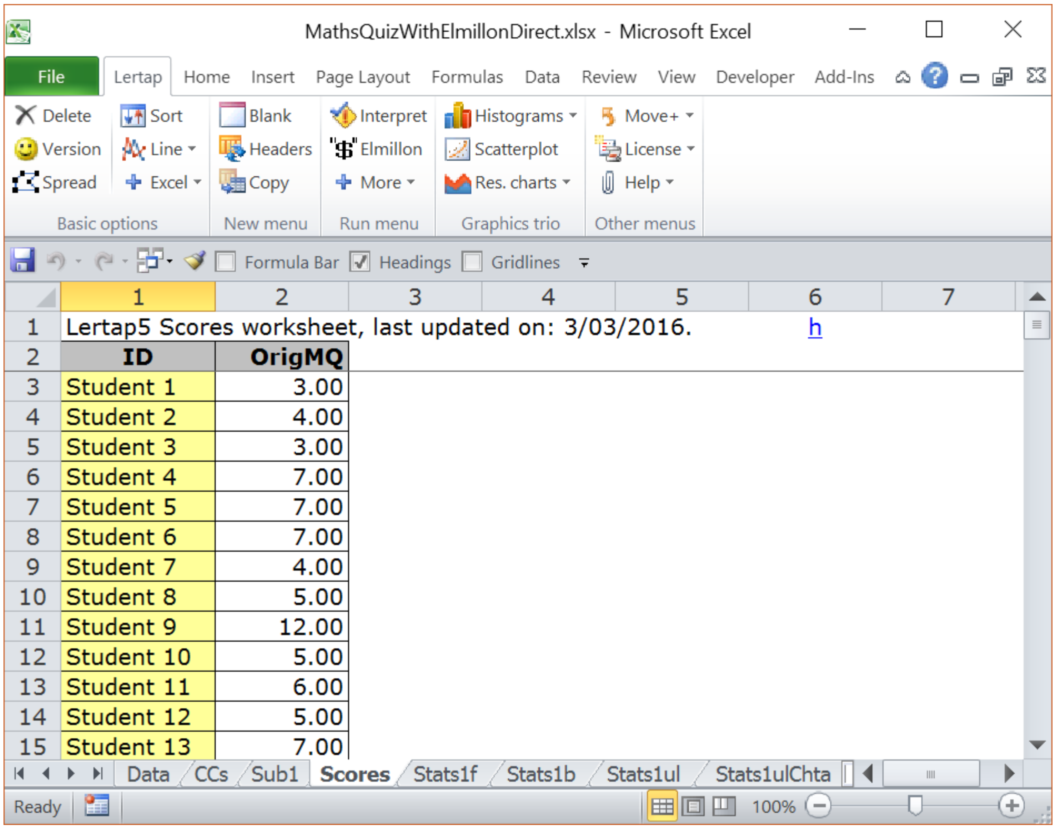


Elmillon completes its tasks and, with no announcement, sets the focus on the Stats1ul sheet. Now there are eight tabs showing at the bottom of the screen, from Data on the left to Stats1ul on the right.

Suppose that at this point the user clicks on the Res. charts option in order to get quantile plots. Lertap adds another worksheet with the plots; Stats1ulChta is its name.



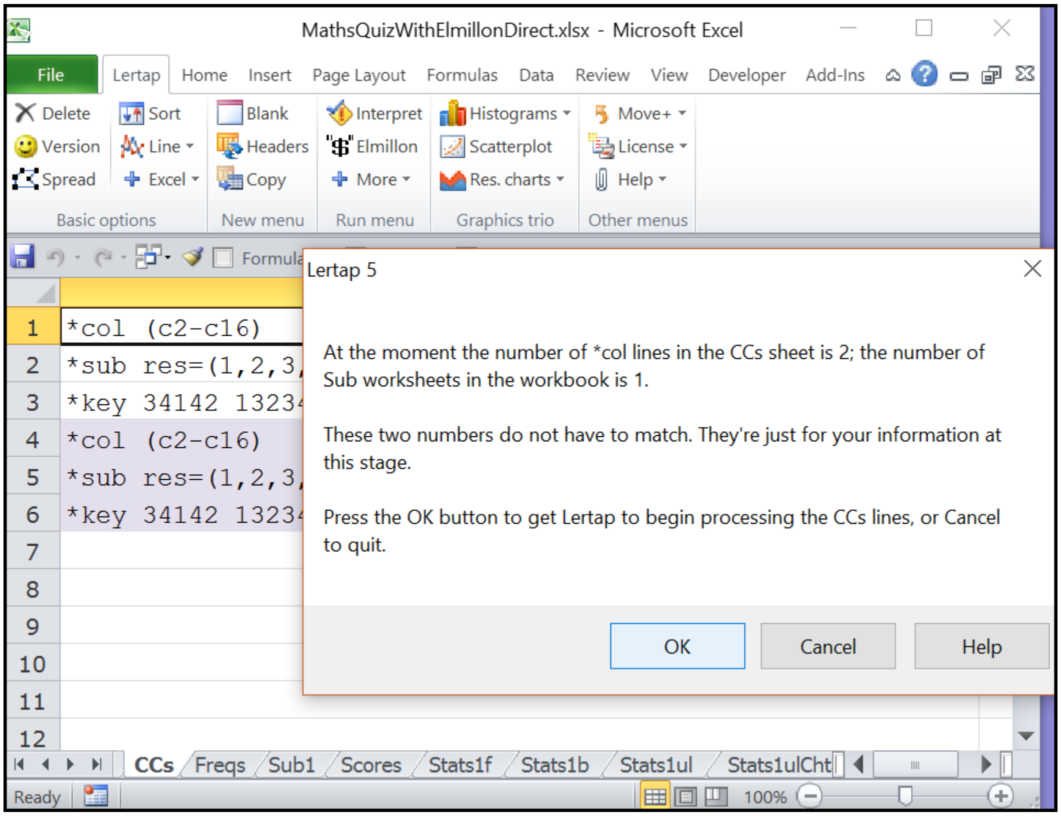
The Scores worksheet will look as follows at this stage:



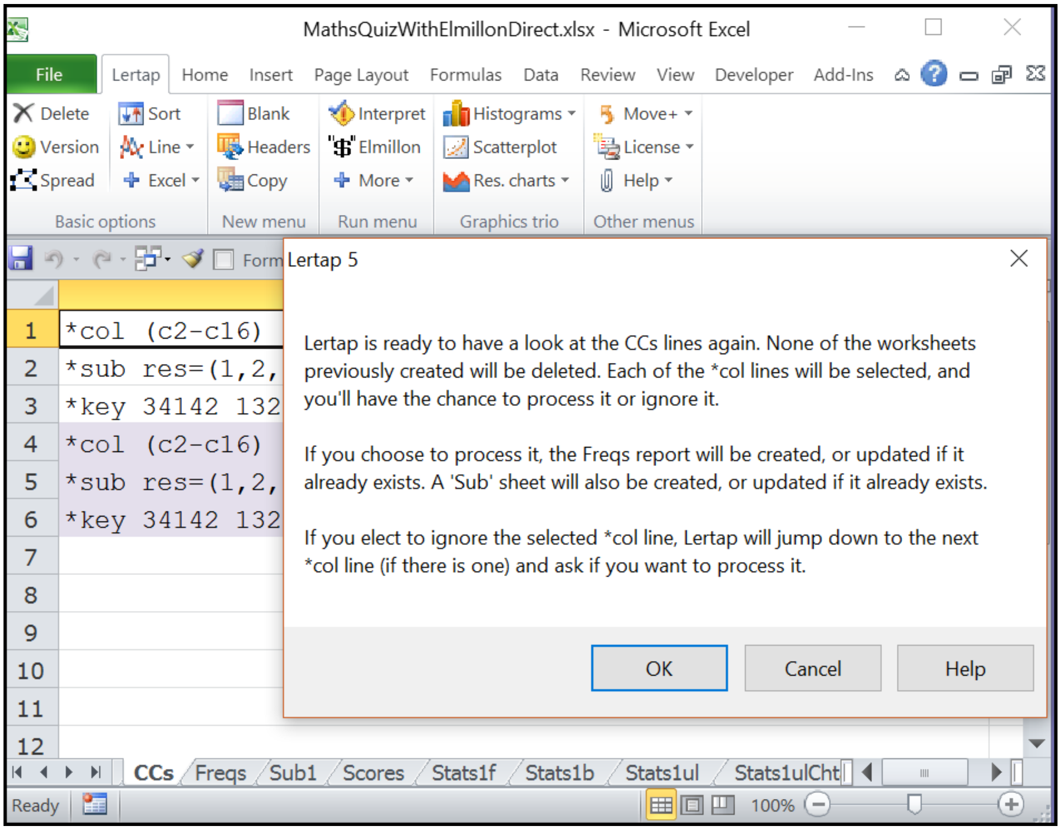
A full analysis of the various Stats reports, and a scan down the quantile plots, will reveal that there was an error in the *key line for the first subtest. The key for the 11th item was supposed to be 4, not 1 as seen in that *key line. This prompted the original researcher to add the second set of *col, *sub, and *key lines seen above in the CCs sheet. Note the name=() and title=() settings on the second *sub card -- they're different. There has also been a change on the second *key line (the entry for the 11th item has changed).

True enough; the researcher could have simply changed the first *key line and run the analysis again without creating a new set of *col, *sub, and *key lines. But that would have ruined this example (!) -- and, from a practical viewpoint, adding the new CCs lines will make it easier to see the effect of correcting the error in the first *key line. But first, of course, we need to process those new lines.

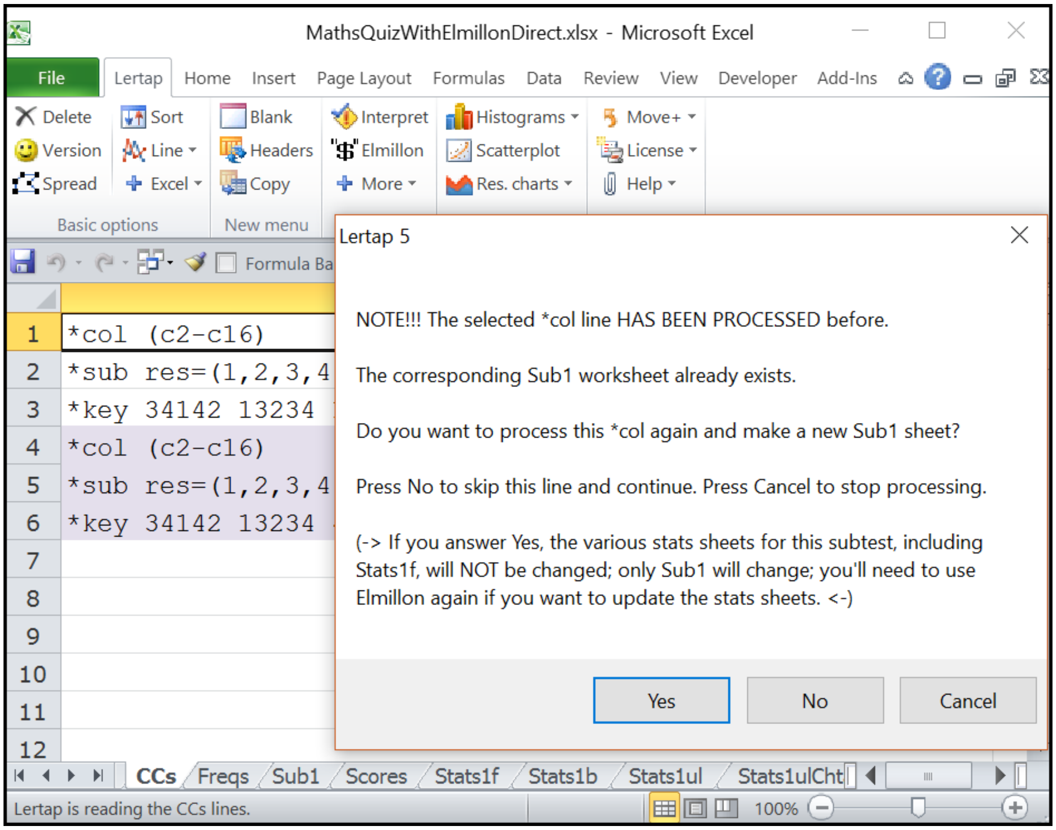
Back to the Interpret option.


MathsQuizWithElmillonDirect.xlsx - Microsoft Excel
File
Lertap
Home
Insert
Page Layout
Formulas
Data
Review
View
Developer
Add-Ins
Delete
Version
Spread
Sort
Line
Excel
Blank
Headers
Copy
Interpret
Elmillon
More
Histograms
Scatterplot
Res. charts
Move+
License
Help
Basic options
New menu
Run menu
Graphics trio
Other menus
Lertap 5
At the moment the number of *col lines in the CCs sheet is 2; the number of Sub worksheets in the workbook is 1.
These two numbers do not have to match. They're just for your information at this stage.
Press the OK button to get Lertap to begin processing the CCs lines, or Cancel to quit.
OK
Cancel
Help
*col (c2-c16)
*sub res=(1,2,3,
*key 34142 1323
CCs
Freqs
Sub1
Scores
Stats1f
Stats1b
Stats1ul
Stats1ulCht
Ready
100%

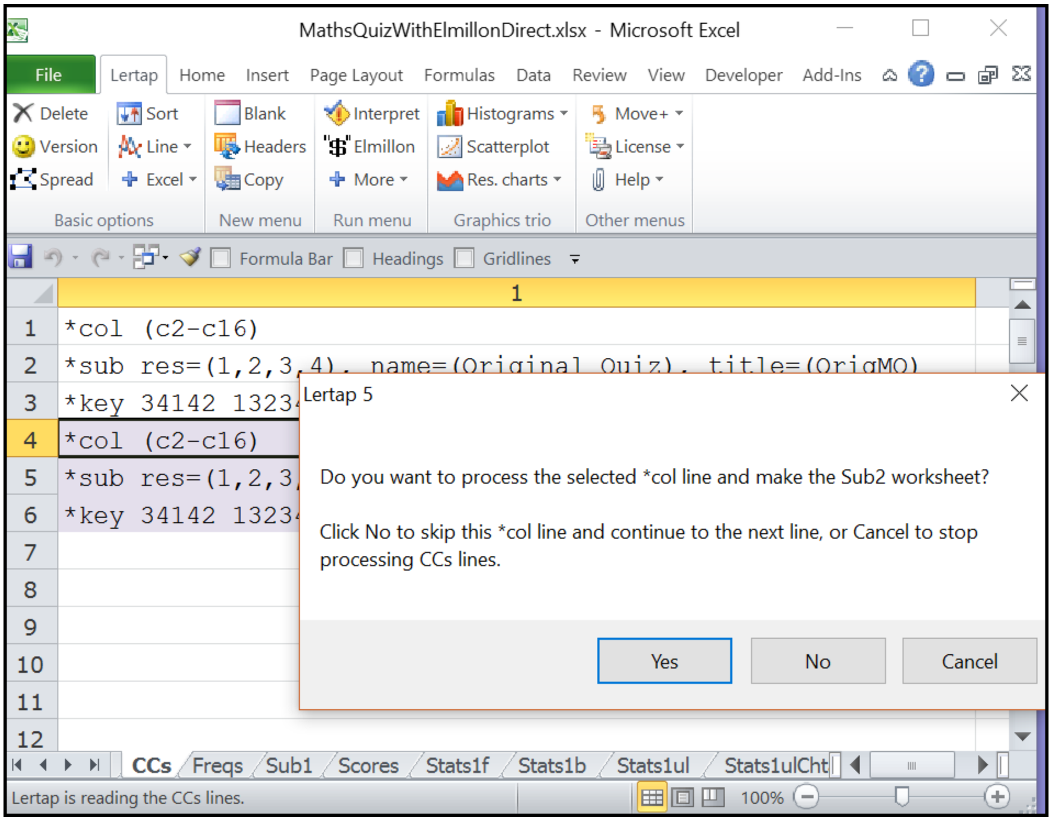
Click OK.


MathsQuizWithElmillonDirect.xlsx - Microsoft Excel
File
Lertap
Home
Insert
Page Layout
Formulas
Data
Review
View
Developer
Add-Ins
Delete
Version
Spread
Sort
Line
Excel
Blank
Headers
Copy
Interpret
Elmillon
More
Histograms
Scatterplot
Res. charts
Move+
License
Help
Basic options
New menu
Run menu
Graphics trio
Other menus
*col (c2-c16)
*sub res=(1,2,
*key 34142 132
*col (c2-c16)
*sub res=(1,2,
*key 34142 132
Lertap 5
Lertap is ready to have a look at the CCs lines again. None of the worksheets previously created will be deleted. Each of the *col lines will be selected, and you'll have the chance to process it or ignore it.
If you choose to process it, the Freqs report will be created, or updated if it already exists. A 'Sub' sheet will also be created, or updated if it already exists.
If you elect to ignore the selected *col line, Lertap will jump down to the next *col line (if there is one) and ask if you want to process it.
OK
Cancel
Help
CCs
Freqs
Sub1
Scores
Stats1f
Stats1b
Stats1ul
Stats1ulCht
Ready
100%

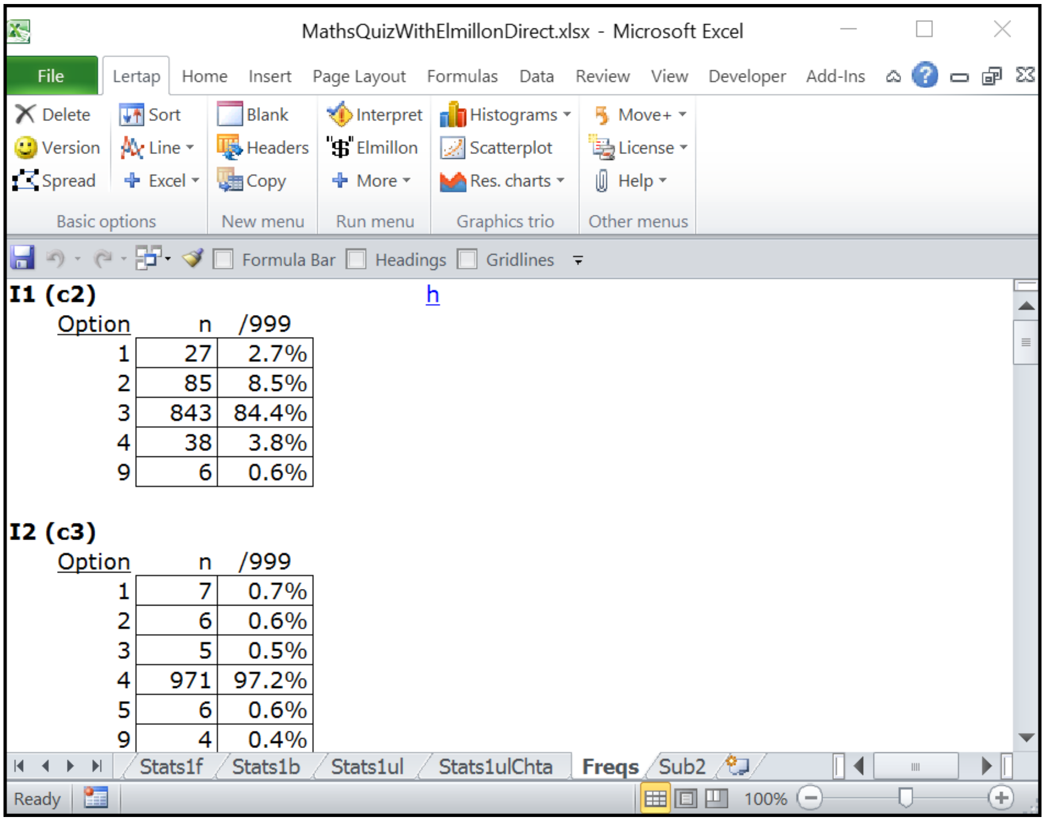
Click OK.


MathsQuizWithElmillonDirect.xlsx - Microsoft Excel
File
Lertap
Home
Insert
Page Layout
Formulas
Data
Review
View
Developer
Add-Ins
Delete
Sort
Blank
Interpret
Histograms
Move+
Version
Line
Headers
Elmillon
Scatterplot
License
Spread
Excel
Copy
More
Res. charts
Help
Basic options
New menu
Formula Ba
*col (c2-c16)
*sub res=(1,2,3,4
*key 34142 13234
*col (c2-c16)
*sub res=(1,2,3,4
*key 34142 13234
Lertap 5
NOTE!!! The selected *col line HAS BEEN PROCESSED before.
The corresponding Sub1 worksheet already exists.
Do you want to process this *col again and make a new Sub1 sheet?
Press No to skip this line and continue. Press Cancel to stop processing.
(-> If you answer Yes, the various stats sheets for this subtest, including Stats1f, will NOT be changed; only Sub1 will change; you'll need to use Elmillon again if you want to update the stats sheets. <-)
Yes
No
Cancel
CCs
Freqs
Sub1
Scores
Stats1f
Stats1b
Stats1ul
Stats1ulCht
Lertap is reading the CCs lines.
100%

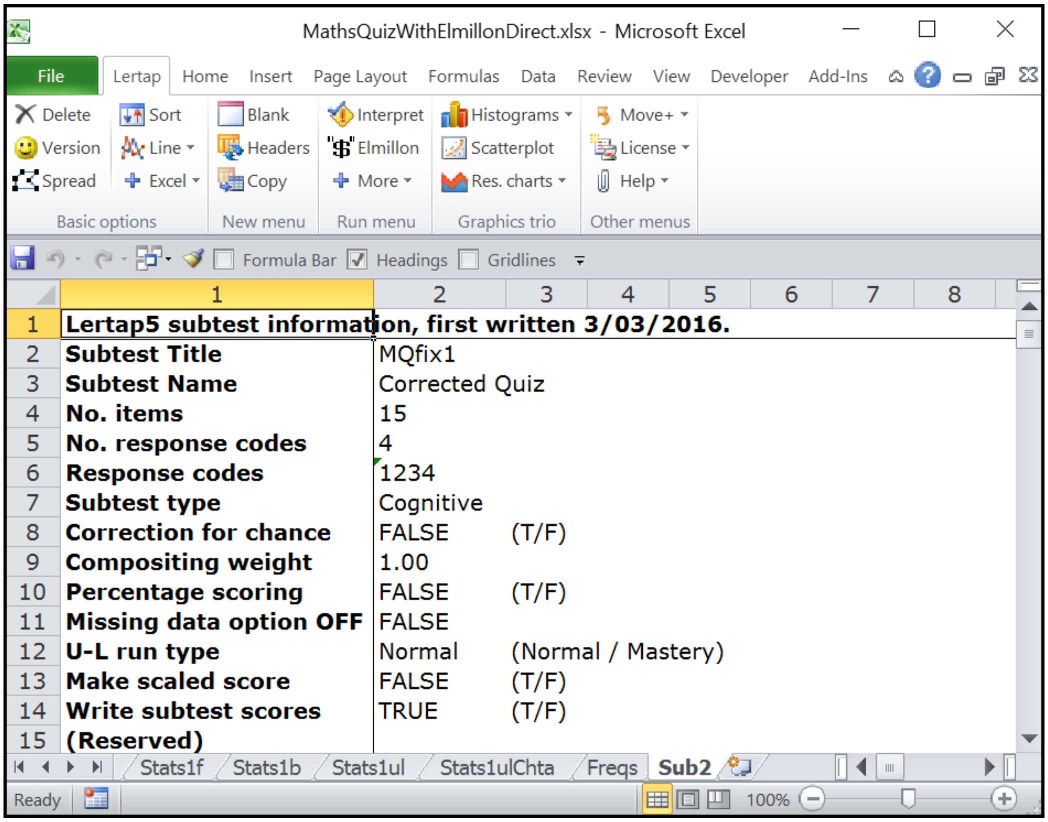
Click No.



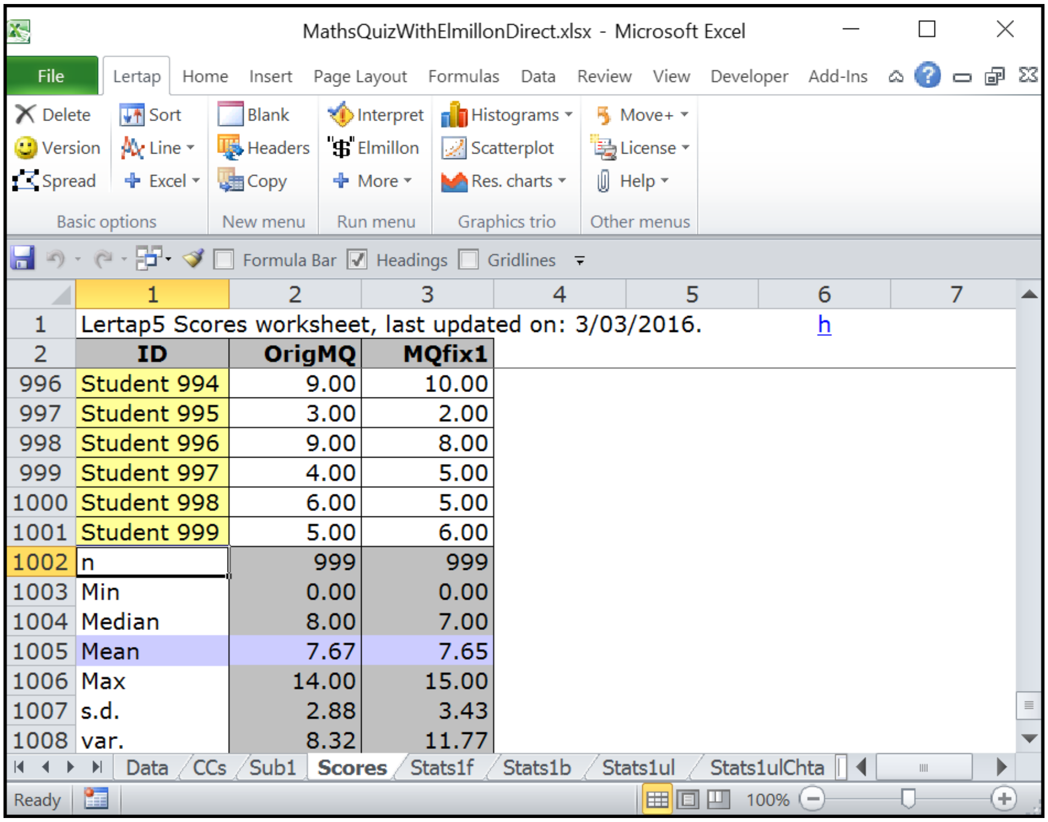
Click Yes. Lertap updates the Freqs report and adds the Sub2 worksheet. The original results for the first subtest, including the quantile plots, are still there.



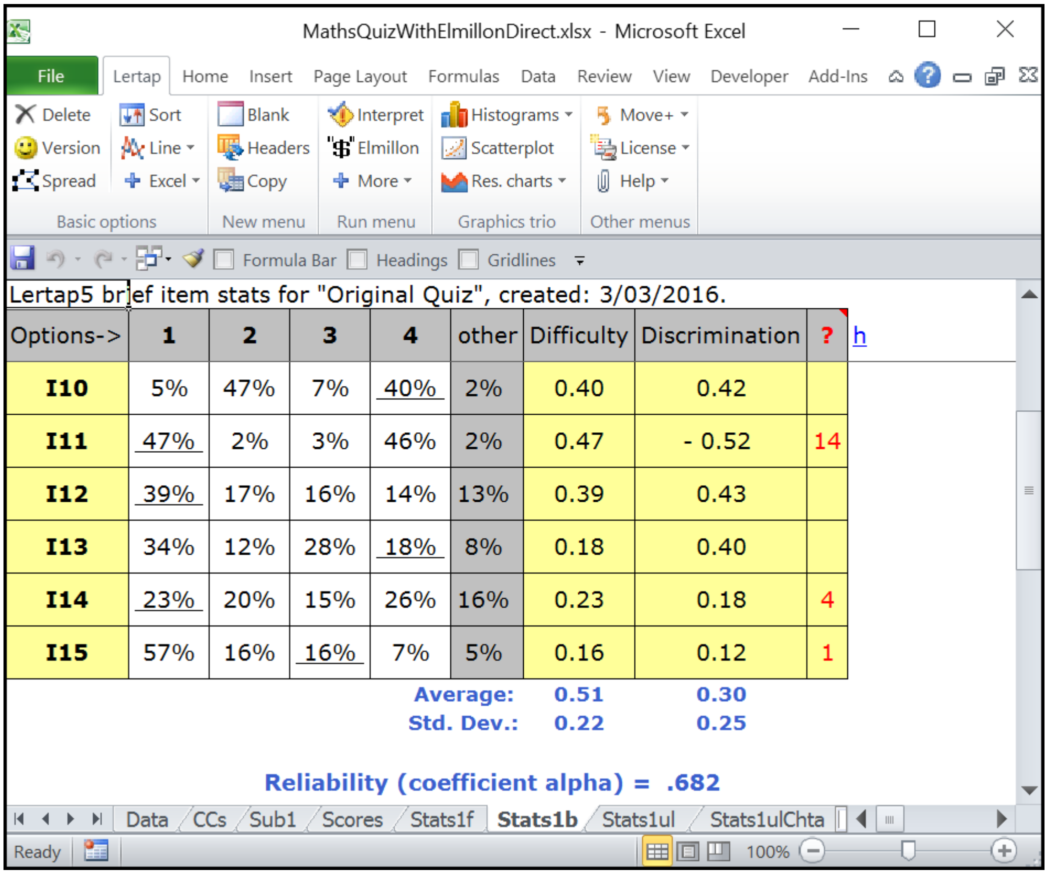
The Sub2 sheet looks like this:



Now a click on the Elmillon option will initiate the creation of Stats2f, Stats2b, csem2, and Stats2ul reports. It will also result in a new score being added to the Scores worksheet:



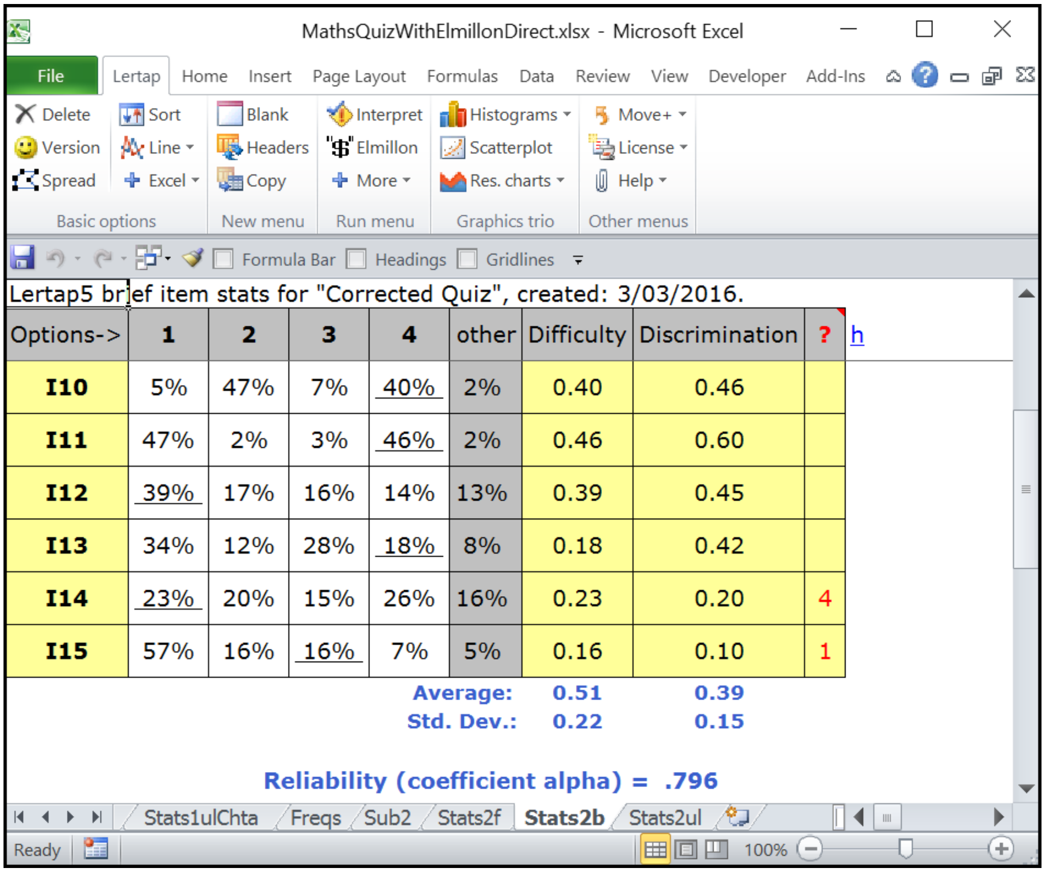
To see the effect of correcting the mis-keyed item, flick back and forth from Stats1b to Stats2b; remember that the problem was with I11, the 11th item:



Lertap5 brief item stats for "Original Quiz", created: 3/03/2016.

Options->	1	2	3	4	other	Difficulty	Discrimination	?
I10	5%	47%	7%	40%	2%	0.40	0.42	
I11	47%	2%	3%	46%	2%	0.47	- 0.52	14
I12	39%	17%	16%	14%	13%	0.39	0.43	
I13	34%	12%	28%	18%	8%	0.18	0.40	
I14	23%	20%	15%	26%	16%	0.23	0.18	4
I15	57%	16%	16%	7%	5%	0.16	0.12	1
					Average:	0.51	0.30	
					Std. Dev.:	0.22	0.25	

Reliability (coefficient alpha) = .682



Lertap5 brief item stats for "Corrected Quiz", created: 3/03/2016.

Options->	1	2	3	4	other	Difficulty	Discrimination	?
I10	5%	47%	7%	40%	2%	0.40	0.46	
I11	47%	2%	3%	46%	2%	0.46	0.60	
I12	39%	17%	16%	14%	13%	0.39	0.45	
I13	34%	12%	28%	18%	8%	0.18	0.42	
I14	23%	20%	15%	26%	16%	0.23	0.20	4
I15	57%	16%	16%	7%	5%	0.16	0.10	1
					Average:	0.51	0.39	
					Std. Dev.:	0.22	0.15	

Reliability (coefficient alpha) = .796

The Interpret option may be run again at any time. It will go down the CCs lines in the same manner as above. The Elmillon option may also be run again, starting from either the Sub1 or the Sub2 worksheets. For example, if at this stage the user clicks on Sub2 and then on Elmillon, all the reports created for the second subtest, such as Stats2f and Stats2b, will first be deleted and then written again, in effect updating them with any changes that may have been made in Sub2.

How would changes be made to Sub2? Most users would begin by making changes in the set of *col, *sub, and *key lines for the second subtest, and then running the Interpret option again to produce a fresh Sub2 sheet. However, for those whose life insurance policies are up to date, it is possible to make changes directly in the Sub2 worksheet itself. This topic explains how and why, using text from an old help system from the version of Lertap5 made to work with Excel 97. (It makes mention of the "Advanced Toolbar". This toolbar can be made to show in recent versions of Excel if the "Add-ins" tab is added to Excel's ribbon -- search Excel's help system for information on how to get the Add-Ins tab to show.)

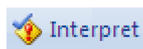
Related tidbits:

This discussion is based on real events and actual test data. The original dataset may be downloaded from this site.

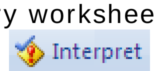
With regard to life insurance policies: making changes directly in Sub worksheets (as opposed to making them in *col, *sub, *key and then running Interpret) very effectively breaks the correspondence between the CCs worksheet and the results worksheets. For example, a subtest's name and title may readily be changed in the Sub worksheet and the changes will be reflected in the reports made by Elmillon; but the name=() and title=() settings in the corresponding CCs *sub line will then be inaccurate. Risky stuff, users might want to be insured :)

4.4.2 Elmillon

Elmillon is the name of Lertap's main item analysis routine, first developed for the Venezuelan Ministry of Education. The name has origins in the Spanish language: "*un millon*", or, in English, "*thanks a million*", is what the chief of the data analysis section, Rogelio Blanco, said when the routine was debugged and delivered. The name Elmillon has been used since then.

Elmillon is always run after the 

option has been taken. Elmillon reads data, makes subtest and scale scores (written to the Scores worksheet), and goes on to produce from one to three item analysis reports. The item analysis reports are Excel worksheets, having names like Stats1f, Stats1b, and Stats1ul.

Once Elmillon has been run, it will be reluctant to perform again until the workbook's secondary worksheets have been deleted. This isn't as difficult as it may sound: selecting 

will delete them.

Related tidbits:

The discussion above relates to running Elmillon when the user run mode is set to "normal". Its behavior is somewhat different when user run mode is set to "Elmillon direct"; in this mode the secondary worksheets referred to above will not necessarily be deleted. Please refer to this topic.

Read more about making sense of Elmillon's various reports in the manual.

4.4.3 External criterion

The item correlation coefficients which form part of Lertap's Stats1f and Stats1b reports are based on correlating item scores with what's referred to as an "internal criterion": the number produced by scoring the remaining items in the subtest or scale to which the item belongs.

It is possible to replace the internal criterion with another score. This other score is referred to as an "external criterion". An external criterion may be used as part of the process of validating test items.

In Lertap, the external criterion score must correspond to a column in the Scores worksheet. The analysis begins by having the user pick the Scores column that has the score to be used as the external criterion. Once this is done, Lertap asks the user to pick out the subtest which has the items to be correlated with this score. This is done by displaying Sub worksheets.

Users of the external criterion analysis feature will often have a criterion measure which needs to be imported to the Scores worksheet. In this case, the criterion measure should be recorded in a column in the Data worksheet. Once it's there, the Move menu on the Lertap toolbar will allow the measure to be copied over to the Scores worksheet.

The effects of part-whole inflation may be examined by using an external criterion analysis. Lertap's item correlation coefficients are always corrected for part-whole inflation (sometimes referred to as part-whole contamination); to see what they'd be without such correction, define a subtest's score, as found in the Scores worksheet, to be the external criterion. (In some texts and other item analysis programs, part-whole inflation is at times referred to as "spuriousness".)

An example of the output corresponding to an external criterion analysis may be seen via a click here.

Related tidbits:

See the "Using an external criterion" section of the manual's Chapter 8.

An easy-to-read technical paper with more details about how Lertap 5 calculates item correlations is available here.

Also see "Using Lertap in a Test Validity Study", click here to branch out to it.

4.4.4 Item scores and correlations

A person gets a score for each item s/he answers. For example, for a cognitive item, the custom is to give one point for a right answer, and zero points otherwise. For affective items, each possible response usually has a certain number of points associated with it. "Strongly agree", for example, might equate to a score of 5 points, while "strongly disagree" might be made equal to just a single point.

In Lertap, a person gets a score on each item even when no answer is given. For cognitive items, a non-response usually equates to zero points, while for affective items Lertap will apply MDO, the "missing data option", to non-responses. MDO usually equals the mean of the possible scores for an item -- for example, if the possible range is 1.00, 2.00, 3.00, 4.00, 5.00, then MDO=3.00 (note).

When asked to do so, Lertap will get all the item scores together, and write them to a new worksheet named IStats. An IStats worksheet may have two sections -- the first section has rows whose cells contain the score each person earned on each item.

The cells may have background shading (highlighting) if they correspond to missing data (read more).

The second section is optional, depending on a setting in the System worksheet.

If row 22 of the System sheet has Yes as its "Present setting", then you'll find IStats will include rows of descriptive statistics for each item, such as the median, mean, and standard deviation. These rows will be followed by an interitem correlation matrix. Under "normal" conditions, a single matrix of Pearson product-moment correlation coefficients is made, with a row of average correlations included at the bottom of the matrix. Lertap uses standard Excel functions for all of these calculations. Average correlations are computed by using the n-1 non-diagonal entries in each column, where "n" is the number of items.

There are two settings in the System worksheet which affect the information found under this section of the IStats worksheet. One of them replaces the diagonal element of the correlation matrix with the "SMC", the squared multiple correlation. The SMC for an item is often used in factor analysis as an initial estimate of the item's common variance.

The standard setting for SMC calculations is "off"; to activate SMC output, change the setting seen in Row 21, Column 2 of the System worksheet. If this setting is changed, it takes effect immediately. (However, this doesn't mean that any correlations matrices you may have already made will immediately change; in fact, they won't -- you'll have to delete or rename the IStats worksheet, and then get Lertap to make a new IStats sheet.)

For additional information about Lertap's calculation of SMC values, just page ahead to the next topic.

You'll also find that Lertap will endeavor to find the eigenvalues of the correlation matrix. Eigenvalues are also known as "latent roots". Read more about them by paying a visit to this topic.

The IStats report will also include an estimate of coefficient omega; see this topic for more information.

Other settings in the System worksheet determine whether or not Lertap might add a matrix of tetrachoric correlations to the IStats output, and possibly make worksheets designed for export to other data analysis systems. There's more about these matters in the following topics.

A special macro, the "IStatsPruner" is available for dissecting an IStats summary. It creates an IScores worksheet with only the item scores, and an ICorrs worksheet with only the correlations.

For more about IStats sheets, please refer to the manual. You'll find an example in Chapter 7, and a fairly thorough discussion under the "Item scores matrix" section in

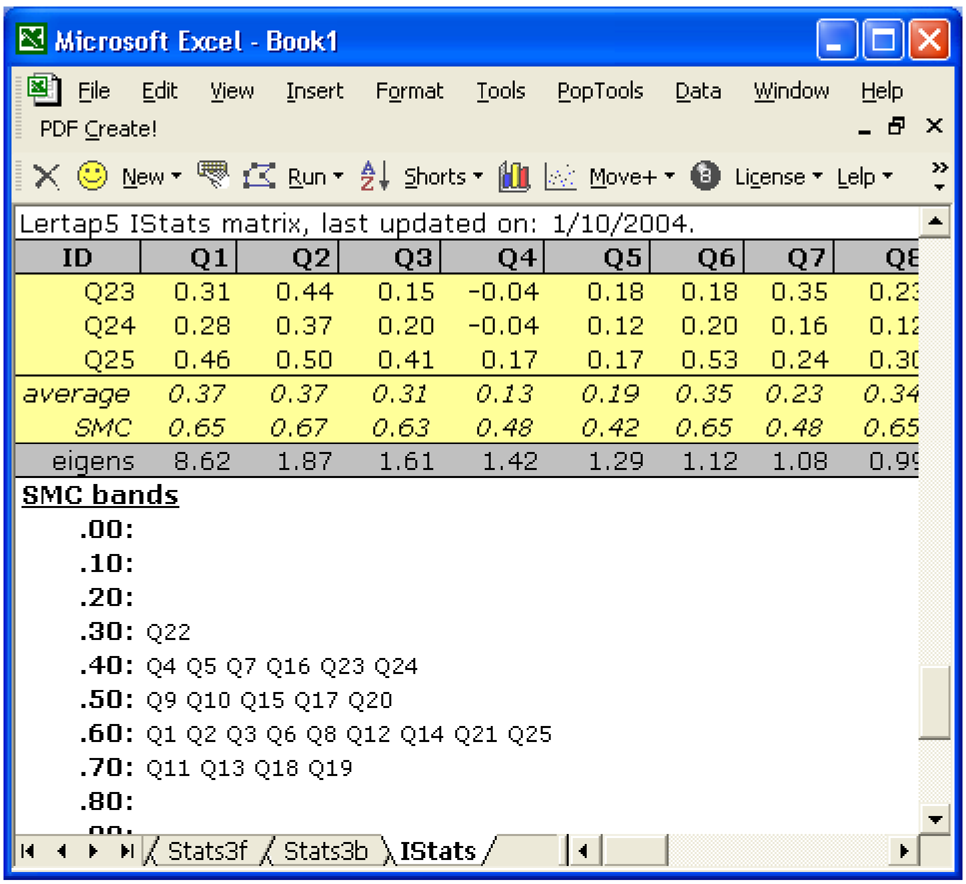
Chapter 10. Note, however, that the SMC, tetrachoric, export worksheets, and eigens options were added after the manual was printed.

4.4.4.1 SMCs

An item's SMC value, its squared multiple correlation, indicates the proportion of the item's variance which may be linked to, or predicted from, the other items in the subtest. As mentioned in the previous topic, the SMC is sometimes used as an estimate of the amount of variance any single item has in common with the other items.

For technical discussions on the SMC, and its calculation, see Pedhazur & Schmelkin (1991, pp. 414-417), Lord & Novick (1968, pp. 265-266), Hays (1973, pp. 705-708), and Glass & Stanley (1970, pp. 186-191), or search the Internet for "multiple correlation coefficient". (To get to the Lertap references page, begin with a click here.)

Lertap's IStats worksheet gives some prominence to SMC values by "banding" them. Help your little self to a squiz of the following sample:



As seen above, the IStats report now has a row with SMC values (just above the eigens row), followed by a display of SMC bands.

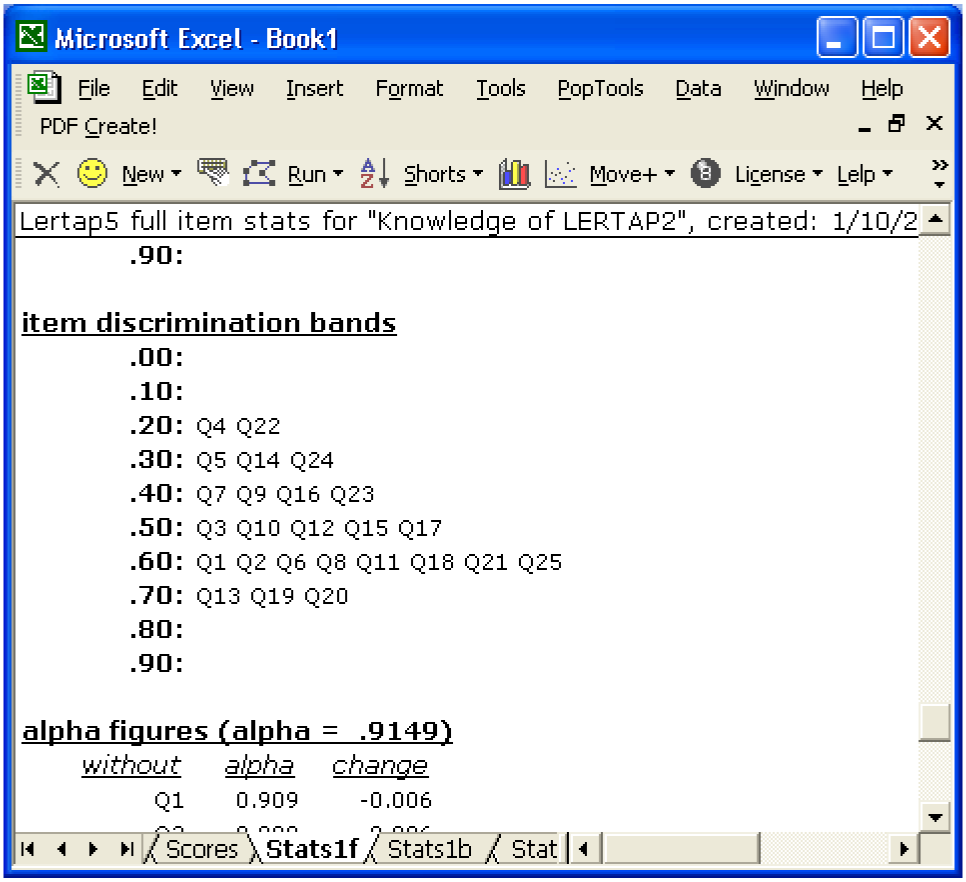
The bands give a quick idea of the spread of SMC values -- we see, for example, that nine subtest items had an SMC value equal to or greater than 0.60, but less than 0.70. Among these nine were Q1, Q2, Q3, Q6, and Q8 -- you can see the actual SMC values for these five items by looking at the row of SMC values showing above the bands.

In this case, we see that Q2's SMC was 0.67. We may interpret this as meaning that 67% of Q2's variance can be explained by the other items in the subtest. If we took the square root of Q2's SMC, we'd get 0.82, a value we may interpret as indicating the correlation between Q2 and the other items. (In somewhat more technical lingo, 0.82 is the value of the Pearson product-moment correlation coefficient between the scores people earned on Q2 and a specially-weighted linear composite score formed from the other items. The "special weights" are determined via a multiple linear regression analysis, as the references point out.)

Now, you'll remember that there is another Lertap report which indicates how an item correlates with the other items in the subtest.

There is?

Sure. Look at the item discrimination bands seen in the following screen snippet:



In Lertap, the standard index of item discrimination (or, for affective subtests, the item correlation) is the correlation between the item and a person's score derived by summing over all the other items in the subtest. (There is more about this in the manual.)

Note where Q2's discrimination index falls: in the 0.60 band. If we could scroll up the Stats1f report, or page over to the corresponding Stats1b report, we'd find Q2's correlation to be 0.66.

We have, then, two measures of Q2's correlation with the other subtest items, 0.82 and 0.66. Both figures represent the correlation between Q2 and a composite score formed by adding together the scores on the other subtest items; the SMC-based correlation will always be equal to or greater than the other correlation as it, the SMC-based value, is derived by using the special weights resulting from the multiple linear regression analysis underpinning the calculation of the SMC.

What do you have to do to get Lertap to produce SMC values? Nothing much, really, whenever you use the Run menu's option to "Output item scores matrix", the resultant IStats report will automatically include the row of SMC values, and the little table with SMC bands, as seen above.

What, then, is the SMC setting mentioned in the previous topic? It's a setting which determines whether or not the diagonal values of the IStats correlation matrix has 1's or SMCs. Why do some users want to have SMCs on the diagonal? Often because they're thinking of using the correlation matrix as input to a factor analysis program.

A discussion and example of using Lertap 5's output with other programs, such as SPSS, is available via this link.

4.4.4.2 Tetrachoric correlations

Tetrachoric correlation coefficients are computed when two conditions are met: (1) the tetrachorics option is set as "Yes" in the System worksheet, and (2), Lertap finds that the item scores are just zeros and ones.

These conditions are in fact easy to satisfy. The tetrachoric option's default setting in the System worksheet is No when Lertap is first installed, but this may quickly be changed to Yes. And cognitive test items are very often scored on just a right/wrong basis, with one point for a correct answer, zero points otherwise.

What are tetrachoric correlation coeffcients? They're estimates of what the correlation between two items would be if responses to the items had an underlying normal distribution, instead of the simple right/wrong dichotomy used to score the items. Some researchers and test developers are at times willing to assume underlying normal distributions, especially when they are interested in aspects of IRT modelling.

For more reading, use Lertap's references page, looking at Crocker and Algina, Lord, and/or Glass and Stanley. Or, search the Internet for definitions and discussions.

To compute the tetrachorics, Lertap uses an algorithm created by Brown (1977) (see References). Brown's algorithm calls for the use of two normal-curve functions: "AlNorm", and "PPND". Lertap uses two in-built Excel functions instead: NORMINV and NORMSDIST.

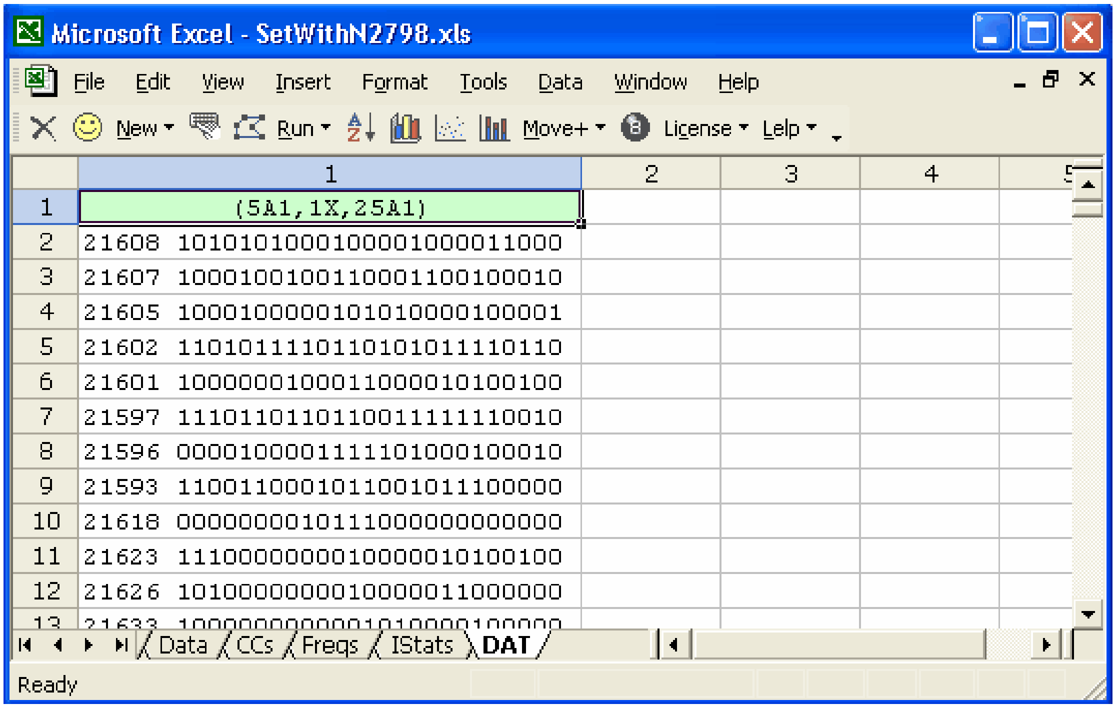
4.4.4.3 A DAT-like worksheet.

The Bilog and Bilog-MG computer programs like to have input formatted as what their authors call a "DAT" file, by which is meant a simple unformatted text file with fields of fixed length, suitable for reading by a FORTRAN Input statement. (Such files are often referred to as 'ASCII' files.)

Bilog-MG assumes that items have been scored on a dichotomous basis, that is, as either right or wrong, with a "1" used to signify right, and a "0" (zero) used for wrong.

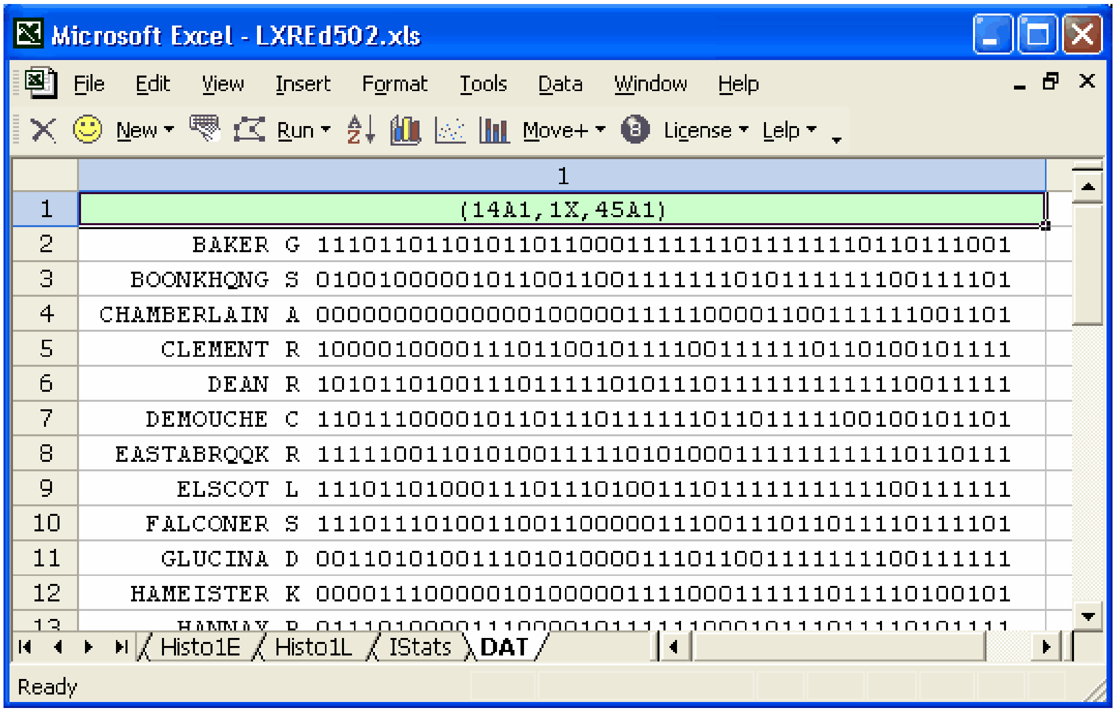
A Bilog-friendly DAT worksheet will be created by Lertap when the appropriate option is set in row 23 of the System worksheet.

Here's a screen shot from a Lertap DAT worksheet, one from a data set having 25 cognitive items, using a record ID field with a number between 15001 and 28000:



Every time Lertap makes a DAT worksheet, it inserts a Fortran format statement at the top. The line above says the data records start with a 5-column ID field, followed by a space, followed by 25 item scores.

Here's a snippet from another DAT worksheet. This one corresponds to a data set with 45 cognitive items, and an ID field with student names. The longest student name was 14 characters wide; Lertap has right-justified the names, using blanks on the left side whenever the name was shorter than 14 characters:



Now, a program such as Bilog, Bilog-MG, or Xcalibre will not read data from an Excel worksheet. The sheet has to be saved as a text file. How? How to save the DAT worksheet as a text file? Click here to find out, or page ahead to the topic titled "Creating a text file".

Once the DAT worksheet has been saved as a text file, the first line, the Fortran format statement, should be deleted -- Bilog won't like it. However, as long as the original DAT worksheet remains part of a Lertap/Excel workbook, the Fortran format line should not be deleted -- it's possible to add more data to the DAT worksheet, and the procedure which does this will cough, sputter, and die if the format line has gone walkabout.

What's this about adding more stuff to the DAT worksheet? Use the "Copy a Data column to the DAT worksheet" option (of course!), as found under the Move+ menu.

4.4.4.4 Xcalibre worksheets

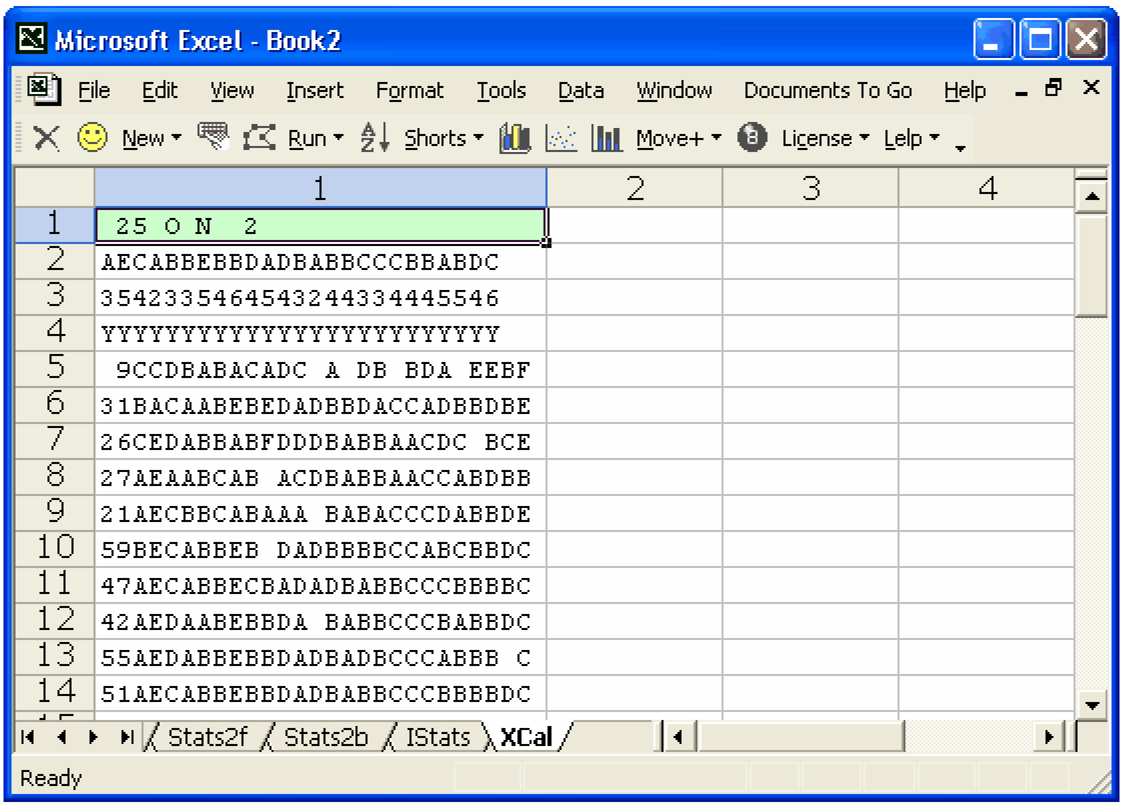
Xcalibre is the name of an IRT program created by Assessment Systems Corporation, ASC, in the United States.

In 2011, ASC released version 4 of Xcalibre. The XCal worksheet described in this topic was created by the Excel 2007 version of Lertap, known as version 5.8.2.1. The XCal worksheet will work with Xcalibre 4.1 or greater, providing you check the box "Data matrix includes an Xcalibre 1.1 Header" found in the Files tab used to set up a run with Xcalibre 4.1.

If you have the Excel 2010 version of Lertap (or later), it will create two files for use by Xcalibre 4.1. Please refer to the following topic for more information.

Lertap has an Xcalibre interface which works much like that for the Bilog-MG program mentioned in the previous topic. To activate the interface, go to the System worksheet in the Lertap5.xlsm file, and activate the setting in Row 24, Column 2 -- set it to "yes". Once "yes" is in the appropriate spot in the System sheet, Lertap will spin out a new worksheet every time the "Item scores and correlations" option is selected from the Run menu, providing that the subtest being processed is a cognitive one. The worksheet will be named "XCal".

Like Bilog-MG, Xcalibre wants to have its input arranged in a very specific format. Here's a screen shot of a Lertap XCal worksheet, one resulting from processing a 25-item cognitive subtest



The first four rows of the XCal worksheet have the control information wanted by the Xcalibre program. The 2nd row has the keyed-correct answers for the items; the 3rd indicates how many options were used by each item; and the fourth, a row of Ys, tells Xcalibre that all items are to be included in its analysis.

What about the 1st row? It's really the most detailed, having four fields of critical information for Xcalibre. The first field, characters 1-3 in the row, give the number of subtest items. The second field must appear in character position 5; this field tells

Xcalibre the code used in the data records to indicate an omitted item -- Lertap follows the Xcalibre convention of using the letter O for this code, but you may change it as wished. The N following the space after the O has to appear in character position 7; it indicates the code for items which have not been reached by a respondent. The final control field appears in character positions 9-10, giving the maximum number of characters of ID information for each test taker. In the example used here, only two ID characters were used.

The actual data records begin in Row 5 of the XCal sheet. Each record has its ID code as the first characters, followed by the item responses. Note that the screen shot above has blanks in some of the records -- a dinkum Xcalibre user would replace these with the code used to indicate omitted items -- in this example that would be the letter O.

Xcalibre users might want to note the caution given in the Xcalibre manual about processing "extremely large data sets", and avail themselves of Lertap's all-conquering "To Halve and Hold" option, an option which randomly splits a data set into halves.

Once Lertap has made its XCal worksheet, are you all set to run the Xcalibre program? No. The Xcalibre program will *not* read an Excel spreadsheet. The XCal worksheet has to be saved as a text file, and Xcalibre wants the text file to have an extension of "DAT". You now have a need to know how to save an Excel worksheet as a text file, and we've got some comments ready for your peepers to peep -- click here to jump to them, or simply page ahead until you get to the "Creating a text file" topic.

4.4.4.4.1 Xcalibre 4.1

The information in this topic has to do with getting Lertap to create the two special files for use with Xcalibre 4.1.

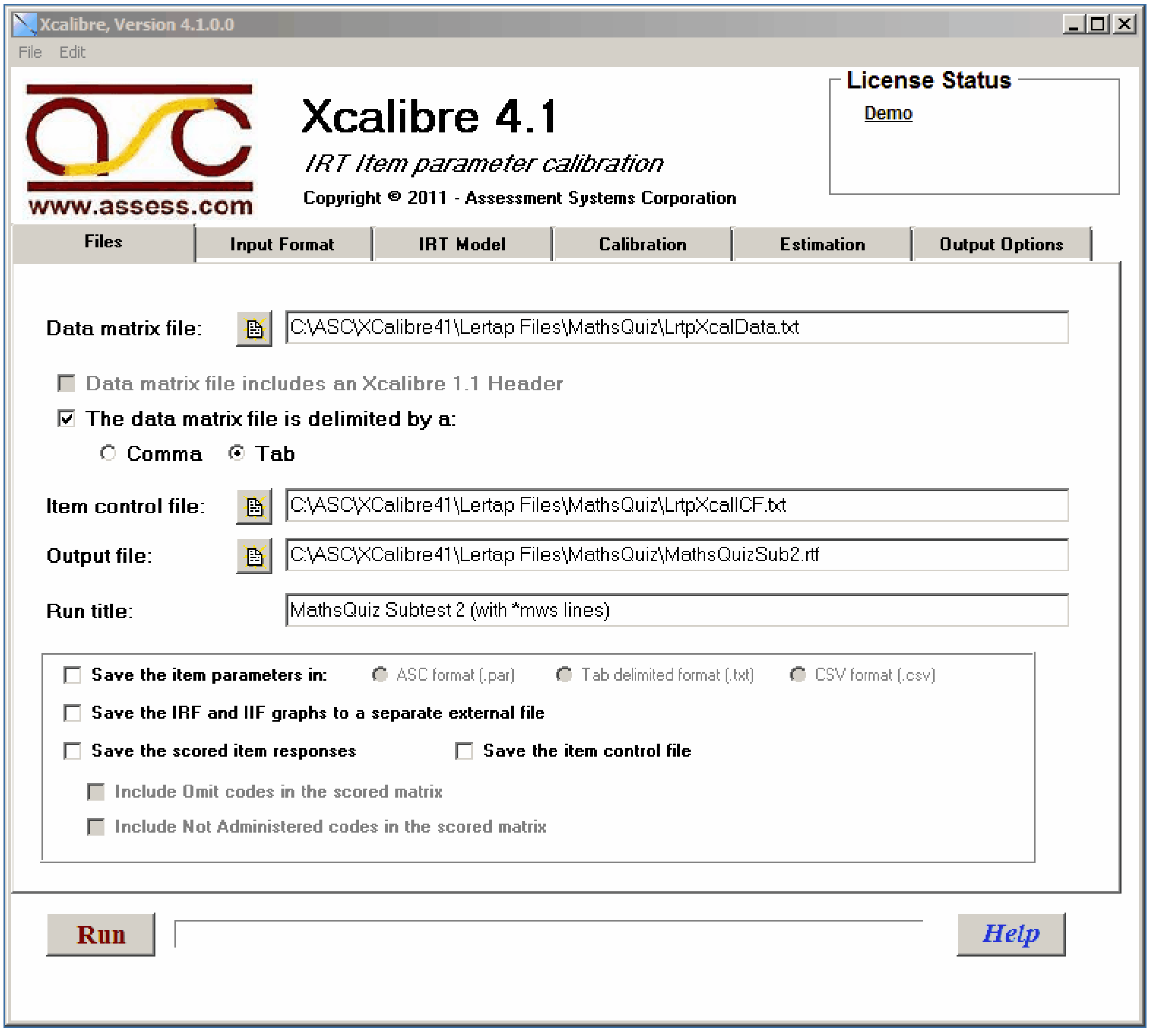
Lertap has an Xcalibre interface which works something like that for the Bilog-MG program mentioned in a previous topic. To activate the interface, go to the System worksheet in the Lertap5.xlsm file, and activate the setting in Row 24, Column 2 -- set it to "yes". Once "yes" is in the appropriate spot in the System sheet, Lertap will spin out two new worksheets, and two corresponding "txt" files, every time the output "Item scores and correlations" option is selected from the Run menu; the subtest selected for processing may be either cognitive or affective.

Xcalibre 4.1 requires a "Data matrix file" and an "Item control file". (This is not strictly true, it will also work with the old data matrix file used by the former version of Xcalibre. The Excel 2007 version of Lertap, 5.8.2.1, was the first able to create such a file -- it's called "XCal", please see the previous topic.)

The data matrix file created by Lertap 5.10 is called LrtpXcalData.txt. It is a "tab delimited text" file. Lertap will create this file in the same folder as the active workbook (that is, the one Lertap was working with when the "Item scores and correlations" option was selected).

The item control file created by Lertap 5.10 is called LrtpXcalICF.txt. It is also a "tab delimited text" file, and it will also be placed in the same folder as the active workbook.

Here's a picture of the File tab in Xcalibre 4.1, showing links to the two files from Lertap:



Lertap also creates worksheets with the same information as that found in the two text files. One of these new worksheets is called XCal41Data; the other is called XCal41ICF. These two worksheets are in no way critical to Lertap -- they may be deleted if they're of no use; deleting the worksheets does not delete the corresponding text files.

There's more information about Lertap 5 and Xcalibre in this topic.

4.4.4.5 RSAdata worksheet

RSA stands for "response similarity analysis" -- you'd be interested in RSA if you wanted to see if the responses of any two test takers were, as Wesolowsky (2000) would say, "excessively similar". In less diplomatic terms, RSA is used, by some, to examine the possible presence of cheating in an examination environment.

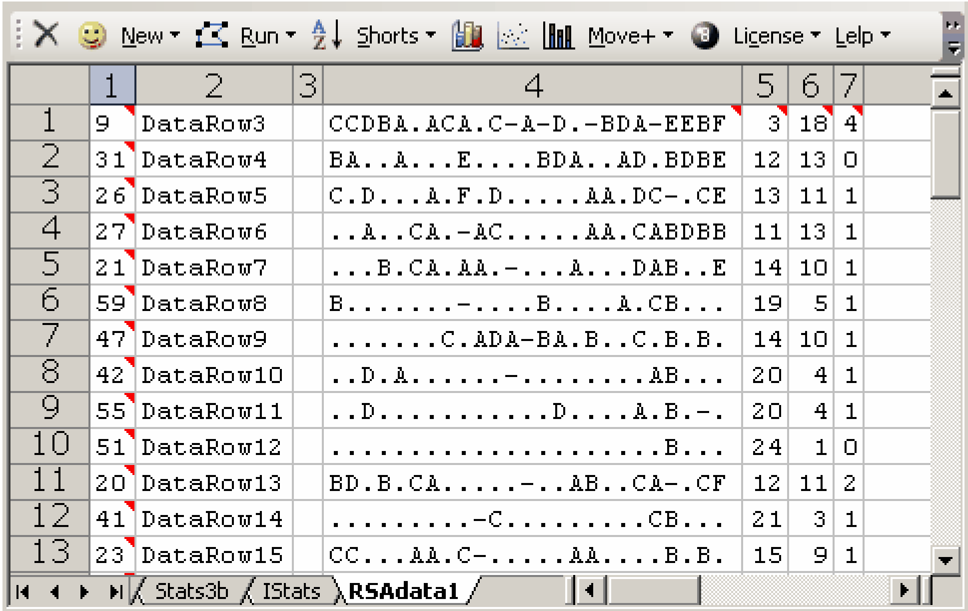
Lertap will produce a worksheet, "RSAdata1", and a special text file, "SCheckData1.DAT", whenever users take the "Item scores and correlations" option from the Run menu's More drop-down menu, and have set the RSA option to "yes" in the System worksheet. If production mode is on, and your data set includes more than one subtest, then there will be additional files: RSAdata2, SCheckData2.DAT, and so on (one pair of files for each subtest).

The screen snapshot below captures the System worksheet's RSA settings -- note the "yes" setting in row 25.

	1	2	3	4
1 2	These are Lertap5 system settings. Change them only if you understand them.	**Present setting:**	**Allowed settings:**	**Usual setting:**
25	Should an **RSA** worksheet be created?	yes	yes / no	no
26	Cutoff value for **Harpp-Hogan** statistic:	1.00	0 to 1	1.00
27	Minimum **EEIC** value:	6	0 to 20	6
28	Mark all records as **pickable** for RSA?	yes	yes / no	yes
29	Run in **production mode**?	no	yes / no	no
30	Include **histograms** in production mode?	no	yes / no	no
31	Include **response charts** in production mode?	no	yes / no	no
32	Include **items scores matrix** in production mode?	no	yes / no	no

Comments / Data / CCs / **System** / Syntax / OldCCs / Problems

The snapshot below indicates how the RSAdata worksheet will generally look:



	1	2	3	4	5	6	7
1	9	DataRow3		CCDBA.ACA.C-A-D.-BDA-EEBF	3	18	4
2	31	DataRow4		BA..A...E....BDA..AD.BDBE	12	13	0
3	26	DataRow5		C.D...A.F.D.....AA.DC-.CE	13	11	1
4	27	DataRow6		..A..CA.-AC.....AA.CABDBB	11	13	1
5	21	DataRow7		...B.CA.AA.-...A...DAB..E	14	10	1
6	59	DataRow8		B.......-....B....A.CB...	19	5	1
7	47	DataRow9		.......C.ADA-BA.B..C.B.B.	14	10	1
8	42	DataRow10		..D.A......-........AB...	20	4	1
9	55	DataRow11		..D..........D....A.B.-.	20	4	1
10	51	DataRow12		.....................B...	24	1	0
11	20	DataRow13		BD.B.CA.....-..AB..CA-.CF	12	11	2
12	41	DataRow14		.........-C.........CB...	21	3	1
13	23	DataRow15		CC...AA.C-.....AA....B.B.	15	9	1

The RSAdata worksheet is made to conform to a format used by Wesolowsky's "SCheck" program. (Refer to the references for the appropriate citation to Wesolowsky's work in this area, and to his web page for information about SCheck software: http://www.business.mcmaster.ca/msis/profs/wesolo/wesolo.htm.)

Each row in the RSAdata worksheet contains seven columns of information.

The first column corresponds to the Lertap ID in use -- Wesolowsky generally refers to this as the student ID number, but it doesn't have to be a true number -- it can be a name.

The second column is referred to as "name" in SCheck; Lertap inserts "DataRowX" instead, where X corresponds to the row number in the Data worksheet.

The third "can be initials", according to the SCheck.exe user guide -- Lertap leaves this column empty.

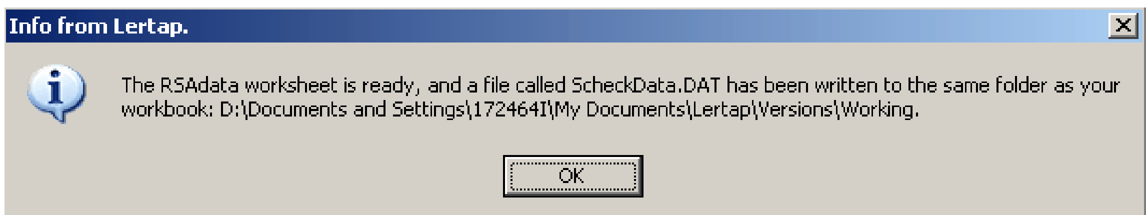
The fourth column's contents has a length equal to the number of items in the subtest, and indicates how each student responded to the items: a full stop (or "period") indicates that the student got the item right; a dash (or hyphen) indicates that the student did not answer the item (or had an answer not recognised by Lertap); and a letter or a digit indicates which wrong answer, which distractor, was selected.

In the snapshot seen above, the first student has an ID of 9; his or her complete data may be found in row 3 of the Data worksheet. The student left four items unanswered (there are four dashes), and got only three items correct (there are three full stops). The student selected distractor C on the first and second items, D on the third, B on the fourth, and so on.

Columns 5, 6, and 7 indicate the number of items a student answered correctly, the number answered incorrectly, and the number of questions having what Lertap refers to as an "other" response -- this is usually the same as the number of unanswered questions.

The red triangles (which may appear black on your screen or printout) seen in the snapshot indicate that the worksheet cells have comments. If you were to hover your mouse over one of the cells, you'd see that the comment is "Pickable for similarity analysis". RSAdata cells which have a red triangle will be included whenever the "RSA similarity analysis" option is taken from the Run menu. (This option has nothing to do with the SCheck program.)

Besides the RSAdata worksheet, Lertap will produce a companion "ASCII" file called SCheckData.DAT. If you've saved your workbook prior to selecting the "Output items scores matrix" option from the Run menu, Lertap will produce a message such as the following:



The purpose of this message is to remind you that you've now got a brand new DAT file to work with, and to tell you where to find it. If you haven't saved your workbook prior to selecting the "Output items scores matrix" option from the Run menu, this message may be a bit ambiguous, but the new ScheckData.DAT file will be on your computer, somewhere. (Please note that this message will not appear if you have Lertap set to run in "Production mode".)

Lertap's SCheckData.DAT file is ready for input to Wesolowsky's SCheck software. The contents of this file are very similar to those of the RSAdata worksheet, with commas used to separate information:

```
9,DataRow3, ,CCDBA.ACA.C-A-D.-BDA-EEBF
31,DataRow4, ,BA..A...E....BDA..AD.BDBE
26,DataRow5, ,C.D...A.F.D.....AA.DC-.CE
27,DataRow6, ,..A..CA.-AC.....AA.CABDBB
21,DataRow7, ,...B.CA.AA.-...A...DAB..E
```

(The information from columns 5, 6, and 7 of the RSAdata worksheet is not carried over to the SCheckData.DAT records.)

An option on Lertap's Run menu, "Response similarity analysis (RSA)" will get Lertap to use the RSAdata worksheet to feed its own response similarity investigator. To read more about this, simply click here.

To read more about the System worksheet, give a wee click here. A downloadable Excel workbook set up to check for cheaters may be found at the sample datasets website.

4.4.4.6 Eigenvalues

The eigenvalues, or "latent roots", or "characteristic roots", of a correlation matrix are sometimes used as a means of estimating the number of factors (or components) which may underpin a test, or a scale. There are often times when researchers would like to be able to say that their test is unidimensional, involving a single factor or construct. Some feel that a test may be said to be unidimensional if it can be shown that the largest eigenvalue underlying the test's correlation matrix is so dominant that it dwarfs the others. (See references and discussion below.)

Eigenvalues are computed if the System worksheet has "yes" in Row 22, Column 2.

Lertap's eigenvalue extraction uses computational routines produced by Leonardo Volpi and the Foxes Group in Italy, made available by the authors' kind permission. The Foxes Group's general matrix package, "Matrix.xla", is freely available at: http://digilander.libero.it/foxes/index.htm. Matrix.xla is a powerful, extensive set of matrix manipulation routines for use with Excel; it includes the ability to produce a complete principal factors / components analysis, with Varimax rotation, something Lertap users may wish to experiment with.

Here's a sample of Lertap's output with "eigens":

Lertap5 IStats matrix, last updated on: 31/10/2004.

ID	Q26	Q27	Q28	Q29	Q30	Q31	Q32	Q33	Q34	Q35
n	60	60	60	60	60	60	60	60	60	60
Min	1.00	2.00	2.00	2.00	1.00	2.00	3.00	2.00	1.00	1.00
Median	3.00	3.00	4.00	4.00	3.00	4.00	4.00	4.00	2.00	3.00
Mean	3.08	2.98	3.75	3.93	3.33	4.10	4.17	3.90	2.00	3.23
Max	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
s.d.	1.28	0.88	0.94	0.89	1.14	0.81	0.66	1.08	0.73	1.02
var.	1.64	0.78	0.89	0.80	1.29	0.66	0.44	1.16	0.53	1.05
Correlations										
Q26	1.00	0.57	-0.02	0.41	0.45	-0.01	0.24	0.74	-0.41	0.66
Q27	0.57	1.00	0.17	0.23	0.34	0.14	0.06	0.45	-0.34	0.34
Q28	-0.02	0.17	1.00	-0.14	-0.12	-0.05	-0.07	-0.19	-0.02	-0.20
Q29	0.41	0.23	-0.14	1.00	0.37	0.01	0.27	0.43	-0.31	0.36
Q30	0.45	0.34	-0.12	0.37	1.00	-0.04	0.04	0.55	-0.36	0.51
Q31	-0.01	0.14	-0.05	0.01	-0.04	1.00	-0.03	-0.07	-0.14	-0.05
Q32	0.24	0.06	-0.07	0.27	0.04	-0.03	1.00	0.26	-0.17	0.26
Q33	0.74	0.45	-0.19	0.43	0.55	-0.07	0.26	1.00	-0.51	0.55
Q34	-0.41	-0.34	-0.02	-0.31	-0.36	-0.14	-0.17	-0.51	1.00	-0.47
Q35	0.66	0.34	-0.20	0.36	0.51	-0.05	0.26	0.55	-0.47	1.00
average	*0.29*	*0.22*	*-0.07*	*0.18*	*0.19*	*-0.03*	*0.10*	*0.25*	*-0.30*	*0.22*
SMC	*0.72*	*0.42*	*0.22*	*0.27*	*0.41*	*0.11*	*0.16*	*0.69*	*0.39*	*0.57*
eigens	3.83	1.25	1.06	0.98	0.69	0.68	0.51	0.46	0.37	0.16
percent	38.3%	12.5%	10.6%	9.8%	6.9%	6.8%	5.1%	4.6%	3.7%	1.6%
p-comp1	0.85	0.61	-0.14	0.60	0.68	0.02	0.34	0.85	-0.65	0.79

In this example, the 10-item "Comfort" affective scale seen in the Lertap Quiz data set, the largest eigenvalue was 3.83, the smallest 0.16. In a well-conditioned correlation matrix with 1's (ones) on the diagonal, the sum of the eigenvalues will equal n, the number of test items (assuming the correlations are Pearson product-moments, not tetrachorics).

The row with the actual eigenvalues is followed by the "percent" row seen above. The percent figures appear whenever the correlation matrix has 1's on its diagonal; when the SMC setting is on, and SMCs are found on the diagonal, two changes are made to the table: the percent figures are not created, and the correlations found in the p-comp1 row are replaced with correlations between the item and the first principal *factor*, with the row's label then changing to p-fact1.

What do the percent values mean? Well, first note that there are ten items in this example, Q26 through Q35. There are also ten eigenvalues. As noted above, the sum of the eigenvalues equals the number of items: 10 in this example. The percent value for the first eigenvalue is 100(3.83/10), or 38.3%.

Each eigenvalue corresponds to what's called a "principal component". If we could look at the multivariate scatterplot of the ten items, and if each item had a distribution meeting the requirements of the normal distribution, the scatterplot would have the form of an n-dimensional ellipsoid, where n is the number of items (10 in this case). If the items are uncorrelated, the ellipsoid is an n-dimensional sphere. If, on the other hand, the items are correlated, the sphere stretches out to an ellipsoid.

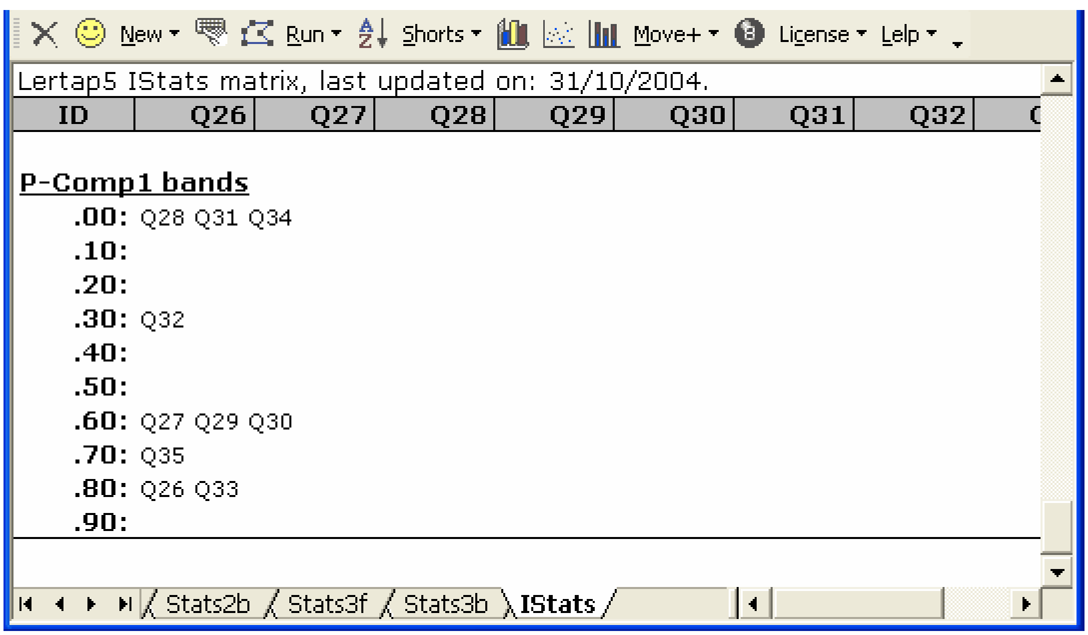
After the percent row comes the "p-comp1" row, giving the correlation of each of the items with the first principal component -- the values found in this row are also sometimes called the "loadings" of the items on the first principal component.

The first principal component corresponds to the ellipsoid's major axis, to its longest axis. Each eigenvalue represents the relative length of one of the ellipsoid's axes. Each of these axes is said to represent, or correspond to, a principal component.

Think for a moment of the case when n=3. If the three items are normally distributed and uncorrelated, their scatterplot will have the form of a soccer ball, a perfect sphere. As the three items begin to correlate, the soccer ball changes shape, morphing into an American football, and then, as the correlation among the items increases, into a cigar shape. The shape of the scatterplot is highly related to the relative sizes of the eigenvalues; if the eigenvalues are all equal, the shape is a sphere. If the first eigenvalue is much greater than the others, the shape is a cigar, and in such a case the multivariate scatterplot is said to have, essentially, one principal component, or dimension.

In the 10-item example above, the first principal component is said to account for 38.3% of the total variance (or volume) found in the multivariate scatterplot. As the size of the first component comes to dwarf the others, some people say there appears to be but one dimension underlying the items, which, in turn, often leads people to say that the items are "measuring the same thing".

Lertap will also "plot" the item-component correlations (or loadings) in bands. It takes the values found in the p-comps1 row, and makes a little table, such as the one below:



The P-Comp1 bands indicate that there's a group of six items, Q26, Q27, Q29, Q30, Q33, and Q35 with high correlations on the first principal component. If we were to create a new subtest using just these items, chances are very good we'd end up with a coefficient alpha value much higher than that obtained for all ten original items.

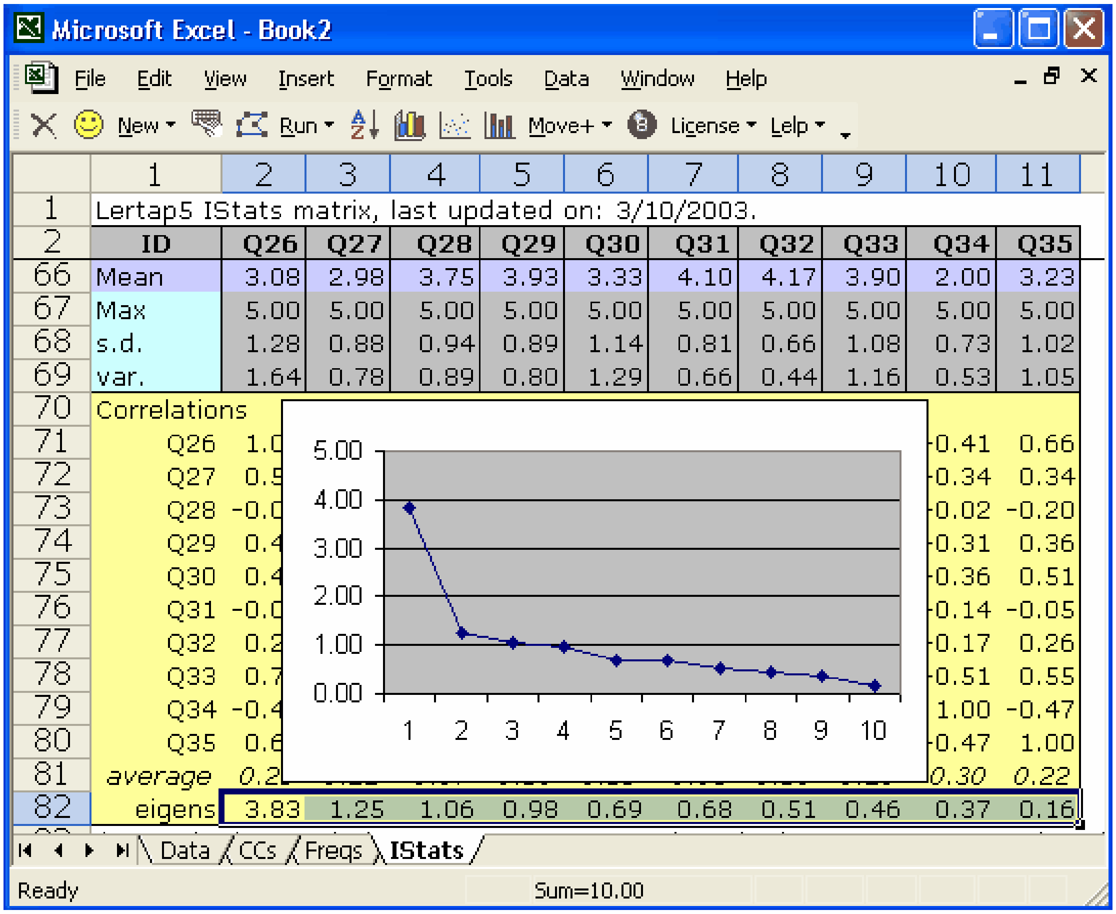
And, speaking of alpha values, did you happen to notice that one of the eigenvalues seen above, the first one, has a little black triangle next to it? (This triangle is really red, not black, but for some reason when we took our snapshot of the original screen the colour changed.)

When you have your own IStats screen showing, find one of these triangles and let your mouse hover above it. Lertap will display the alpha value for the corresponding principal component; in this case the value turns out to be 0.821 -- it can be shown that this value, 0.821, is the maximum possible value which coefficient alpha could assume for any linear combination of the items comprising the subtest. (Please refer to the technical paper cited below for more information, and also please note that these small triangles will appear only when the corresponding alpha value is equal to or greater than 0.60.)

The Scree Test / Plot

When we think about the first eigenvalue possibly "dwarfing" the others, we might well long for a picture of some type. The scree test was invented by Cattell way back in 1966 to meet these longings. Cattell suggested we graph the eigenvalues from highest to lowest to see if the first eigenvalue, or the first few eigenvalues, dwarf the others. His suggestion remains popular to this day.

We can graph our 10 eigenvalues using a couple of methods. The plot shown below was obtained by selecting the eigenvalues, and then using Excel's Insert / Chart (Line) options. An easier way to accomplish much the same thing is to use an option from the Lertap tab on the Excel ribbon: there is a "Line" option in the "Basic options" icon group.



The so-called scree test for the number of factors involves nothing more than eye-balling a line graph such as the one above, and deciding where the scree begins. In case you've forgotten, the scree is all the loose rocks at the base of the cliff your friends want to climb, those pesky fallen chunks where your boot will slip in and get stuck, twisting your ankle, granting access to a face-saving retreat to the beer tent in case you were really too chicken to climb the cliff to begin with.

Does the first eigenvalue dwarf the others? Does our scree begin with the 2nd eigenvalue, or the 5th? This question will remain unanswered here; many times the start of the scree is much easier to detect. For references on the scree test, see Catell (1966), Pedhazur and Schmelkin (1991), or search the Internet.

Note that eigenvalues can go negative. This is likely, for example, when SMCs are used on the diagonal of the correlation matrix, when one of the items has no variance, or (especially) when tetrachoric correlations are used. Also note that it is possible for the eigenvalue extraction method used by Lertap to fail; the method is an iterative one which concludes when the iteration process appears to converge. Under some circumstances convergence will not occur -- eigenvalues will not be returned in such cases (but it may be worthwhile to try again, that is, to return to the Run menu, click on the "More" option, and again request "Item scores and correlations").

The computation of eigenvalues can be a timely, labour-intensive task for your computer. If you will not be making use of eigenvalues, and have no desire to become an avid scree plotter, then you'll want to turn off the eigenvalue option in the System worksheet (the option's setting is found in Row 22, Column 2 -- set it to "no").

More timely comments may be found by paging ahead to the time trials topic.

Related tidbit:

For more about these topics, see "Some observations on the scree plot, and on coefficient alpha", a 16-page document with lots of little tables and some wonderful screes, available via the Internet: click here if you're connected.

4.4.4.7 Omega

Omega is a reliability coefficient said to be superior to coefficient alpha as it makes fewer assumptions, and can be expected to result in a more accurate estimate of a test's or scale's reliability. Lertap5 uses the "closed-form" method to derive an estimate of omega. This working paper has related reference material and citations.

When the "Item scores and correlations" option is used to generate an "IStats" report, statistics related to coefficient omega will be found towards the bottom of the report.

Covariances

Inter-item covariance values will be displayed in matrix form, with item variances on the diagonal. If any of the covariance values are negative they will be highlighted in yellow. These are unwanted -- all items will hopefully "go well and hang together" (correlate positively with each other). Reliability coefficients such as alpha and omega will be highest when all covariance values are positive and, ideally, at least 0.30 in magnitude.

Closed-form lamda estimates

These indicate the degree to which each item contributes to, "loads on", or correlates with the single "general factor" postulated by the underlying model, estimated by using the closed-form method. Ideally all lamdas will be above 0.50 in magnitude.

Score variance is the variance of the total test/scale scores as found in the Scores report.

The omega estimate is calculated by summing the lamda estimates, squaring the sum, and then dividing the result by the score variance.

In theory the maximum value the omega estimate can obtain is 1.00 (indicating perfect reliability). However, if the covariance matrix has numerous negative entries, it is possible to find the omega estimate exceeding its theoretical maximum. When this happens, Lertap will raise a warning message and indicate that the omega estimate should be ignored. This does not mean that omega cannot be computed - rather, the indication is that the closed-form method itself has failed and users should make use of another of the methods mentioned in the working paper to obtain an omega estimate.

4.4.4.8 Factor analysis

As mentioned in a previous topic, when the SMC setting is on, Lertap will output a row with correlations between each item and the first principal factor. Witness:

Q34	-0.41	-0.34	-0.02	-0.31	-0.36	-0.14	-0.17	-0.51	0.39	-0.47
Q35	0.66	0.34	-0.20	0.36	0.51	-0.05	0.26	0.55	-0.47	0.57
average	*0.29*	*0.22*	*-0.07*	*0.18*	*0.19*	*-0.03*	*0.10*	*0.25*	*-0.30*	*0.22*
SMC	*0.72*	*0.42*	*0.22*	*0.27*	*0.41*	*0.11*	*0.16*	*0.69*	*0.39*	*0.57*
eigens	3.37	0.55	-0.29	0.25	0.22	-0.21	-0.14	0.11	0.08	0.02
p-fact1	0.39	0.26	-0.06	0.24	0.29	0.01	0.13	0.39	-0.27	0.34

Lertap's first principal factor is exceedingly simple -- its initial estimate of an item's communality is the same as its final estimate: the item's SMC (no iterations are undertaken to improve on the SMC).

Some users may find Lertap's p-fact1 row to be a useful tool for indicating relative item loadings on the first factor, but those interested in a more complete factor analysis, or, for that matter, a complete principal components analysis, will want to step out to a program such as SPSS or SAS.

How to move beyond Lertap, to get Lertap's item scores into, for example, SPSS?

Not too difficult. Have a look here.

4.4.4.9 Creating a text file.

A text file, sometimes referred to as an ASCII file, is a file devoid of special formatting characters. Examples of text files are those which may be processed on a Windows-based computer with the Notepad or WordPad programs, or on a Macintosh with a program such as TextEdit. On Windows computers, text files often have an extension of "TXT".

In the data processing world, data files are often text files (ASCII files), and they frequently have an extension of "DAT". Many of the programs from SSI, Scientific Software International (www.ssicentral.com), and from ASC, Assessment Systems Corporation (www.assess.com), enjoy receiving their input from text files saved with an extension of DAT.

As you now well know, Lertap's repertoire of data processing capabilities includes a provision for creating Excel worksheets formatted so as to be compatible with some of the ASC and SSI programs. For example, Lertap's XCal worksheet is made for use with

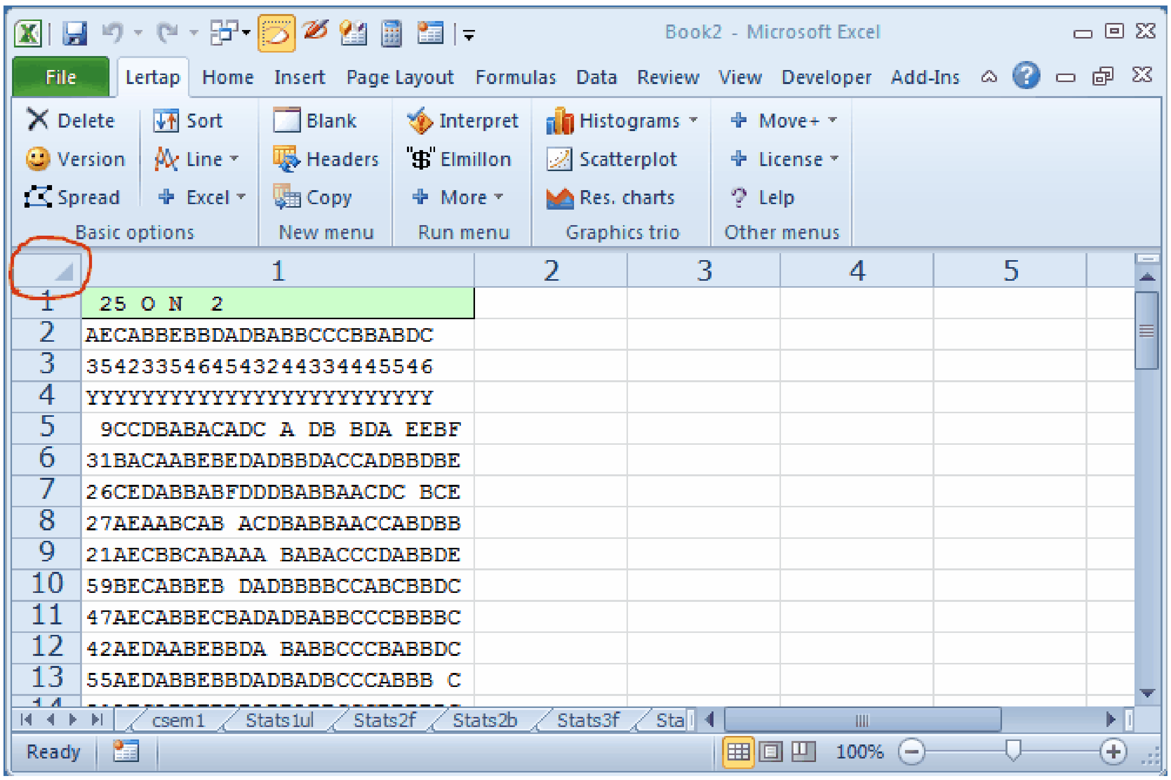
ASC's XCALIBRE program, while Lertap's DAT worksheet is designed to be friendly to SSI's Bilog-MG program.

But the ASC and SSI programs cannot (yet) read from Excel worksheets. We require a way to save Lertap's XCal and DAT sheets so that they're text files ready for input to the other programs. Can do?

Sure. In fact, there's more than one way. First, make sure that the DAT or XCal worksheet is the active one, the one currently in view.

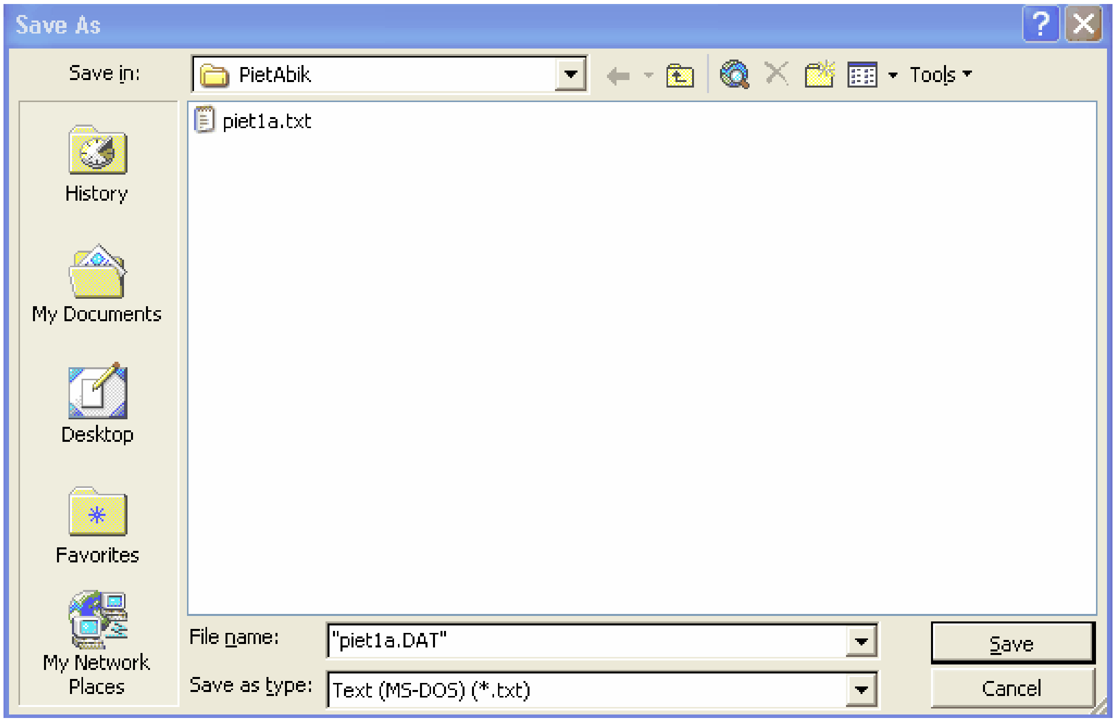
About the easiest way to prepare an ASCII file is to select, copy, and paste. Select all of the rows in the DAT or XCal worksheet, copy, open the Notepad, WordPad, or TextEdit program, and paste. Simple.

How to select all of the rows in an Excel worksheet? It's real easy, but not obvious if you haven't done it before. Just click at the top-left of the worksheet, where the red circle is seen in the picture below.

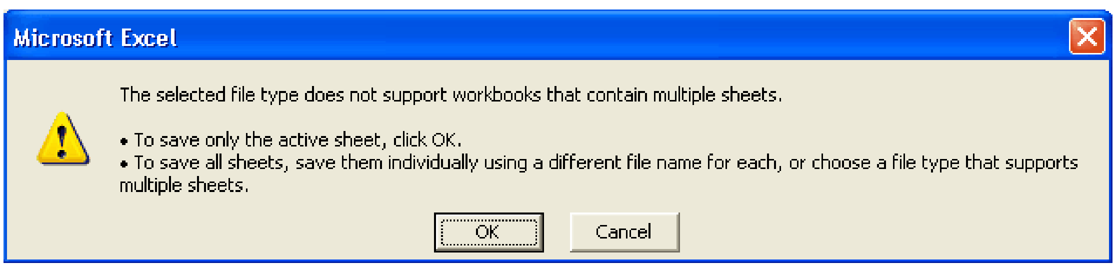


Another way, not as easy: use Excel's / Save as TXT (MS-DOS) option. This will create a simple text file with an extension of TXT; this file may then be renamed so as to have an extension of DAT.

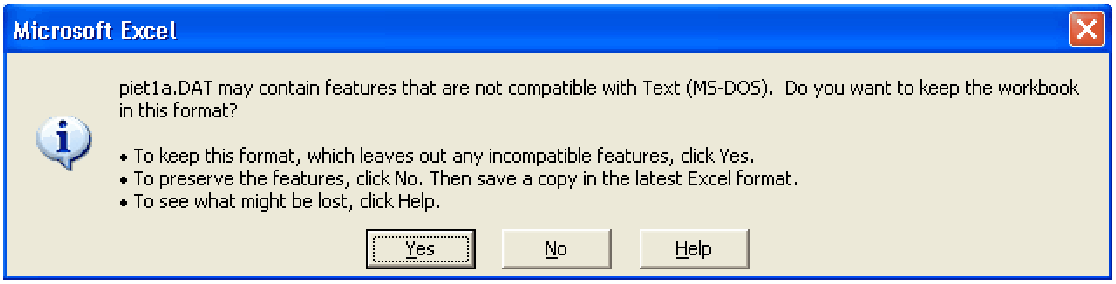
Then, here's still another way to get the job done: in the process of saving the file as TXT (MS-DOS), using quotation marks around the file name will allow it to be saved directly as a DAT file -- for example:



When saving TXT or DAT files in this manner, Excel is likely to send a message such as this 'un:

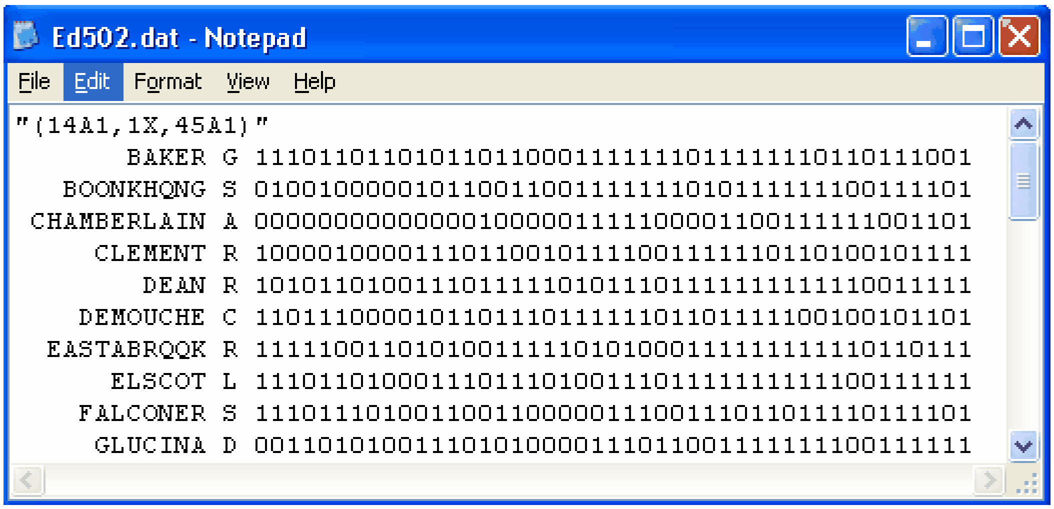


Click OK, and be prepared for Excel to say something like this:



This is also okay -- click Yes.

To see if the file you've saved looks okay, get out Notepad, or WordPad, or, for Mac users, TextEdit, and use the File menu to Open your work of art. The screen snippet below shows an example, in this case a Lertap DAT worksheet saved as Ed502.dat, and viewed on a Windows computer with the Notepad program:

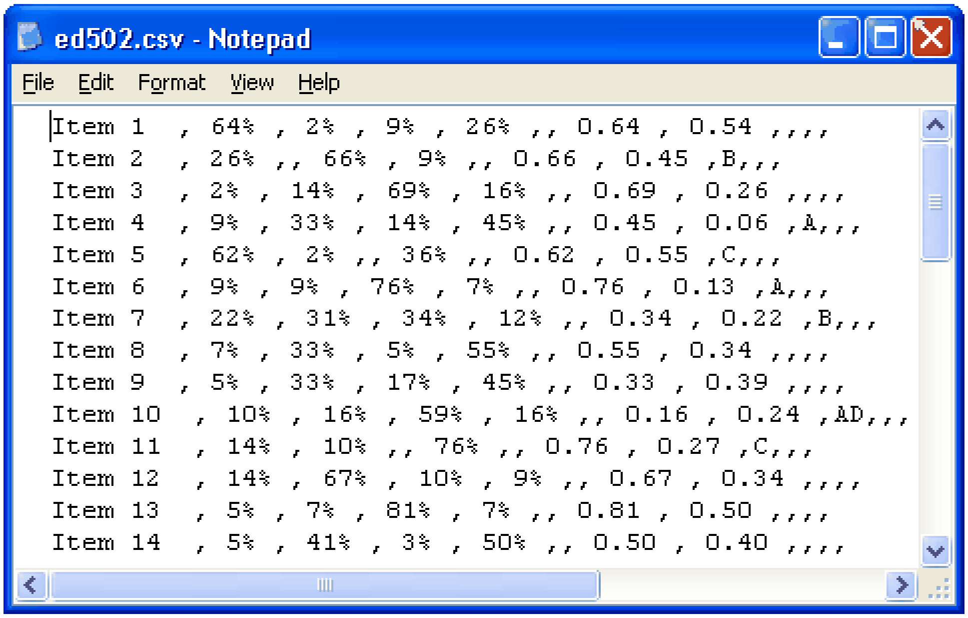


4.4.4.10 Creating a csv file.

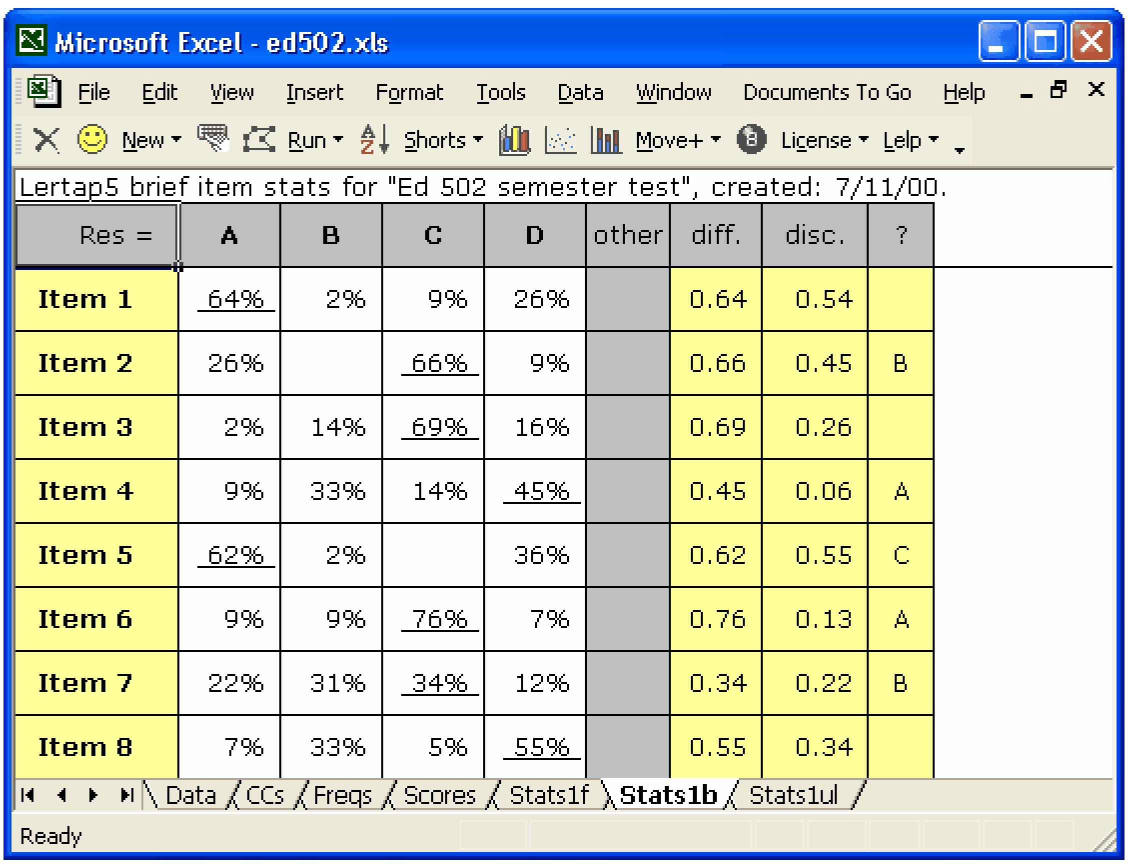
A "csv" file is a text file with a certain number of "fields", with each field representing a value of some sort or another.

CSV means comma-separated values. The records (or lines) in a csv file have a series of values (or fields), with commas used to separate them.

Here's an example (two commas with nothing between them corresponds to an empty field):



The csv file above came from a Lertap Stats1b worksheet which looked like this:



Res =	A	B	C	D	other	diff.	disc.	?
Item 1	64%	2%	9%	26%		0.64	0.54	
Item 2	26%		66%	9%		0.66	0.45	B
Item 3	2%	14%	69%	16%		0.69	0.26	
Item 4	9%	33%	14%	45%		0.45	0.06	A
Item 5	62%	2%		36%		0.62	0.55	C
Item 6	9%	9%	76%	7%		0.76	0.13	A
Item 7	22%	31%	34%	12%		0.34	0.22	B
Item 8	7%	33%	5%	55%		0.55	0.34	

How did we get from the Stats1b worksheet to the csv file? We followed a procedure almost identical to that described in the previous topic, "Creating a text file". However, instead of asking Excel to Save as TXT (MS-DOS), we directed it to Save as CSV (Comma delimited) (*.csv).

Lertap users may have a variety of needs which prompt them to save worksheets as csv files. Among these would be a desire to use Lertap's statistics with an item banking and test development system such as FastTEST from ASC, Assessment Systems Corporation (www.assess.com). The latest versions of FastTEST have an Import Wizard which makes it a straightforward matter to pick up values in a csv file. In the example above, we'd tell FastTEST to pick up csv field #7 as the "P-Value", and csv field #8 as the "ItmTtlCorr".

The number of columns seen in a Stats1b report depends on the number of response options, or alternatives, used by a subtest's items. At times there will be too many columns, too many fields when the worksheet is saved as a csv file, for easy use with FastTEST. In this case you'll want to delete some of Stats1b's columns before making the move to save as a csv file.

Is it difficult to delete Stats1b columns? Nope; it's real easy. Use the toolbar's Shorts menu to "Turn row and column headings on/off". Then get out your mouse, and right-click on, say, column 2. Left-click on Delete, and guess what? Bingo! -- the column is gone.

Now, you know how we've been saying there may be too many Stats1b fields, and how you might want to delete some if you're making a csv move to FastTEST? Well, come to think of it, you might want to insert a new column in the Stats1b worksheet before saving it as a csv file. Yes. FastTEST assigns and carries a UniqueID field for each item. Your work might be a bit easier if you inserted a new column in the Stats1b worksheet, and typed each item's FastTEST UniqueID into it before saving as a csv file. This may speed up the task of importing the item stats.

Is it difficult to insert a new column in the Stats1b worksheet? Yep, it's real tough, about as hard as having to quaff a few ice-cold Emu Exports on a hot summer's day. (Be sure to use the Shorts menu to turn column headings on first.)

The item discrimination value seen in Lertap's Stats1b report, "disc.", is a point-biserial correlation coefficient corrected for part-whole inflation. (The manual discusses Lertap's statistics in some detail.) It is possible to get the Stats1b report to include the biserial equivalent, something which is done by turning on Lertap's "Experimental Features" option. Please refer to the following URL for a discussion of these features:

http://www.lertap5.com/Documentation/ExperimentalFeatures.htm

4.4.4.11 Time trials

--> The most up-to-date data on time trials is maintained here.

The information below is presented only for historical reasons. These days (year 2018), a laptop running a 7th or 8th generation i5 or i7 processor would take less than a minute to process the examples tabled below.

The data minuted in the secs-y table below were obtained in October 2003, on a Pentium 4 running at 2 GHz. N is the number of data records in the Data worksheet; Nits is the number of items to be scored; Item scores indicates the amount of time to produce IStats' matrix of item scores; No eigens indicates the total time to completion, that is, time to create the item scores, the matrix of Pearson correlations, the matrix of tetrachoric correlations, and the DAT worksheet; With eigens indicates how much longer the job took when eigenvalues were also computed.

N	Nits	Item scores	No eigens	With eigens
450	40	13 secs.	27 secs.	51 secs.
1,400	25	24 secs.	33 secs.	42 secs.
2,800	25	48 secs.	1:02 mins.	1:11 mins.
5,600	25	1:45 mins.	2:12 mins.	2:21 mins.
11,200	25	4:04 mins.	4:51 mins.	5:08 mins.

Note: we received a trial data set from a Lertap user with N=267 and Nits=150. It took Lertap a total of 14 minutes to produce its IStats report for this data set (!) -- of this, 11 minutes were required to extract eigenvalues from the correlation matrix. This test used the same computer mentioned above, a Pentium 4 running at 2 GHz. At the end of 2004, this Pentium would be considered as being quite dated (it was new in early 2002); if you're wanting IStats reports for data sets with large Nits, hopefully you'll have a more powerful computer to work on.

Note July 2010: fresh tests with another data set. This one had N=4,700 and Nits=77. In this case, it took Lertap 3 minutes to produce its IStats report, a figure which includes the 43 seconds required to extract eigenvalues from the correlation matrix. We used Excel 2010 for this, running on a Macintosh MacBook Pro (2.53 GHz; 4 GB RAM) with Parallels and Windows 7 (if you have a more conventional Windows computer, not a Mac, you'd likely observe better figures).

Note December 2012: for a comparison of Excel 2010 and Excel 2013 running times, you might have a look at this document with a few selected time trials.

4.4.5 Rasch analysis

Lertap5 employs JML, joint maximum likelihood estimation, to derive Rasch estimates of item difficulty and student ability.

A comprehensive example exemplifies how JML works by using an Excel workbook. See Wu et al (p.123) for a description of JML in the context of Rasch scaling.

Note: There is another way to undertake a Rasch analysis with Lertap5. Pay a visit to the topic having to do with the "IRT with TAM" package, and the Omega1 special macro. An example of a report made by using TAM is here.

Tests must involve cognitive items scored on a right/wrong basis. That is to say, the program is exclusively for use with "dichotomous" test items.

Examples of the program's output are shown below. They're based on test results collected from students in Japan as part of the FIMS project.

FIMS results for Japan sample

Item	CTT Diff	Rasch Diff	Error	Infit	Outfit	Rasch item difficulty bands	
Q1	0.80	-1.07	0.06	0.85	0.83	4.00	
Q2	0.88	-1.77	0.06	**0.67**	**0.48**	3.50	
Q3	0.92	-2.17	0.07	**0.69**	**0.65**	3.00	
Q4	0.72	-0.56	0.05	0.83	0.79	2.50	
Q5	0.33	1.37	0.05	0.73	**0.65**	2.00	
Q6	0.90	-1.97	0.07	0.72	**0.60**	1.50	Q5 Q7 Q12
Q7	0.34	1.31	0.05	1.03	**1.12**	1.00	Q8 Q9 Q11 Q13
Q8	0.36	1.20	0.05	0.97	0.99	0.50	
Q9	0.40	0.99	0.05	0.83	0.81	0.00	Q10
Q10	0.65	-0.20	0.05	0.86	0.82	-0.50	Q4
Q11	0.38	1.12	0.05	0.75	**0.69**	-1.00	Q1 Q14
Q12	0.29	1.56	0.05	**1.23**	**1.38**	-1.50	
Q13	0.39	1.03	0.05	0.84	0.79	-2.00	Q2 Q3 Q6
Q14	0.76	-0.83	0.05	0.74	**0.63**	-2.50	
Average:	**0.58**	**0.00**	**0.06**	**0.84**	**0.80**	-3.00	
Median:	**0.53**	**0.40**	**0.05**	**0.83**	**0.79**	-3.50	
s.d.:	**0.23**	**1.33**	**0.01**	**0.15**	**0.22**	-4.00	

Number of student scores in each of the logit bands		
4.00	0	0%
3.50	0	0%
3.00	112	6%
2.50	0	0%
2.00	164	8%
1.50	180	9%
1.00	423	21%
0.50	263	13%
0.00	282	14%
-0.50	393	20%
-1.00	100	5%
-1.50	55	3%
-2.00	23	1%
-2.50	0	0%
-3.00	12	1%
-3.50	0	0%
-4.00	0	0%

Scores Summary	
n	2,007
Average	0.52
Median	0.42
s.d.	1.18

Note that the program has centered the item difficulty estimates, "Rasch Diff", at zero (the average difficulty is 0.00 as shown above). This is in keeping with the Rasch examples used in the Bond & Fox texts: "*Applying the Rasch Model*" (references are here).

Red highlighting is used to draw attention to item INFIT and OUTFIT values which may be less than or greater than desired.

values less than 0.7 will be in red no matter what the sample size (number of students)
values greater than 1.1 will be in red when sample size is 1000 and above
values greater than 1.2 will be in red when sample size is between 500 and 1000
values greater than 1.3 will be in red when there are less than 500 students

Another small section of the output is shown below. (The tables have come from two distinct workbooks in this example. The JPN results in the table below are not from the same JPN sample used above.)

Logit ranges (JPN)

	Lowest	Highest
Rasch Diffs:	-2.46	1.76
Student scores:	-3.47	3.30

Logit ranges (AUS)

	Lowest	Highest
Rasch Diffs:	-2.29	2.55
Student scores:	-3.45	3.58

When all iterations have been completed, the student (person) INFIT and OUTFIT values produced by the program will be found, along with Rasch scores, in the Scores worksheet. When these values are less than 0.75, or greater than 1.3, they are highlighted in red (no matter what the sample size might be). The Sort option may be used to have a closer look at the range of these values and scores.

Some of the examples in Bond & Fox are based on the "BLOT" test. The next page compares the program's output with that published on p.56 of Bond & Fox (2007).

4.4.5.1 BLOT statistics

From p.56 of Bond & Fox (2007):

TABLE 4.1
BLOT Item Difficulty Estimates With Associated Error Estimates for Each Item

Item	Difficulty Estimate	Error Estimate	Infit Mean Square	Outfit Mean Square	Infit t	Outfit t
1	−0.77	0.26	0.98	0.69	0.0	−0.8
2	−0.70	0.26	1.01	0.75	0.1	−0.6
3	0.74	0.2	0.98	0.9	−0.2	−0.5
4	0.00	0.22	1.00	0.88	0.0	−0.4
5	−0.98	0.28	0.98	0.76	−0.1	−0.5
6	−2.42	0.47	1.06	0.83	0.3	0.1
7	−0.64	0.25	0.97	0.65	−0.1	−1.0
8	0.85	0.19	0.91	1.00	−1.1	0.1
9	0.18	0.21	1.07	0.97	0.7	0.0
10	−0.19	0.23	0.92	0.68	−0.7	−1.1
11	0.18	0.21	1.02	0.96	0.2	−0.1
12	−1.76	0.36	0.69	0.24	−1.1	−1.5
13	1.00	0.19	1.16	1.32	2.0	1.8
14	−0.70	0.26	1.15	1.32	1.0	0.9
15	1.00	0.19	0.96	0.84	−0.4	−0.9
16	−0.30	0.23	1.13	1.03	1.0	0.2
17	0.39	0.2	0.87	0.75	−1.4	−1.2
18	−0.05	0.22	0.9	0.74	−0.9	−1.0
19	0.47	0.2	1.01	1.05	0.1	0.3
20	−0.84	0.27	0.91	0.81	−0.5	−0.4
21	2.33	0.2	1.27	1.75	2.6	3.4
22	−1.06	0.29	0.91	1.69	−0.4	1.4
23	0.35	0.21	1.06	0.92	0.7	−0.3
24	0.22	0.21	0.89	1.03	−1.1	0.2
25	0.51	0.2	1.07	1.26	0.8	1.2
26	0.78	0.2	0.89	0.75	−1.3	−1.4
27	−0.91	0.27	0.85	0.62	−0.8	−0.9
28	1.63	0.19	1.12	1.23	1.4	1.4
29	−0.46	0.24	0.94	0.71	−0.4	−0.8
30	1.07	0.19	1.19	1.15	2.3	0.9
31	0.18	0.21	1.07	1.55	0.7	2.0
32	1.14	0.19	0.96	0.85	−0.5	−0.9
33	−0.52	0.25	1.1	0.93	0.7	−0.1
34	−0.41	0.24	1	0.79	0.1	−0.6
35	−0.30	0.23	0.93	0.73	−0.5	−0.9

Note. Fit statistics are shown in their natural (mean square) and standardized forms (standardized as t).

Results produced by the program; compare "Rasch Diff" below with "Difficulty Estimate" above:

Item	CTT Diff	Rasch Diff	Error	Infit	Outfit
V1	0.86	-0.79	0.26	0.98	0.69
V2	0.86	-0.72	0.26	1.01	0.75
V3	0.65	0.76	0.20	0.98	0.90
V4	0.77	0.00	0.22	1.00	0.88
V5	0.88	-1.01	0.28	0.98	0.76
V6	0.97	-2.49	0.47	1.06	0.83
V7	0.85	-0.66	0.25	0.97	0.65
V8	0.63	0.88	0.19	0.91	1.00
V9	0.74	0.18	0.21	1.07	0.97
V10	0.80	-0.20	0.23	0.92	0.68
V11	0.74	0.18	0.21	1.02	0.96
V12	0.94	-1.81	0.36	0.69	0.24
V13	0.60	1.03	0.19	1.16	1.32
V14	0.86	-0.72	0.26	1.15	1.32
V15	0.60	1.03	0.19	0.97	0.84
V16	0.81	-0.31	0.23	1.13	1.03
V17	0.71	0.40	0.20	0.87	0.75
V18	0.78	-0.05	0.22	0.90	0.74
V19	0.69	0.48	0.20	1.01	1.05
V20	0.87	-0.86	0.27	0.91	0.81
V21	0.35	2.40	0.20	1.27	1.75
V22	0.89	-1.09	0.29	0.91	1.69
V23	0.71	0.36	0.21	1.06	0.92
V24	0.73	0.23	0.21	0.89	1.03
V25	0.69	0.52	0.20	1.07	1.26
V26	0.64	0.80	0.20	0.90	0.75
V27	0.88	-0.94	0.27	0.85	0.62
V28	0.48	1.68	0.19	1.12	1.23
V29	0.83	-0.47	0.24	0.94	0.71
V30	0.59	1.10	0.19	1.19	1.15
V31	0.74	0.18	0.21	1.07	1.55
V32	0.57	1.17	0.19	0.96	0.85
V33	0.84	-0.53	0.25	1.10	0.93
V34	0.82	-0.42	0.24	1.00	0.79
V35	0.81	-0.31	0.23	0.93	0.73
Average:	0.75	0.00	0.24	1.00	0.95
Median:	0.77	0.00	0.22	0.98	0.88
s.d.:	0.13	0.97	0.05	0.11	0.31

4.4.5.2 Steps

Take the program for a test drive by doing these things.

(1) Get a suitable dataset to use, one having cognitive items (e.g. multiple choice items)

Such as MathsQuiz or FIMs

Note1: only 100 or so data records are needed for a test run - if the dataset has more, get into the Data worksheet and insert a new row after record 102 or thereabouts - whenever Lertap encounters an empty row in the Data sheet it thinks it has come to the end of the records. You would insert after row 102 because row 102 has the 100th data record; the first two rows in the Data sheet are "headers".

(2) Run the "Interpret" option

This will produce the Freqs report.

(3) Run the "Elmillon" option

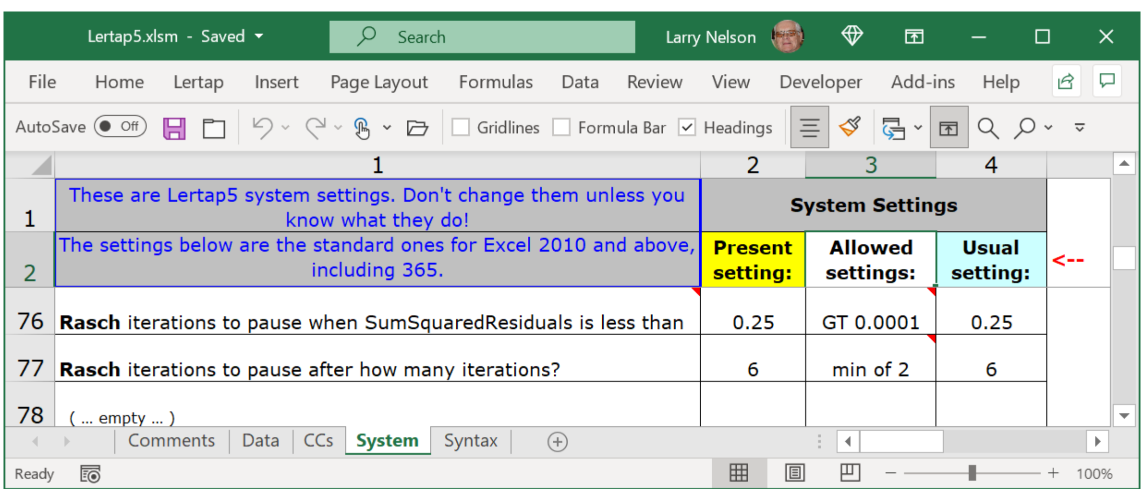
This makes the Stats reports, Stats1f, Stats1b, Stats1ul.

(4) Run the "Item scores and correlations" option from the Run menu

It makes an IStats worksheet.

(5) Select the RaschAnalysis1 option from the Run menu

How it works: the program follows joint maximum likelihood iteration steps much as exemplified by Moulton. Two settings in Rows 76 and 77 of Lertap5's System worksheet provide pausing rules:



When one of the two conditions in rows 76 and 77 is met, the program pauses its automatic iterations and asks users what to do next: stop iterating, or continue. The conditions may be changed by entering new values under the "Present setting" column; if this is done, the Lertap5.xlsm workbook (the main workbook, the one that's always open) should be saved so that the settings are preserved.

As the program steams along it provides feedback in the status bar at the bottom of the Excel screen, indicating the "SumSqrdResiduals" at the end of each iteration. This value will (should) decrease steadily, usually getting very close to zero after 5 or 6 iterations. Note that Excel uses exponential formatting as numbers get very small. A figure such as 0.3062 E-02, for example, would mean 0.003062 (should a value such as this appear there would be no need to ask the program to continue iterations -- the objective is to iterate until Sum is, say, less than .5 or so).

4.4.5.3 Notes

JML, joint maximum likelihood estimation, involves a series of computational iterations and, as a result, a Rasch analysis will require time to complete, depending on the number of test items, the number of students, and the speed of the computer used.

As an example, when processing the BLOT data with 35 items and 150 students, Excel 365, running on a Windows10 laptop with an I5-8250U Intel processor @ 1.60 GHz, took 43 seconds to complete the five iterations needed to get the sum of squared residuals below the default cutoff value of 0.25 -- Excel 2010 running on the same laptop took 18 seconds. On a MacBookPro, also with an i5 processor and running Excel 365, much more time was required: 350 seconds, almost six minutes.

Note: the Excel 2010 times reported here are probably closer to those that would obtain using a current laptop or desktop computer running Excel 365 but having a processor better than an i5. (As of the year 2021 in Australia, Excel 2010 was no longer available for purchase. Note that Excel "365" is a generic label for users having an annual Office 365 subscription; these users will have their version of Excel automatically updated by Microsoft.)

With workbooks from the FIMS study, using Excel 365 to process the whole 14-item dataset with all 6,300 students required about 20 minutes to work through 6 iterations. This reduced to just under 7 minutes using Excel 2010.

A dataset with 60 items and 1,767 students also took 7 minutes using Excel 2010 (compared to 16.5 minutes with Excel 365). Using data from just the first 500 of these students saw this reduce to less than 2 minutes. This dataset may be found here.

The program insures that it will have a complete data matrix by deleting those items and students having either a perfect or a zero score. In the case of BLOT, three students had perfect scores and were automatically excluded from the analysis.

What about missing data? What happens when a student does not answer an item? This document has the asnswer.

INFIT and OUTFIT t-tests have not been included but may be added in the future. Such tests are so highly influenced by sample size as to have, in the opinion of some, limited utility.

INFIT and OUTFIT values can also relate to sample size; Bond and Fox (2015, Appendix B) suggest that INFITs/OUTFITs above 1.3 may hint at problematic items (not fitting the model) for samples of less than 500 students. This might be dropped, they say, to 1.2 for samples between 500 and 1,000, and then down to 1.1 for samples over 1,000.

For a *superior* discussion of the interpretation and use of INFIT and OUTFIT statistics see the entry for a text by Wu et al. (2016) in the Lertap5 references. One of the many useful points raised in this text is that classical test reliability and item discrimination values must be used in conjunction with a Rasch analysis -- it is possible the find acceptable INFIT and OUTFIT values for test items even when the test itself has unacceptable reliability (see p.154 in the text).

4.4.5.4 Exercises

There are a variety of practical exercises that might be undertaken with the program.

An example that could be done would be to run FIMS results through the program, comparing output for Japan students with those from Australia -- the summaries of Rasch diff values made by the macro will exemplify that Rasch results, like those from CTT (classical test theory) and general IRT (item response theory), can indeed be sample dependent.

One could also note the effect of sample size on INFIT and OUTFIT statistics: the larger the sample, the closer these values will be to 1.00. This might be done using the 60-item "M.Nursing" dataset twice: first with all 1,767 students, then with just the first 500.

Wu (2016) makes reference to the relationship between the classical item discrimination index (pb(r) values in Stats1f reports) and the INFIT statistic often used in Rasch item analysis (see p.152 in Wu's text). A related exercise would be to correlate and plot item pb(r) and INFIT values -- Wu suggests that an inverse relationship will be found: low INFIT values can be expected to correlate with higher pb(r) values, or, in other words, items with low INFIT values can be expected to have higher discrimination.

Lertap5 hosts other macros (program modules) for IRT work. While not as easy to use as the RaschAnalysis1 macro, a payoff is that they lead to plots of item performance and fit. Refer to this topic.

Of note is that the macro copies the "Rasch" scores it makes over to the Scores sheet. This makes it possible to correlate original test scores with those from Rasch - students new to this area may be surprised to find near perfect correlations, usually at least r=0.98; the Rasch scores will not histogram too well until they're multiplied by a factor of 10 or 100; this little option makes it easy to do that, while this one makes it easy to update the Scores worksheet's correlations. Students could also readily get a scatterplot of the scores; they will look like those seen in Wu's text (Chapter 7, p.130). (Note that some records in the Scores sheet will not have a Rasch score if a student had a perfect test score to begin with.)

The bottom of the Stats1f worksheet houses a little area with "bands". It can be illuminating to compare the item difficulty bands in the Stats1f bands with the Rasch item difficulty bands made by the macro (as seen here).

The possible impact of mis-keyed items can be investigated using the MathsQuiz dataset. Item 11 was mis-keyed. Has this affected the CTT and Rasch results, and, if so, how?

Two things, two hints, useful when using Lertap5 to undertake exercises such as those suggested above: (1) it is very easy to make a copy of a Lertap5 workbook, and (2) inserting a blank row in the Data worksheet will get Lertap5 to stop reading data records at that point -- when a Data sheet has many hundreds or thousands of records, inserting a blank row after, say, 250 Data rows will at times result in enough data for purposes of the exercise, and will always speed up processing. Note that the "to halve and hold" option may be used to extract random samples of the Data sheet records.

4.4.5.5 User messages

Here are some messages that will or might appear when the RaschAnalysis1 program runs.

> Whoops! This program is designed to work exclusively with test items that have been scored on a right/wrong basis, with 1 point for right and zero points for wrong, a scoring method known as 'dichotomous scoring'.
>
> Sorry, but your test results are not suitable for Rasch scoring.
>
> The empty Rasch worksheet has to be deleted.

The program expects to find that all items have been scored on a right/wrong basis, with zero points for a wrong answer and one point for a correct answer.

It determines if this condition has been met by looking at two rows in the IStats worksheet, the ones labeled "Min" and "Max". An example is shown here from a 56-item multiple choice test - one of the items, "I42", has been found to have a "Min" of 1.00:

Lertap5 IStats matrix, created: 25/11/19.

ID	I1	I2	I3	I4	I40	I41	I42	I43	I44	I55	I56
n	137	137	137	137	137	137	137	137	137	137	137
Min	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
Median	1.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00
Mean	0.62	0.82	0.93	0.46	0.76	0.82	1.00	0.88	0.62	0.45	0.41
Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
s.d.	0.49	0.39	0.25	0.50	0.43	0.39	0.00	0.32	0.49	0.50	0.49
var.	0.24	0.15	0.06	0.25	0.18	0.15	0.00	0.10	0.24	0.25	0.24
Correlations											
I1	1.00	0.33	0.10	0.18	0.09	0.10	0.00	0.14	0.16	0.08	0.13
I2	0.33	1.00	0.10	0.06	0.09	-0.03	0.00	0.06	0.06	0.24	0.01
I3	0.10	0.10	1.00	0.07	0.13	0.03	0.00	0.09	0.16	0.06	0.04
I4	0.18	0.06	0.07	1.00	0.21	0.13	0.00	0.02	0.18	0.10	0.22
I5	0.23	0.13	-0.03	0.17	0.15	0.18	0.00	-0.02	0.23	0.04	0.18
I6	0.13	-0.01	-0.06	0.13	-0.04	0.08	0.00	-0.08	0.13	0.12	0.11
I7	0.22	-0.04	0.17	0.24	0.14	0.11	0.00	0.05	0.16	0.05	0.29
I8	0.05	0.19	0.12	0.06	0.02	-0.11	0.00	0.08	-0.03	-0.06	-0.06

... Stats1ul IStats

Now it just so happens that the program is being a bit too smart. I42 was indeed scored on a right/wrong basis but, of the 137 students, not one got the item wrong. As a result, the minimum score on the item was 1 (one) and the macro wasn't happy.

So the error message wasn't quite accurate in this case, but the result is the same: I42 cannot be processed. It has to be excluded from the test. In Lertap5 this is done by adding a *exc line in the CCs sheet, as shown here:

1	*col (c2-c57)
2	*sub res=(A,B,C,D,E), PER
3	*key DCEBC BABCE ABDDB CBCD
4	*exc (c43)
5	

Data CCs Freqs

Ready

Why c43? Because that's where I42 responses were found in the Data sheet - the Freqs report says so, have a look:

I41 (c42)

Option	n	/137
A	5	3.6%
B	112	81.8%
D	17	12.4%
E	2	1.5%
?	1	0.7%

I42 (c43)

Option	n	/137
C	137	100.0%

I43 (c44)

Option	n	/137
A	1	0.7%

Note that excluding items where everyone got it right, or wrong, is common in Rasch analyses. Such items will not fit the Rasch model. Many other Rasch programs don't sound an error as the RaschAnalysis1 macro does; they, the other programs, make their own processing adjustments internally without saying anything. Incidentally, the same goes for students: those who get all items right, or all wrong, cannot be accommodated by the Rasch model. Dexter is an R package that uses "inf" (infinity) for a student's Rasch score (called "theta" in Dexter) when a student has a perfect score, as exemplified below with results from the 35-item BLOT test:

Copied from DexterGUI, 26-Nov-1

Test	Raw score	theta
BLOT	34	3.924795
BLOT	32	2.702082
BLOT	31	2.354125
BLOT	28	1.615642
BLOT	35	Inf
BLOT	34	3.924795
BLOT	34	3.924795
BLOT	29	1.829436
BLOT	35	Inf
BLOT	30	2.071273

4.4.5.6 Time trials

Refer to this page for a few time trials, measures of the time necessary for Lertap5's Rasch analysis to complete on Windows and Macintosh laptops.

4.4.6 Response similarity analysis

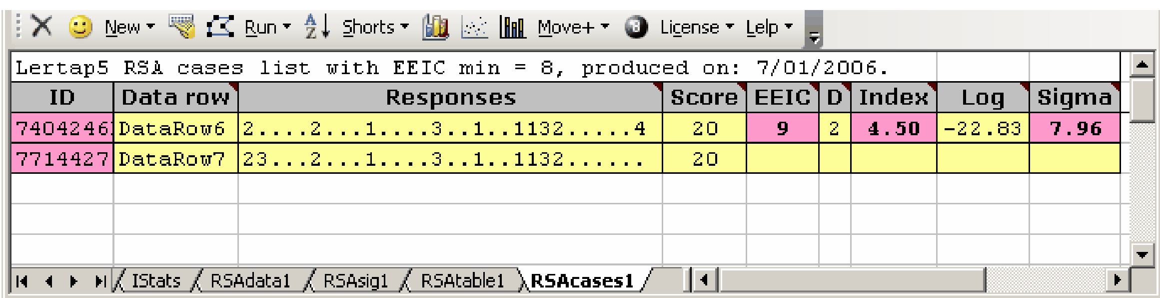
Response similarity analysis, RSA, involves getting Lertap to examine the answers from pairs of students to see if each pair's item responses might be unexpectedly similar. This sort of analysis is generally undertaken to see if some students might have colluded in creating their answers, something which is often considered to be "cheating".

If your data set involves "N" students, and if the objective is to compare all possible student pairs, the number of pairs will equal (N)(N-1)/2. When N=100, there will be (100)(99)/2 = 4,950 student pairs to compare. When N=5,000 there will be more than twelve million (> 12,000,000) student pairs to compare (!). But fear not: Lertap will crunch your pairs without a whinge, asking only that you muster some patience when N gets over 800 or so (see time trials below, and note that it's easy to pare the number of pairs (as it were) by selecting a subset of students, such as, perhaps, all those whose percentage-correct score is from, say, 40% to 95%).

An RSAdata worksheet forms the base for similarity analyses. RSAdata worksheets are made whenever the "Item scores and correlations" option is taken from the Run menu, and the RSA option has been set to "yes" in the System worksheet.

Once an RSAdata worksheet has been created, another option on the Run menu, "Response similarity Analysis (RSA)", will get Lertap to produce its three RSA reports: RSAcases, RSAtable, and RSAsig.

The RSAcases report is the bread and butter of Lertap's RSA analysis. Here's a typical sample:



An RSAcases report presents data for those pairs of students whose item responses have been judged to be "suspect", using criteria developed by Professors Harpp & Hogan.

The item responses given by each pair of students are found under the Responses column, using a format suggested in the "SCheck" program from Wesolowsky (2000):

a full stop (or "period") indicates a correct answer. Each of the two students above had 20 correct answers.

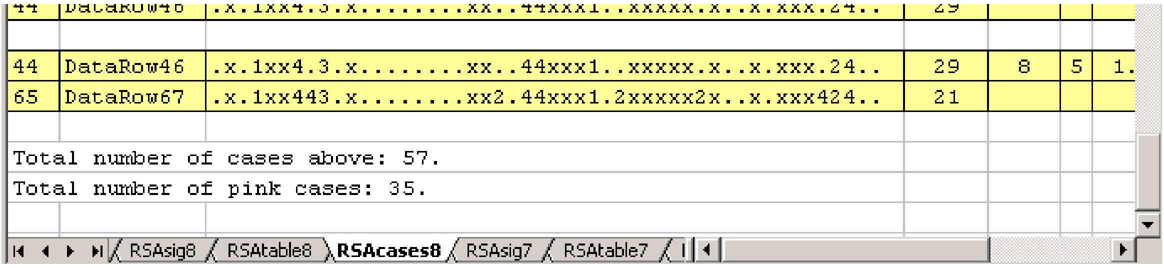
The 2 seen at the start of each response string seen above indicates that both students selected "2" as their response to the first item. On this item, both students made what RSA classifies as an "error": they failed to find the correct answer. Not only did they both make an error, but they made an identical error on the item. Wherever the student incorrect responses match, they have an "exact error in common". It's pretty easy to see that the two students had nine matching errors, nine "exact errors in common". Over all 30 of their item responses, there were only two response differences.

The values of the Harpp-Hogan measures are found under the EEIC, Index, and Sigma columns of the RSAcases report. Briefly, Harpp-Hogan methods are based on (1), determining EEIC, the number of exact errors found in common in student responses; (2), comparing EEIC to "D", the total number of response differences found, a comparison made by dividing EEIC by D, producing the "Index"; (3), developing a response probability measure for the pair of students, and comparing it to a distribution of similar measures formed from non-suspect pairs. The probability measure is found under the "Log" column, with "Sigma" indicating how significant the pair's probability measure was.

EEIC, Index, and Sigma measures are computed for all pairs of students, not just for those whose results come to feature in the RSAcases report. When a pair's EEIC and Index measures are above preset cutoff values, the pair's results are said to be "suspect", meaning that the pattern of their answers to test items was, *perhaps*, suspicious.

All pairs found to be "suspect" are entered in the RSAcases report. To these the final Harpp-Hogan criterion is applied: if the Sigma measure for a suspect pair is above the preset Sigma cutoff value, the pair's results become "significantly suspect", or "very suspect". Their results receive special highlighting in RSAcases: a pink highlight is added to their ID, EEIC, Index, and Sigma entries (an example is seen above). It's these pairs which we might then investigate further. Did they have the opportunity to cheat during the exam? Were they seated close to each other? Were they seen to be using mobile phones, or noted to share the same eraser?

It is the nature of the RSA business to want to have a number representing the extent of possible cheating. In Lertap's RSA analysis, that number corresponds to the number of RSAcases pairs whose entries are "in the pink". When an RSAcases report has more than five entries, a small section at the end of the report summarizes results, as exemplified below:



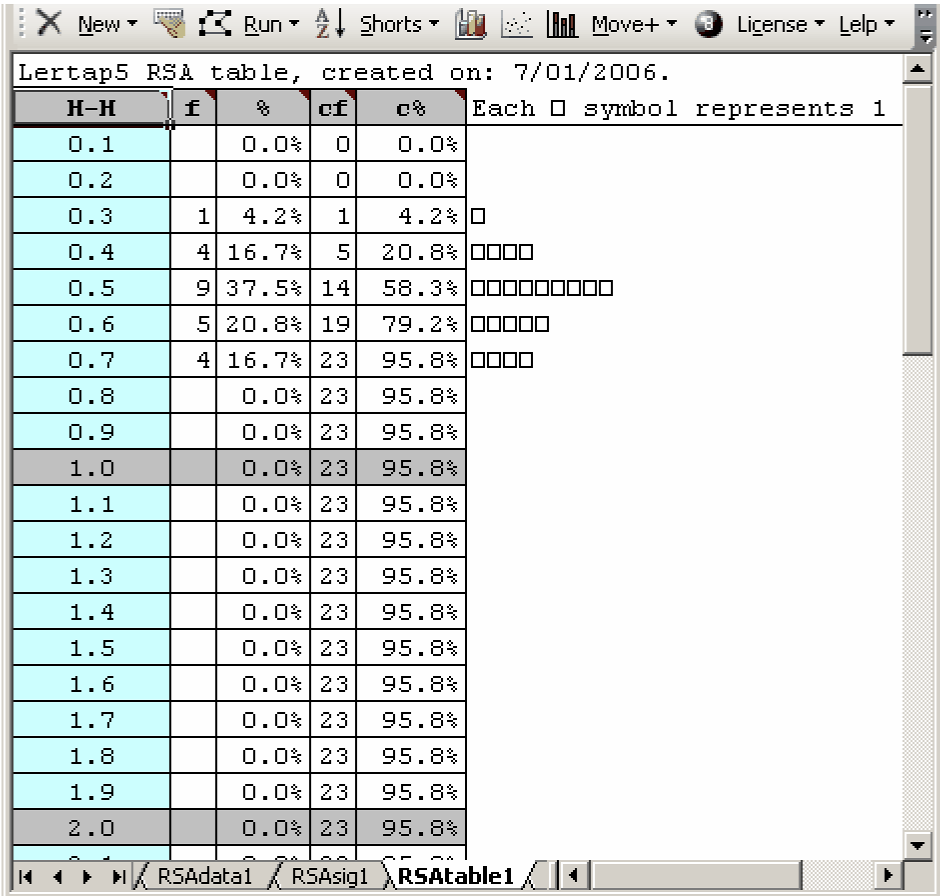
44	DataRow46	.x.1xx4.3.x........xx..44xxx1..xxxxx.x..x.xxx.24..	29			
44	DataRow46	.x.1xx4.3.x........xx..44xxx1..xxxxx.x..x.xxx.24..	29	8	5	1.
65	DataRow67	.x.1xx443.x........xx2.44xxx1.2xxxxx2x..x.xxx424..	21			
Total number of cases above: 57.						
Total number of pink cases: 35.						

In this example, the RSAcases report had 57 entries, 57 paired student results. Of these, 35 were "in the pink". We might say that our RSA analysis uncovered 35 pairs whose item responses were "significantly suspect", or, in Wesolowsky's terms, "excessively similar". We can't yet say for sure that they cheated, but we may well have reason to question their results.

More than one RSA analysis may be applied to the same RSAdata worksheet. As discussed below, there are several options which control how an RSA analysis runs; it is quite common to specify an analysis which looks not at all students, but only at those whose test scores fall within a certain range. In some cases, we might want to exclude "weak" test items from the analysis, as described below.

The RSAcases report conveys the essence of Lertap's analysis, but two other reports are produced for those who care to delve further into the results.

One of these is RSAtable, exemplified in the screen snapshot below:



Lertap5 RSA table, created on: 7/01/2006.

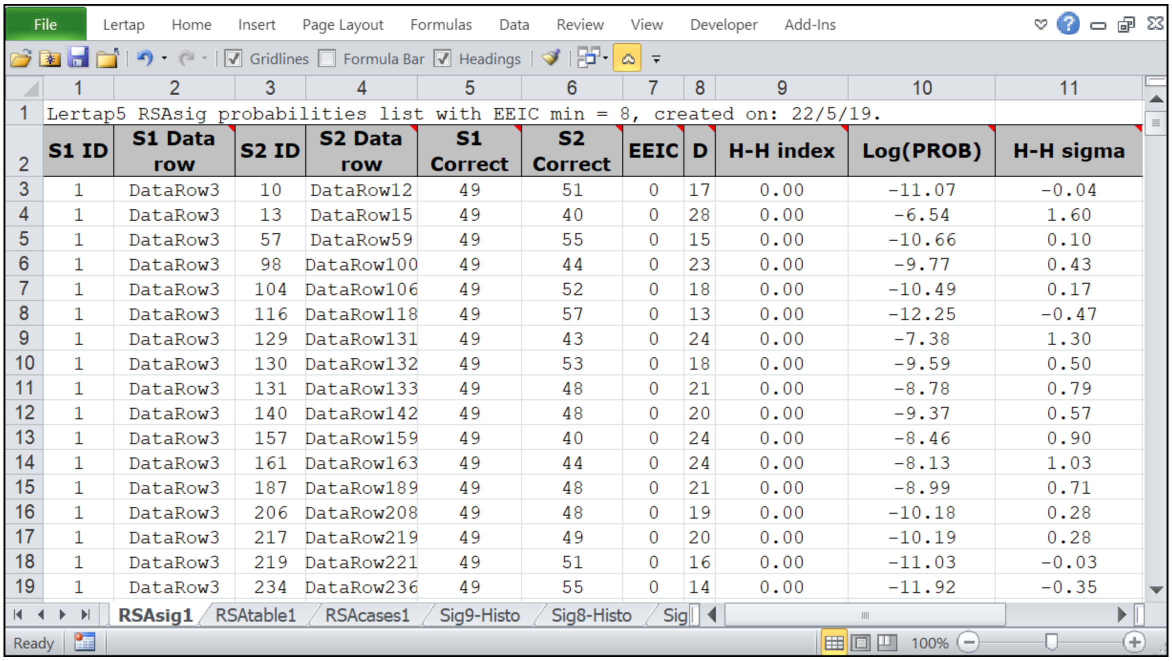
H-H	f	%	cf	c%	Each □ symbol represents 1
0.1		0.0%	0	0.0%	
0.2		0.0%	0	0.0%	
0.3	1	4.2%	1	4.2%	□
0.4	4	16.7%	5	20.8%	□□□□
0.5	9	37.5%	14	58.3%	□□□□□□□□□
0.6	5	20.8%	19	79.2%	□□□□□
0.7	4	16.7%	23	95.8%	□□□□
0.8		0.0%	23	95.8%	
0.9		0.0%	23	95.8%	
1.0		0.0%	23	95.8%	
1.1		0.0%	23	95.8%	
1.2		0.0%	23	95.8%	
1.3		0.0%	23	95.8%	
1.4		0.0%	23	95.8%	
1.5		0.0%	23	95.8%	
1.6		0.0%	23	95.8%	
1.7		0.0%	23	95.8%	
1.8		0.0%	23	95.8%	
1.9		0.0%	23	95.8%	
2.0		0.0%	23	95.8%	

An RSAtable report "plots" the values of the Harpp-Hogan Index measure for all those pairs of students having an EEIC value at or above the preset cutoff. Most H-H Index values will be less than 1.0 in magnitude. To be noted is a special case: the H-H Index is a ratio, one whose denominator, "D", may be zero. When this occurs, Lertap sets H-H Index equal to a value of 999.

Lertap's RSAtable report is made to resemble Figures1, 2, and 3 in Harpp, Hogan, & Jennings (1996).

The RSAtable report is a hold-over from Lertap 5.5 where it was used as the main indicator of potential cheating, a role which has now been assumed by the RSAcases report.

The third Lertap RSA report is RSAsig, a worksheet which contains a wealth of information. RSAsig has three main areas: top, lower-left, and lower-right.



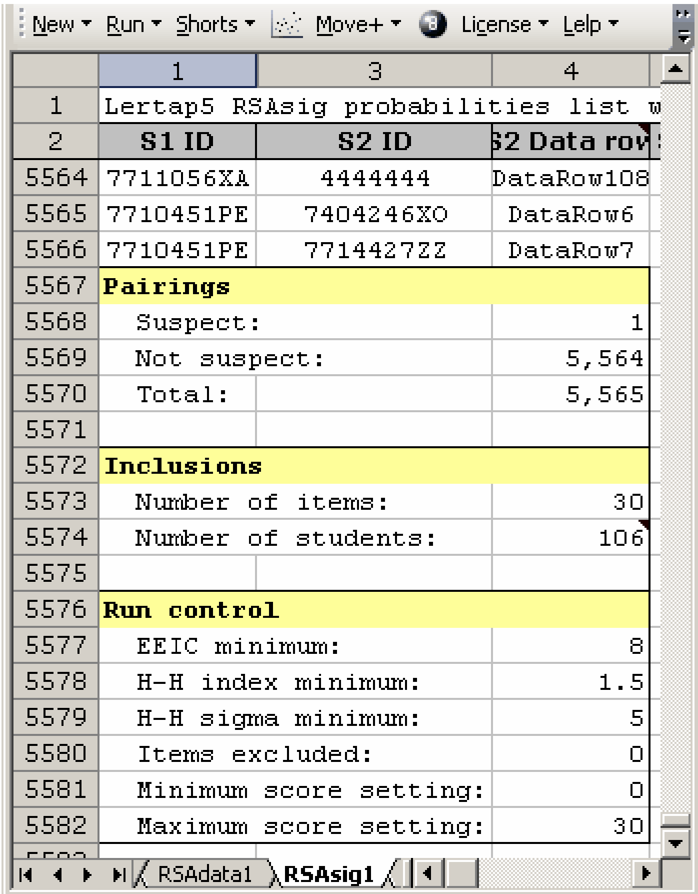
Lertap5 RSAsig probabilities list with EEIC min = 8, created on: 22/5/19.

S1 ID	S1 Data row	S2 ID	S2 Data row	S1 Correct	S2 Correct	EEIC	D	H-H index	Log(PROB)	H-H sigma
1	DataRow3	10	DataRow12	49	51	0	17	0.00	-11.07	-0.04
1	DataRow3	13	DataRow15	49	40	0	28	0.00	-6.54	1.60
1	DataRow3	57	DataRow59	49	55	0	15	0.00	-10.66	0.10
1	DataRow3	98	DataRow100	49	44	0	23	0.00	-9.77	0.43
1	DataRow3	104	DataRow106	49	52	0	18	0.00	-10.49	0.17
1	DataRow3	116	DataRow118	49	57	0	13	0.00	-12.25	-0.47
1	DataRow3	129	DataRow131	49	43	0	24	0.00	-7.38	1.30
1	DataRow3	130	DataRow132	49	53	0	18	0.00	-9.59	0.50
1	DataRow3	131	DataRow133	49	48	0	21	0.00	-8.78	0.79
1	DataRow3	140	DataRow142	49	48	0	20	0.00	-9.37	0.57
1	DataRow3	157	DataRow159	49	40	0	24	0.00	-8.46	0.90
1	DataRow3	161	DataRow163	49	44	0	24	0.00	-8.13	1.03
1	DataRow3	187	DataRow189	49	48	0	21	0.00	-8.99	0.71
1	DataRow3	206	DataRow208	49	48	0	19	0.00	-10.18	0.28
1	DataRow3	217	DataRow219	49	49	0	20	0.00	-10.19	0.28
1	DataRow3	219	DataRow221	49	51	0	16	0.00	-11.03	-0.03
1	DataRow3	234	DataRow236	49	55	0	14	0.00	-11.92	-0.35

The top of a typical RSAsig report has been captured here. S1 and S2 refer to the first and second students in each pair. Such reports contain data pertaining to all student pairs whose item responses are not suspect; these are all those pairs with an EEIC value, and/or an Index value less than respective preset cutoff figures.

(Note: to be included in the RSA analysis, a student must have at least one answer wrong. Students with perfect scores, or totally imperfect scores (not a single correct answer) are excluded.)

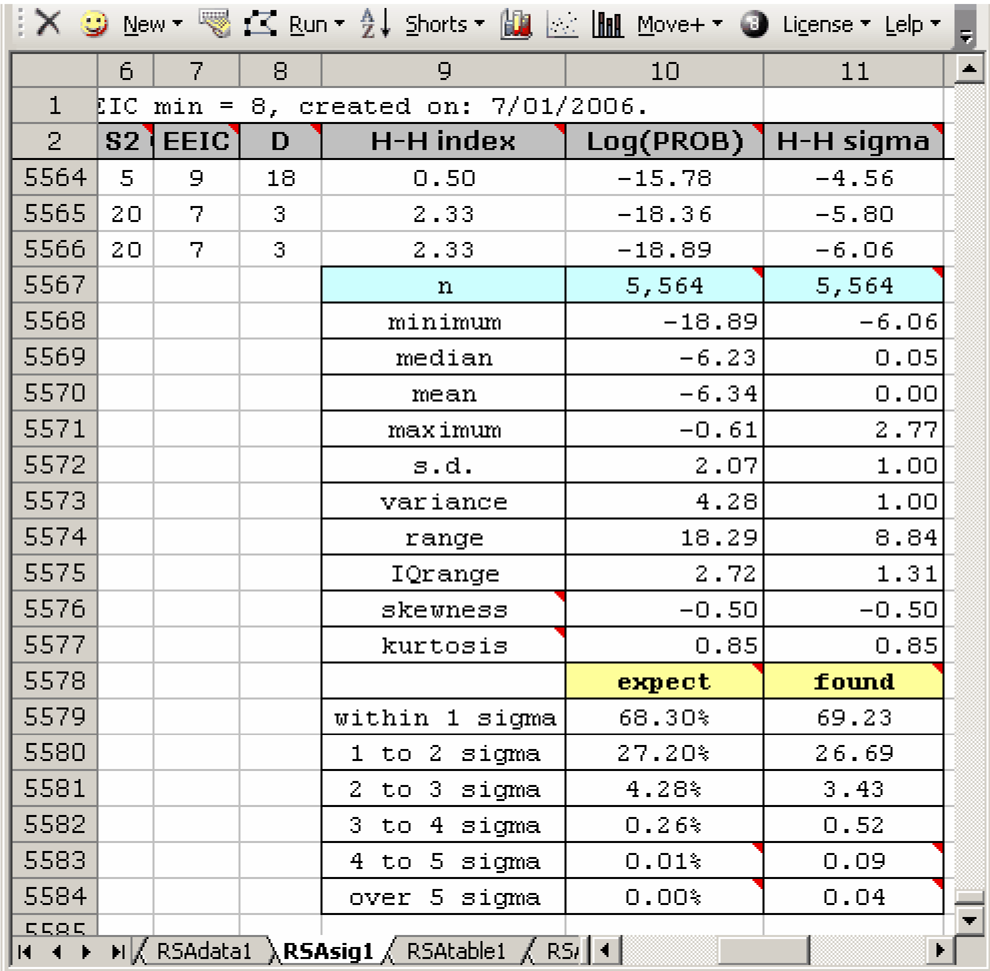
The entries in the RSAsig report are sorted on column 9, from lowest H-H Index to highest. The Log(PROB) column, abbreviated as "Log" in RSAcases, is the logarithm of the Harpp-Hogan response probability measure, "PROB", described in Harpp & Hogan (1993).



	1	3	4
1	Lertap5 RSAsig probabilities list w		
2	S1 ID	S2 ID	S2 Data row
5564	7711056XA	4444444	DataRow108
5565	7710451PE	7404246XO	DataRow6
5566	7710451PE	7714427ZZ	DataRow7
5567	**Pairings**		
5568	Suspect:		1
5569	Not suspect:		5,564
5570	Total:		5,565
5571			
5572	**Inclusions**		
5573	Number of items:		30
5574	Number of students:		106
5575			
5576	**Run control**		
5577	EEIC minimum:		8
5578	H-H index minimum:		1.5
5579	H-H sigma minimum:		5
5580	Items excluded:		0
5581	Minimum score setting:		0
5582	Maximum score setting:		30

The lower-left portion of an RSAsig report is shown above. Only one suspect student pair was found in this analysis of 5,565 total student pairings. Thirty (30) items were involved in the analysis, and 106 students. Cutoff figures for the three Harpp-Hogan criteria are shown as "minimum" values under the "Run control" heading. No items were excluded from the analysis, and a score range of 0 to 30 was processed, corresponding to 0% correct to 100% correct since there were 30 test items.

Note: the "Number of students" excludes students with perfect scores (no wrong answers at all), and it also excludes students with totally imperfect scores (not even a single right answer).



New Run Shorts Move+ License Lelp

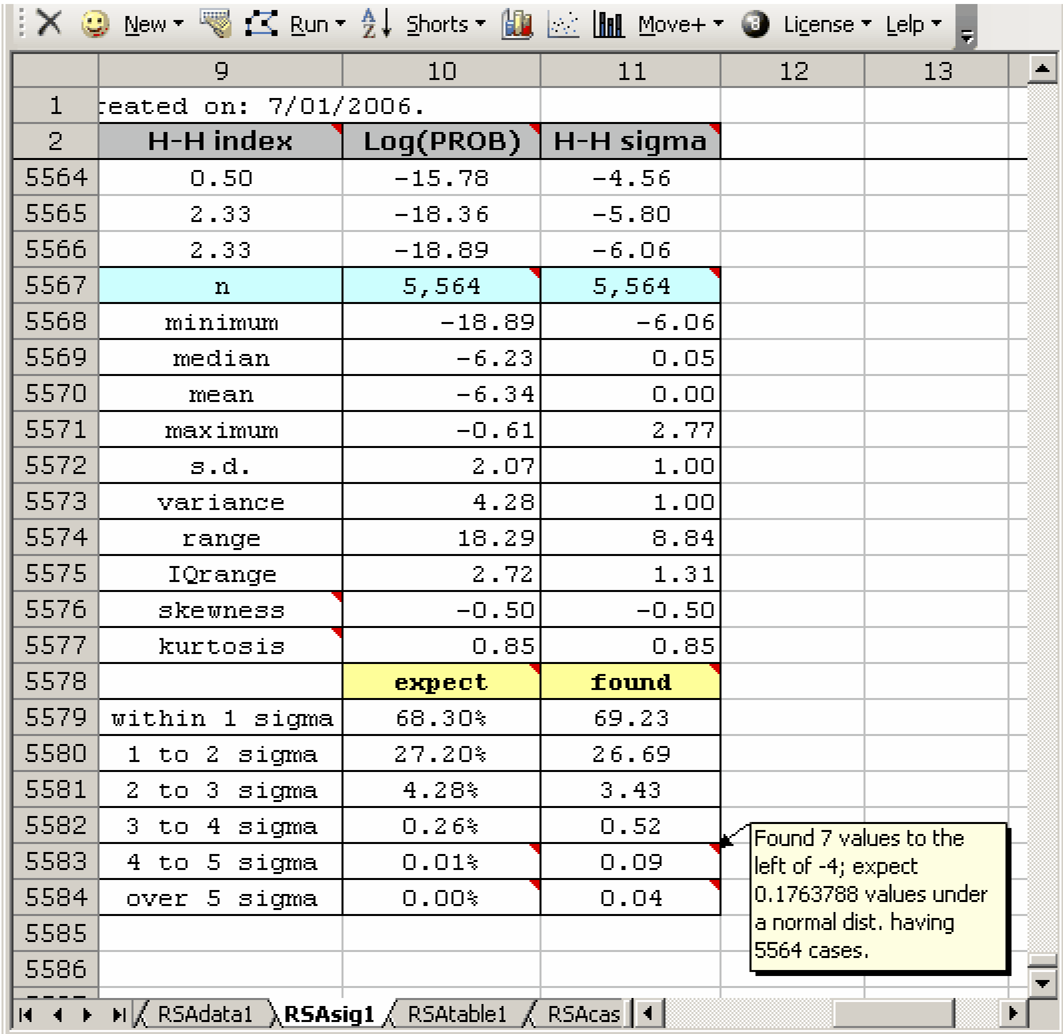
	6	7	8	9	10	11
1	EIC min = 8, created on: 7/01/2006.					
2	S2	EEIC	D	H-H index	Log(PROB)	H-H sigma
5564	5	9	18	0.50	-15.78	-4.56
5565	20	7	3	2.33	-18.36	-5.80
5566	20	7	3	2.33	-18.89	-6.06
5567				n	5,564	5,564
5568				minimum	-18.89	-6.06
5569				median	-6.23	0.05
5570				mean	-6.34	0.00
5571				maximum	-0.61	2.77
5572				s.d.	2.07	1.00
5573				variance	4.28	1.00
5574				range	18.29	8.84
5575				IQrange	2.72	1.31
5576				skewness	-0.50	-0.50
5577				kurtosis	0.85	0.85
5578					expect	found
5579				within 1 sigma	68.30%	69.23
5580				1 to 2 sigma	27.20%	26.69
5581				2 to 3 sigma	4.28%	3.43
5582				3 to 4 sigma	0.26%	0.52
5583				4 to 5 sigma	0.01%	0.09
5584				over 5 sigma	0.00%	0.04

RSAdata1 RSAsig1 RSAtable1 RS

Above is a snapshot of the lower-right area of an RSAsig report. The descriptive statistics, from "minimum" to "kurtosis", have to do with the 5,564 Log(PROB) and H-H Sigma values found in rows 3 through 5566 of the worksheet.

The little "expect - found" table is used to gain an idea of how closely the Sigma values found followed those corresponding to the normal curve. Under a normal, or "Gaussian" distribution, 27.20% of all cases will lie between one and two standard deviations on either side of the mean; for the dataset above, 26.69% of actual cases were found in this region, slightly less than expected. It's clear that the results found for this dataset did not identically match what would have been expected under a true normal distribution, but they're perhaps not too bad.

Small triangles to the upper-right of an Excel cell signify that a comment has been attached to the cell. Letting the mouse hover over such a cell will cause the comment to appear, as seen below:

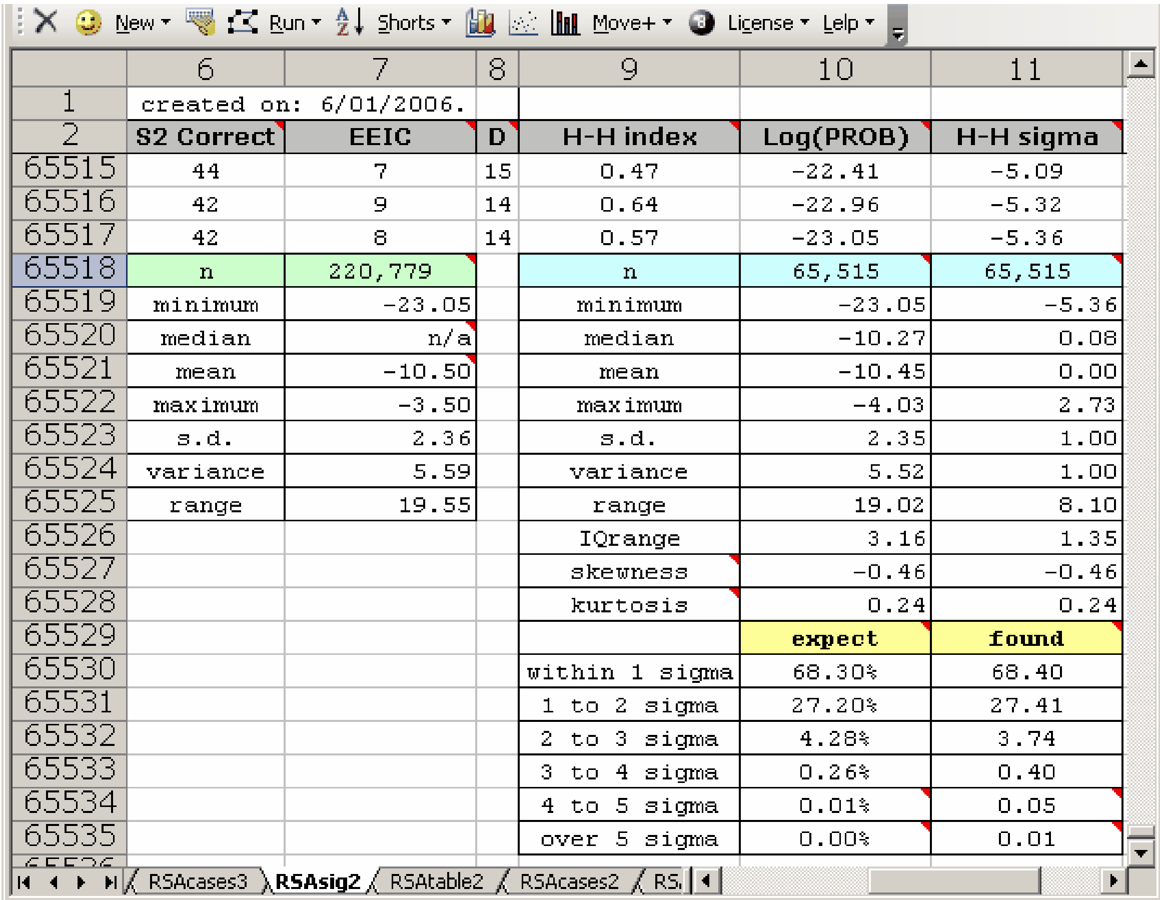


	9	10	11	12	13
1	:eated on: 7/01/2006.				
2	H-H index	Log(PROB)	H-H sigma		
5564	0.50	-15.78	-4.56		
5565	2.33	-18.36	-5.80		
5566	2.33	-18.89	-6.06		
5567	n	5,564	5,564		
5568	minimum	-18.89	-6.06		
5569	median	-6.23	0.05		
5570	mean	-6.34	0.00		
5571	maximum	-0.61	2.77		
5572	s.d.	2.07	1.00		
5573	variance	4.28	1.00		
5574	range	18.29	8.84		
5575	IQrange	2.72	1.31		
5576	skewness	-0.50	-0.50		
5577	kurtosis	0.85	0.85		
5578		expect	found		
5579	within 1 sigma	68.30%	69.23		
5580	1 to 2 sigma	27.20%	26.69		
5581	2 to 3 sigma	4.28%	3.43		
5582	3 to 4 sigma	0.26%	0.52		
5583	4 to 5 sigma	0.01%	0.09		
5584	over 5 sigma	0.00%	0.04		
5585					
5586					

In this case, the comment informs us that seven (7) Sigma values were found to the left of -4 standard deviations, compared to the "0.1763788" values which we would expect to find under a normal curve.

It is easy to get Lertap to graph the Log(PROB) values. Do so by using the histogrammer routine.

The RSAsig report will, at times, differ a bit to the samples seen above. There's a limit to the number of rows an Excel worksheet may have; in Excel 2016 the limit was 1,048,576 rows. Whenever the number of student pairs exceeds a bit less than this number, Lertap stops entering results in RSAsig, but continues to compute a subset of the descriptive statistics (1,048,250 is the precise number of pairs which Lertap will presently report on, corresponding to 1,448 students). It then adds a small table with selected results for all student pairs, as exemplified here:



	6	7	8	9	10	11
1	created on: 6/01/2006.					
2	S2 Correct	EEIC	D	H-H index	Log(PROB)	H-H sigma
65515	44	7	15	0.47	-22.41	-5.09
65516	42	9	14	0.64	-22.96	-5.32
65517	42	8	14	0.57	-23.05	-5.36
65518	n	220,779		n	65,515	65,515
65519	minimum	-23.05		minimum	-23.05	-5.36
65520	median	n/a		median	-10.27	0.08
65521	mean	-10.50		mean	-10.45	0.00
65522	maximum	-3.50		maximum	-4.03	2.73
65523	s.d.	2.36		s.d.	2.35	1.00
65524	variance	5.59		variance	5.52	1.00
65525	range	19.55		range	19.02	8.10
65526				IQrange	3.16	1.35
65527				skewness	-0.46	-0.46
65528				kurtosis	0.24	0.24
65529					expect	found
65530				within 1 sigma	68.30%	68.40
65531				1 to 2 sigma	27.20%	27.41
65532				2 to 3 sigma	4.28%	3.74
65533				3 to 4 sigma	0.26%	0.40
65534				4 to 5 sigma	0.01%	0.05
65535				over 5 sigma	0.00%	0.01

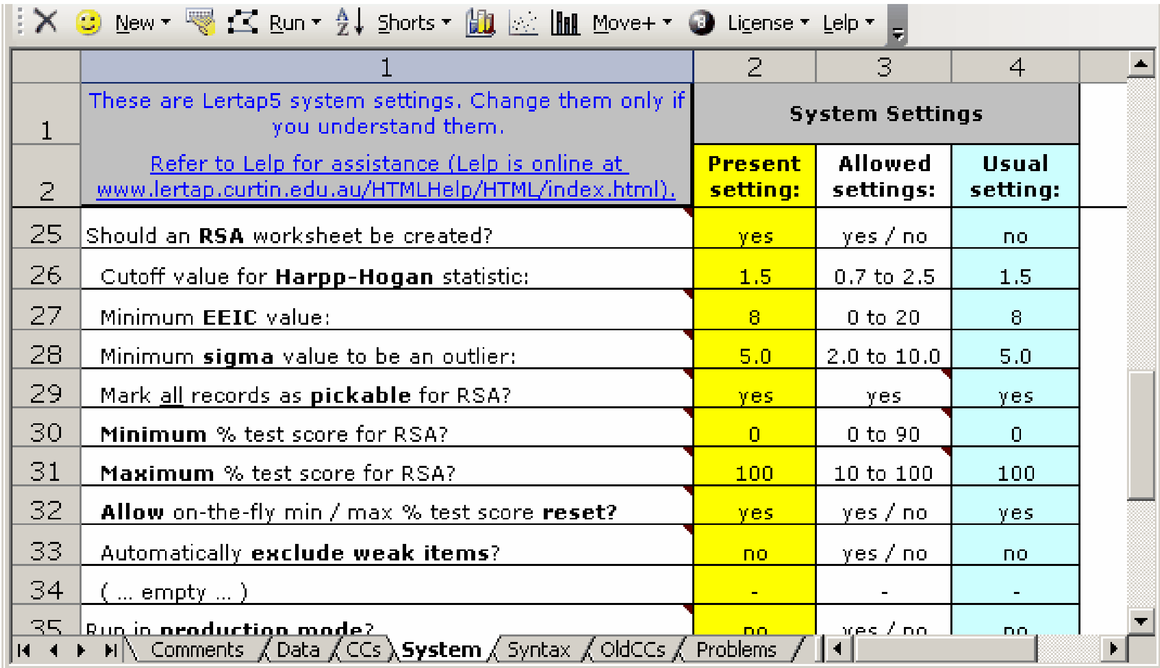
The little table on the left has Log(PROB) minimum, mean, maximum, s.d., variance, and range data for the 220779 student pairs involved in this analysis. We might now consider the 65515 cases whose statistics are given in the right-most table to be a sample from the whole; comparing the sample Log(PROB) mean and s.d. values (-10.45 and 2.35) to those for the population (-10.50 and 2.36) suggests that the sample data are representative.

(Please note that this example is from the Excel 2003 / 2004 versions of Lertap, where the number of rows in a worksheet was limited to 65,536. In the case of Excel 2010, the little table on the left will appear only when the number of students is greater than 1,448, corresponding to about 1,048,000 paired results.)

To read more about response similarity analysis, be sure to refer to the "Related tidbits" at the end of this topic. Of these, if you have time to read only one, make it "Using Lertap 5.6 to monitor cheating on multiple-choice exams".

Lertap's RSA settings

There's a fair smorgasbord of options which control how Lertap goes about its RSA work. Look at the following rows from the System worksheet:



	1	2	3	4
1	These are Lertap5 system settings. Change them only if you understand them.	System Settings		
2	Refer to Lelp for assistance (Lelp is online at www.lertap.curtin.edu.au/HTMLHelp/HTML/index.html).	Present setting:	Allowed settings:	Usual setting:
25	Should an **RSA** worksheet be created?	yes	yes / no	no
26	Cutoff value for **Harpp-Hogan** statistic:	1.5	0.7 to 2.5	1.5
27	Minimum **EEIC** value:	8	0 to 20	8
28	Minimum **sigma** value to be an outlier:	5.0	2.0 to 10.0	5.0
29	Mark all records as **pickable** for RSA?	yes	yes	yes
30	**Minimum** % test score for RSA?	0	0 to 90	0
31	**Maximum** % test score for RSA?	100	10 to 100	100
32	**Allow** on-the-fly min / max % test score **reset?**	yes	yes / no	yes
33	Automatically **exclude weak items**?	no	yes / no	no
34	(... empty ...)	-	-	-
35	Run in **production mode**?	no	yes / no	no

Should an RSA worksheet be created?

If this option is set to "yes", Lertap will produce a worksheet called RSAdata1 whenever the "[Item scores and correlations]" option is selected from Lertap's Run menu. This is the core worksheet for all of Lertap's RSA calculations. If Lertap is running in "[production mode]", there will be one RSAdata worksheet for each subtest. Once an RSAdata worksheet has been created, the "[Response similarity analysis (RSA)]" option may be taken from the Run menu. It is this option which produces Lertap's RSA reports.

Cutoff value for the Harpp-Hogan statistic:

This refers to the H-H index. Harpp and Hogan suggest a minimum of 1.5 for this index.

Minimum EEIC value:

EEIC means "exact errors in common". The recommended minimum is 8, a value which might be lowered to 6 or 7 whenever the number of test items is less than 40.

Minimum sigma value to be an outlier:

Sigma refers to how far a student pair's probability measure is from the mean of the distribution of probability measures. Sigma is a z-score. If the probability measures are normally distributed, a z-score of +5.0 or -5.0 more is a very rare outcome indeed—only 0.0000003 of the area under a normal distribution lies beyond a z-score of 5.0. In practical terms, an exam given to three thousand students will produce about five million pairings of students; if the students have not colluded in their item responses, only about two of the student pairs can be expected to have a sigma greater than 5.0, *assuming that the distribution of probability measures follows a normal distribution*.

Mark all records as pickable for RSA?

This option is, in fact, not yet an option. It may be activated at a future date. In the present version of Lertap, students may be excluded from an RSA analysis by removing the comment (the red triangle) from their RSAdata records; students will also be excluded if their test score does not fall within the range of scores specified by the minimum % and maximum % test score values set in the System worksheet (see immediately below).

Minimum % test score for RSA?

Maximum % test score for RSA?

These two settings determine which students will be included in any RSA analysis. A minimum of 0 (zero) and maximum of 100 will see all students included. Note that experienced users of Harpp Hogan methods will often run several RSA analyses for any given test. They may start with a 0%-100% range for these settings, or 40%-95%, and then reprocess the data with revised settings.

Allow on-the-fly min / max % test score reset?

If this option is set to “yes”, then Lertap will ask you to enter the minimum and maximum % test scores each time you select the “Response similarity analysis (RSA)” option from the Run menu. This completely over-rides the Minimum and Maximum % test score settings in the System worksheet.

Automatically exclude weak items?

For RSA work, “weak items” are those where the number of students selecting the item’s correct answer is less than the number selecting one of the distractors, or less than the number of students who omitted the item. If this option is set to “no”, then Lertap will pause every time it encounters a “weak item”, asking if you’d like to exclude it from the RSA analysis. If the option is set to “yes”, then weak items are automatically excluded. Excluding weak items is strongly recommended; if a test has weak items, the EEIC measure will be inflated, resulting in more “suspects pairs”, that is, more student pairs whose item responses may be judged suspiciously similar (possibly implying cheating). Is it common for tests to have weak items? Yes, it is; difficult items with poorly-functioning distractors will often fall under this definition of a weak item. Note that a “weak item”, in RSA terms, does not necessarily mean a bad item—bad items are, generally, those with a negative discrimination index; it is possible for an item to be weak, in RSA terms, but still have an adequate discrimination figure.

SCheck (Wesolowsky)

The RSA analyses mentioned above all have to do with how Lertap looks at the matter of response similarities. Lertap's procedures are based on those first developed by Harpp & Hogan at McGill University, Canada.

At another Canadian university, Wesolowsky has developed other methods for detecting excessive response similarities. Wesolowsky's SCheck program is based on them. Lertap's RSA procedures will automatically produce a file which will slip right into SCheck -- more about this in steps 2 and 4 below.

Summary of RSA steps

To review, here are the steps required in order to have Lertap do its RSA magic:

1. You have to say "yes" to RSA in the right spot in Lertap's System worksheet. As this topic went to press, the right spot was row 25, column 2.

2. You must go to the Run menu, and click on "Item scores and correlations". This will produce the RSAdata worksheet, and also the SCheckData.DAT file. You'll be able to see the RSAdata worksheet right away as it will form part of your Excel workbook, but the SCheckData.DAT file becomes a separate entity, a file on its own, stored on your computer's hard disk. Where? Well, if you had saved your workbook prior to taking this step, it'll be saved in the same folder as your workbook (otherwise you may have to dig around to find it).

3. Next, back to the Run menu, and a click on "Response similarity analysis" if you want Lertap to make its RSAsig, RSAtable, and RSAcases reports. This option may be selected more than once, each time a new set of reports is created. Note: Lertap5 allows three of the RSA settings to be changed "on the fly" -- the Harpp-Hogan Index cutoff value, the minimum EEIC value, and the minimum sigma value may be changed as the RSA routine runs, over-riding the settings in the System worksheet for the duration of the run. This makes it possible to easily and quickly assess the impact of changing, for example, the minimum EEIC value -- this webpage has a related example.

4. For users with access to Professor Wesolowsky's SCheck.exe program: start SCheck.exe, and get it to work with the SCheckData.DAT file created by Lertap. Read more about SCheck by clicking here. (Note: another SCheck-related RSA worksheet is made by Lertap; it's a special one called "RSAcasesNosort". In some cases this usually-hidden worksheet will be useful even to general RSA users, not only those with SCheck.)

Related comments

What about selecting a subset of data records before getting into Lertap's response similarity analysis? For example, what if you wanted to select only those students who took the exam in the Business school's main lecture hall? There are two ways you could get Lertap to cull out only the records you want.

One way is to use Lertap's *tst card on the CCs worksheet to select the desired records. Of course, you'd have to have a column in the Data worksheet which gives exam location information. Let's say this was column 3, in which case the *tst card might look like this:

```
*tst c3=(Business)
```

Another way is to use the 'Recode' option found under the Move+ menu, entering 'delete' for those records of no interest. The Recode option is more flexible than the *tst method, but it can involve more steps in some cases.

A third way (maybe the easiest) is to use the "NumericFilter2" macro, as described in this topic.

How about using Lertap's RSA support to simply get an estimate of the similarity problems which may pertain to a large data set? Maybe there's too much data, thousands or tens of thousands of students -- too many -- can we possibly get a random sample to work with? But of course. You'd want to read about Lertap's ability to let you Halve and Hold.

Time trials

Having Lertap do RSA things can take time, as you might expect.

From Lertap's viewpoint, there are usually two things to do: make the RSAdata worksheet, and then, when requested, make the suite of RSAsig, RSAtable, and RSAcases worksheets.

As mentioned, the RSAdata worksheet is created when the "Item scores and correlations" option is taken, and the "Should an RSA worksheet be created?" question has been set to "Yes" in the System worksheet.

The "Item scores and correlations" option usually runs fairly quickly; having it make the RSAdata worksheet is basically a trivial matter -- what makes the option work hard is something else: inverting the correlation matrix so that eigenvalues will be computed (row 22 in the System worksheet can be used to stop the inverter from being used).

What is most time consuming for Excel is the creation of the RSAsig, RSAtable, and RSAcases worksheets. For a dataset with 60 items, 500 students, EEIC min at 8, H-H cutoff at 1.5, and with all possible pairs of students included, Excel 2016 took slightly more than 10 minutes to complete its computations when running in the year 2019 on a laptop computer having an Intel Core i5-8250U CPU and 8.00 GB of installed RAM (quite a good laptop but there are certainly faster computers; the processors in laptops are often "clocked down" to save the battery). On the other hand, running on the same computer, Excel 2010 needed just under 3 minutes to do the job.

One thing to keep in mind here: it might not always make much sense to run RSA with data sets housing students from more than one exam venue. Because why? Well, think of what we're trying to figure out: are the item responses from any given pair of students surprisingly similar? If Joe sits the exam in Engineering, and Sally sits the same exam in Commerce, would we want to pose this question? What chance do Joe and Sally have to share exam answers?

We might have all test results in one Lertap Data worksheet, true, but when it comes time for RSA we might well want to break out records according to their exam venue. Interested in this idea? If yes, back up a few paragraphs and read the text around where *tst c3=(Business) is mentioned.

Note: *some users of collusion detection programs suggest that it can be a good idea to refrain from breaking out results by venue as keeping all records, from all venues, may serve as a test of the program's ability to limit the number of false positives; the* *reasoning is that if we know that students like Joe and Sally had no way of sharing answers, then the detection algorithm should not flag them (the flag would be a false positive) -- if it does raise a flag then the quality of the algorithm may be called into question, especially when there might be quite a few of these 'false alarms'.*

Finally, a closing comment: the literature in this area is interesting, and not ambiguous: make it unnecessary to use RSA software by randomly assigning students to seats in the exam venue, and, if possible, by using different test forms, with item scrambling.

Related tidbits:

RSA was used in early 2020 to see if examinees might be finding ways to cheat on exams when able to sit tests at home, under the eye of the "remote proctoring" methods that came to the fore during the COVID19 pandemic. Results from two certified accreditation examinations were compared. On Exam-A no evidence of cheating was found with the 3,000 examinees who took the test at various on-site proctoring examination centers. The same was found with Exam-B and the 2,000 examinees taking the test under on-site proctoring conditions. When testing centers closed for the COVID lockdown period, examinees had to shift to taking both exams under remote proctoring conditions, and Lertap5's RSA analyses did find some evidence of possible cheating among a small (but not insignificant) number of students. (An NCCA conference held late in 2020 was expected to focus, in part, on the rise of remote proctoring and any associated problems.)

Two practical examples of the application of RSA in a university setting may be found starting here; there's another real-life sample, a larger one used in a certification center, available by clicking here.

Lertap5 versions 5.10.9.2 dated after 30 June 2019 are programmed to sort the entries in RSAcases and RSAsig reports on the "Index" column (in RSAcases) and on the "H-H Index" column (in RSAsig), with highest values at the bottom. These are the same statistics, the "Harpp-Hogan index", with slightly different titles. It's possible to see what effect the EEIC cutoff setting has by looking at the last H-H Index values in the RSAsig report. If, for example, the EEIC cutoff is set at 8, those student pairs with EEIC at 6 and 7 will not be seen in RSAcases, but will be seen at the bottom of the RSAsig report. If the corresponding sigma values for such pairs are above the sigma cutoff level then these pairs will also show in RSAcases when the EEIC setting is lowered from 8 to 6. See this topic for a practical example.

For even more about these topics, see "Response Similarity Analysis", a 17-page Word document with lots of similar topics, available via the Internet: click here if you're connected. (Note that this document is quite dated but may be of interest for historical and background reasons.)

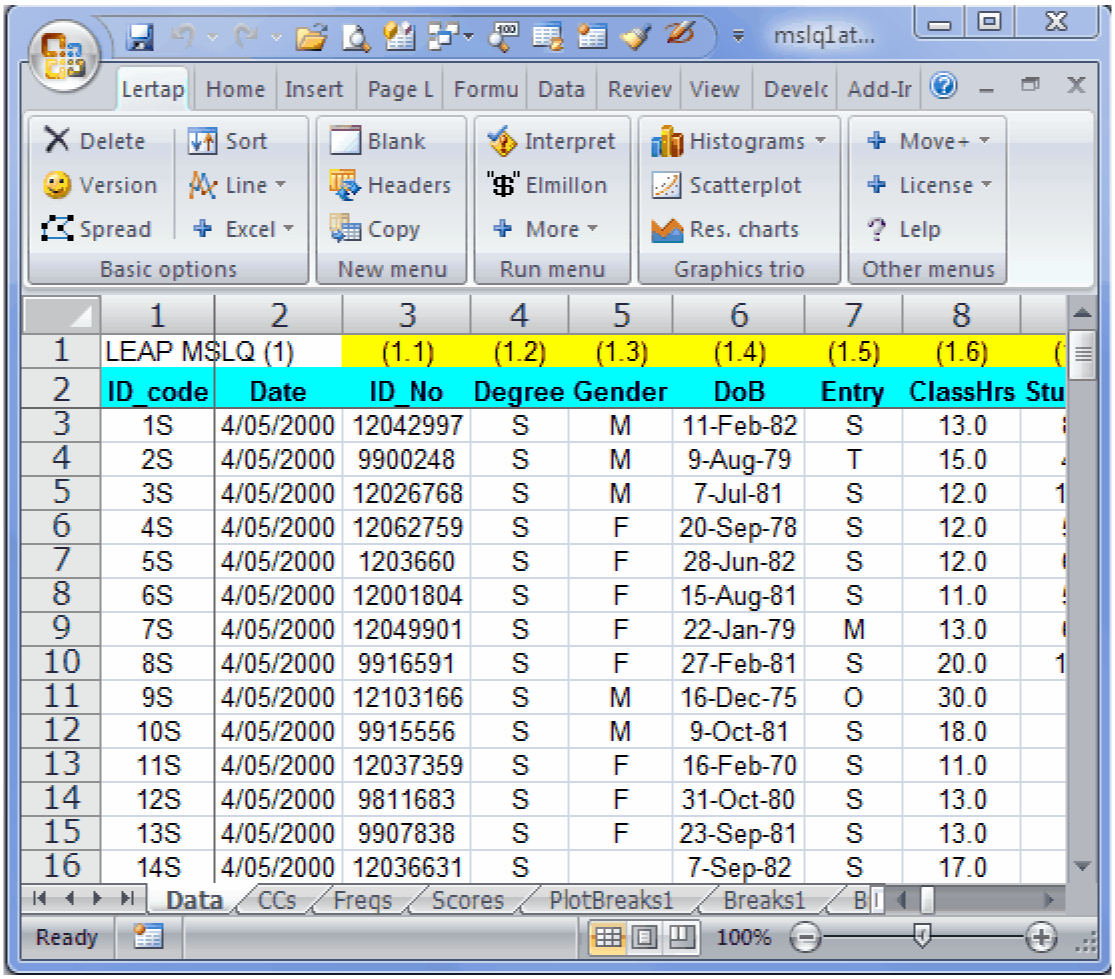
You'll surely want to take in a journal article submitted for publication in 2006: "Using selected indices to monitor cheating on multiple-choice exams", a PDF document, some 15 pages in length. This article mentions other software working in the area of cheating detection, such as *Scrutiny!*, *Integrity*, and *SCheck*.

Then, having looked at the journal article, which was critical of Harpp-Hogan methods, you'll have to take in the best-selling, riveting sequel, a paper which explains how Lertap was modified after Harpp & Hogan revised their original guidelines in response to the journal article. See "Using Lertap 5.6 to monitor cheating on multiple-choice exams".

4.4.7 Breakouts

Suspected you were heading for a breakdown? Lertap can help: use its "Breakout scores by groups" option to obtain a summary table and graph comparing score results for various groups.

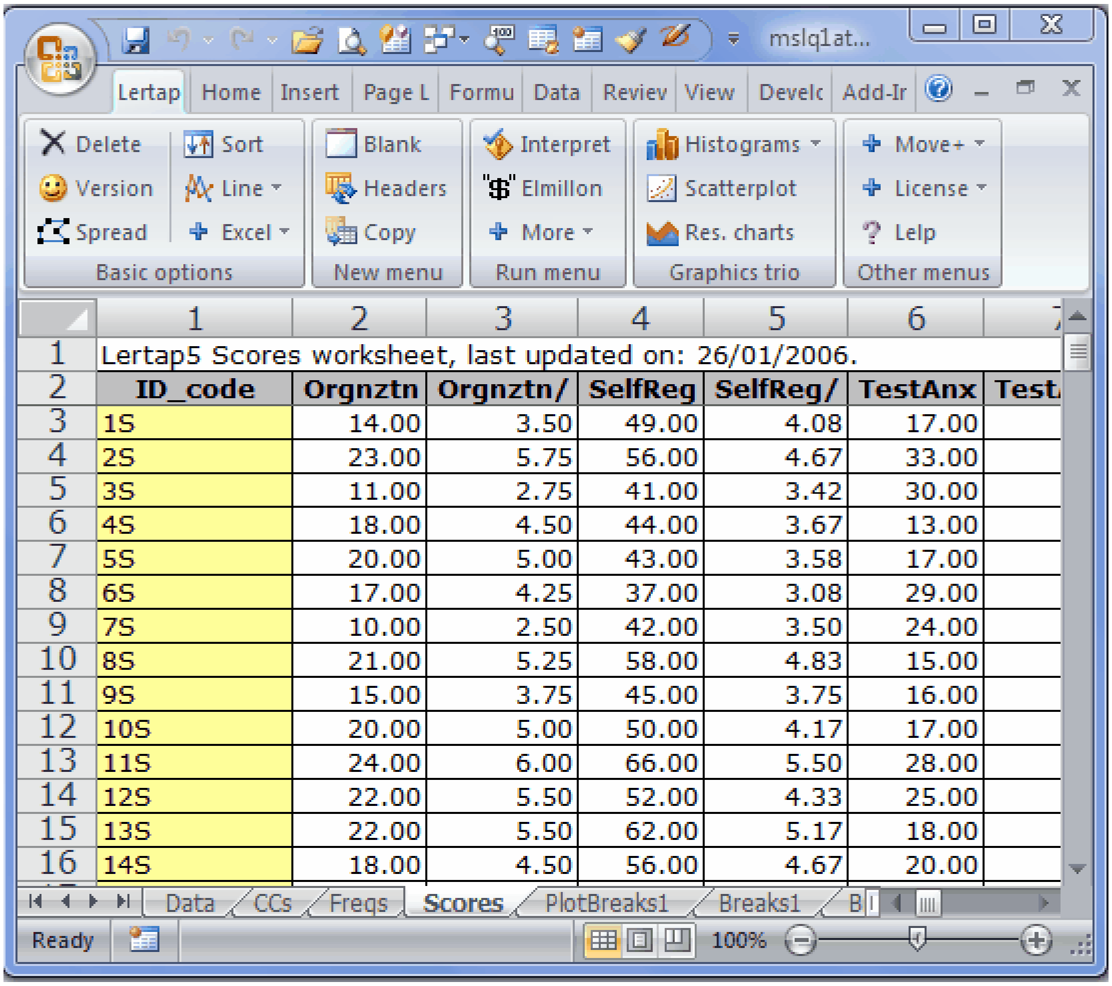
To use this option you will have a column in the Data worksheet which identifies groups.



	1	2	3	4	5	6	7	8
1	LEAP MSLQ (1)		(1.1)	(1.2)	(1.3)	(1.4)	(1.5)	(1.6)
2	ID_code	Date	ID_No	Degree	Gender	DoB	Entry	ClassHrs
3	1S	4/05/2000	12042997	S	M	11-Feb-82	S	13.0
4	2S	4/05/2000	9900248	S	M	9-Aug-79	T	15.0
5	3S	4/05/2000	12026768	S	M	7-Jul-81	S	12.0
6	4S	4/05/2000	12062759	S	F	20-Sep-78	S	12.0
7	5S	4/05/2000	1203660	S	F	28-Jun-82	S	12.0
8	6S	4/05/2000	12001804	S	F	15-Aug-81	S	11.0
9	7S	4/05/2000	12049901	S	F	22-Jan-79	M	13.0
10	8S	4/05/2000	9916591	S	F	27-Feb-81	S	20.0
11	9S	4/05/2000	12103166	S	M	16-Dec-75	O	30.0
12	10S	4/05/2000	9915556	S	M	9-Oct-81	S	18.0
13	11S	4/05/2000	12037359	S	F	16-Feb-70	S	11.0
14	12S	4/05/2000	9811683	S	F	31-Oct-80	S	13.0
15	13S	4/05/2000	9907838	S	F	23-Sep-81	S	13.0
16	14S	4/05/2000	12036631	S		7-Sep-82	S	17.0

In the sample above, the columns labeled Degree, Gender, and Entry would be typical examples of columns which carry some sort of group information.

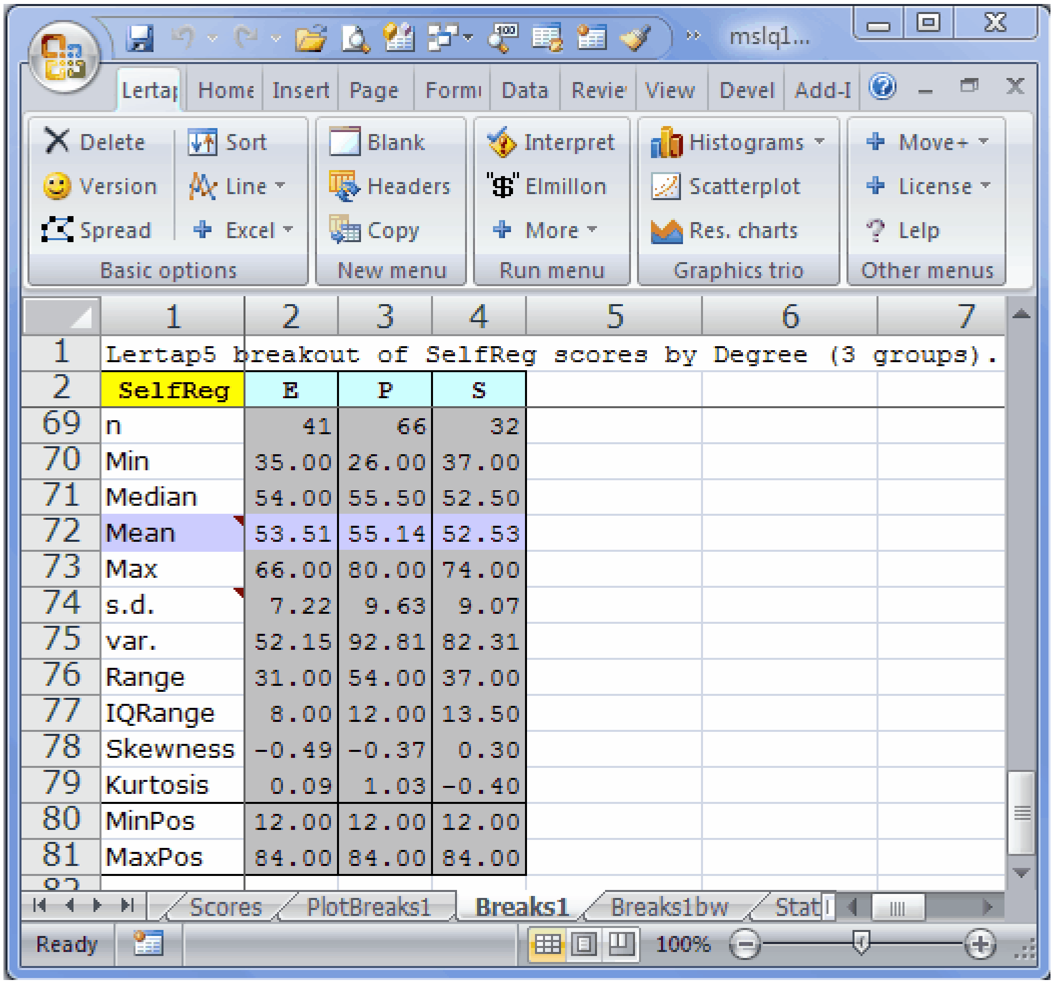
Notes: (1) you can change the codes used in columns such as these using the "Recode macro" available via the Move+ Menu. It is also possible to exclude certain cases from the breakouts, such as, for example, cases with missing data. Click here to read more. (2) the codes should not be numbers -- if, for example, one of the columns is Gender, with codes of 1 and 2, use the Recode macro to make a new column with M and F, or Male and Female, or L and P, or whatever is appropriate, as long as the new codes begin with a letter. *Lertap's breakouts are very likely to be in error if the codes are numbers.*

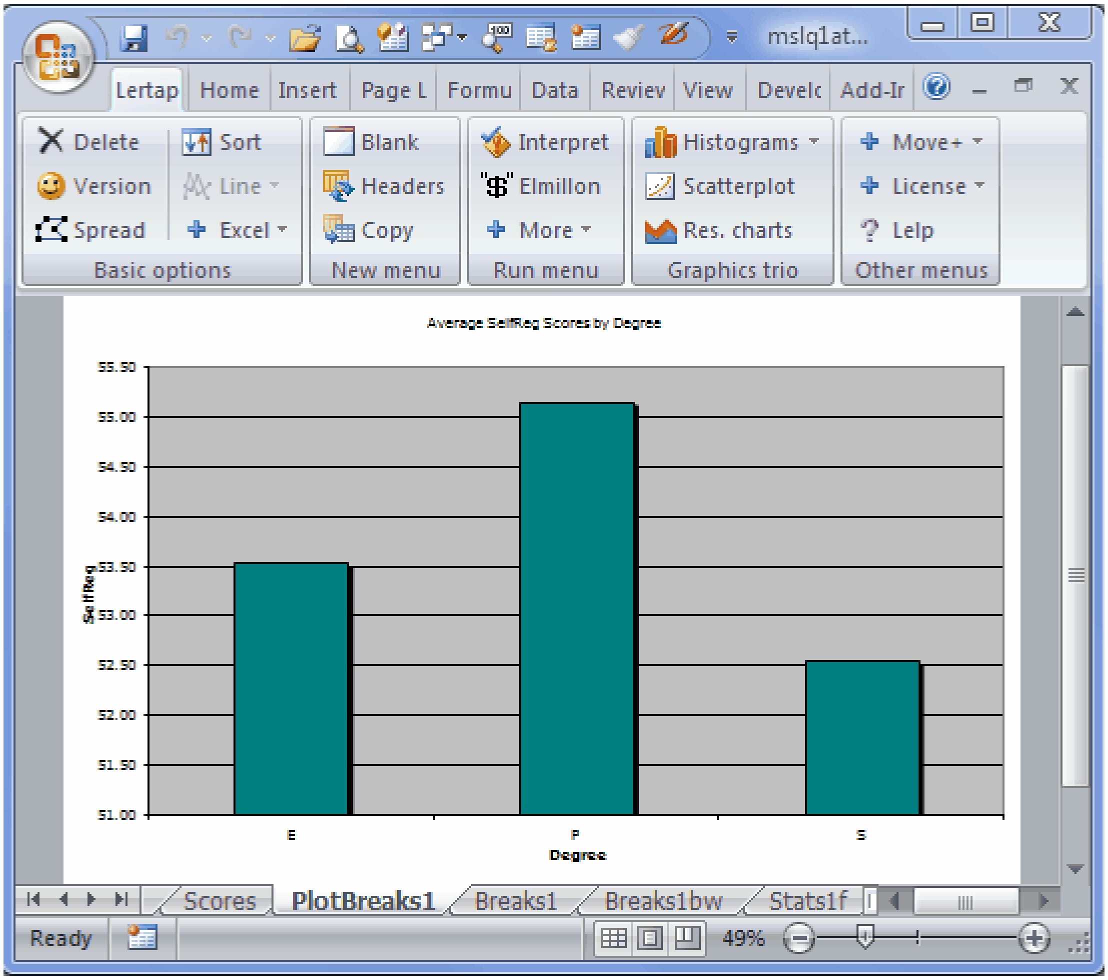


	1	2	3	4	5	6	7
1	Lertap5 Scores worksheet, last updated on: 26/01/2006.						
2	ID_code	Orgnztn	Orgnztn/	SelfReg	SelfReg/	TestAnx	Test/
3	1S	14.00	3.50	49.00	4.08	17.00	
4	2S	23.00	5.75	56.00	4.67	33.00	
5	3S	11.00	2.75	41.00	3.42	30.00	
6	4S	18.00	4.50	44.00	3.67	13.00	
7	5S	20.00	5.00	43.00	3.58	17.00	
8	6S	17.00	4.25	37.00	3.08	29.00	
9	7S	10.00	2.50	42.00	3.50	24.00	
10	8S	21.00	5.25	58.00	4.83	15.00	
11	9S	15.00	3.75	45.00	3.75	16.00	
12	10S	20.00	5.00	50.00	4.17	17.00	
13	11S	24.00	6.00	66.00	5.50	28.00	
14	12S	22.00	5.50	52.00	4.33	25.00	
15	13S	22.00	5.50	62.00	5.17	18.00	
16	14S	18.00	4.50	56.00	4.67	20.00	

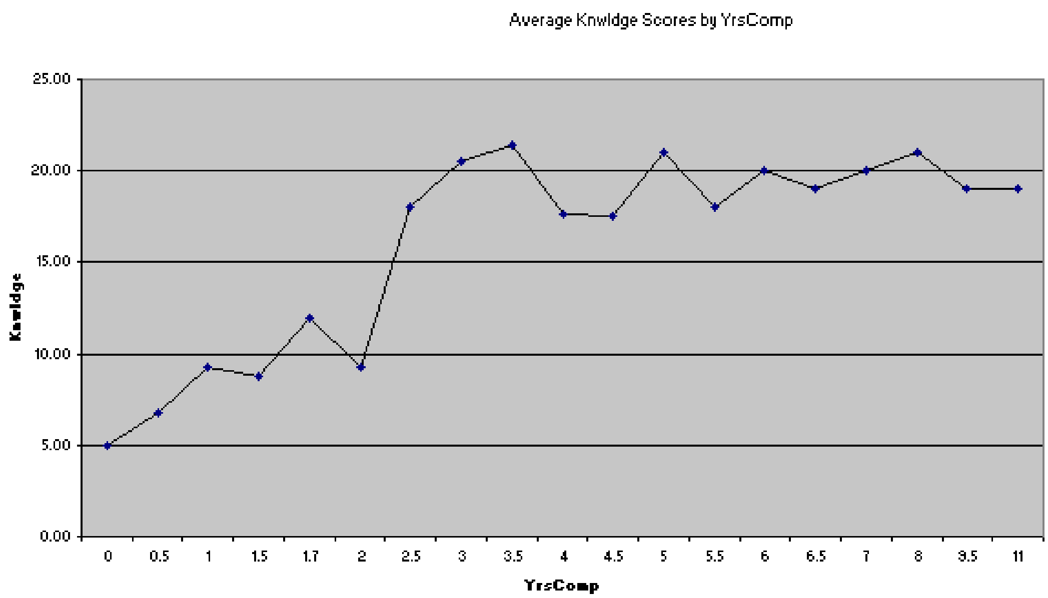
Now, say we had a Scores sheet such as the one above. We might want to cross, say, Degree, column 4 in the Data sheet, with SelfReg, column 4 in the Scores sheet.

We zip up to the Run menu, click on "+ More", and then on "Breakout score by groups", asking for Data column 4 to be broken out using Scores column 4. Lertap produces a breakout report, and a corresponding plot (the statistics in the Breaks1 report are the same as those seen at the bottom of a Scores report):


mslq1...
Lertap Home Insert Page Formu Data Revie View Devel Add-I
Delete Sort Blank Interpret Histograms Move+
Version Line Headers Elmillon Scatterplot License
Spread Excel Copy More Res. charts Lelp
Basic options New menu Run menu Graphics trio Other menus
Lertap5 breakout of SelfReg scores by Degree (3 groups).
SelfReg E P S
n 41 66 32
Min 35.00 26.00 37.00
Median 54.00 55.50 52.50
Mean 53.51 55.14 52.53
Max 66.00 80.00 74.00
s.d. 7.22 9.63 9.07
var. 52.15 92.81 82.31
Range 31.00 54.00 37.00
IQRange 8.00 12.00 13.50
Skewness -0.49 -0.37 0.30
Kurtosis 0.09 1.03 -0.40
MinPos 12.00 12.00 12.00
MaxPos 84.00 84.00 84.00
Scores PlotBreaks1 Breaks1 Breaks1bw Stat
Ready 100%



There can be up to 200 levels in the group column. Values in the column may have any length, and may even be numeric. When there are more than 15 levels, Lertap outputs a line graph instead of a bar graph:



It's possible to change just about everything in Excel charts. Right-click here and there on a chart, and see what happens. Change colours, graph styles, and maybe caffeinated coffee to decaffeinated.

P.S.: we need to whisker something in your ear: there's an option which will let you get a boxplot of group results. Give a click about here.

Analysis of variance table

A Breaks report, as seen in worksheets with names such as "Breaks1", "Breaks2", and so on, terminates with "ANOVA", a small analysis of variance table, rather like the one pictured below:

Analysis of variance			
	df	SS	MS
Between	2	164	82
Within	136	10898	80
Total	138	11062	
F ratio:	1.02	.363	(sig.)
eta sqrd:	0.01		

ANOVA tables provide information which may be used to index the extent of group differences. In this regard, perhaps the most critical statistic shown in the table is "eta sqrd.", short for "eta squared". This statistic has a range of 0 (zero) to 1 (one). If the groups differ greatly with regard to the "dependent variable", SelfReg in this case, eta sqrd. will be close to its maximum possible value of 1.00. If there's little difference among the groups, eta sqrd. will be low, as seen in this example where a value of 0.01 has been found.

Eta squared is referred to as an index of "practical significance"; it's also commonly referred to as an "effect size" estimator: the larger eta squared, the greater the differences among the groups. As Pedhazur and Schmelkin (1991) point out, effect size estimators are often interpreted as being measures of "meaningfulness": the greater the effect size, the more meaningful the differences among the groups.

The F ratio seen in the table is used to test a statistical hypothesis, the so-called "null hypothesis": the average value of the dependent variable, SelfReg, in the populations of people from which our groups have been sampled, is the same: the population groups means are equal (so goes the null hypothesis). The F ratio above, 1.02, results from dividing MS (Between) by MS (Within). To test the null hypot, we used to refer to tables of F values -- these days we can simply ask the computer to see how "significant" the F ratio is. Lertap gets Excel to do this, using Excel's in-built "FDist" function. In our case, FDist says that, were the null hypothesis true, an F Ratio equal to or greater than 1.02 would be observed 36.3% of the time, given the sample sizes used in our "study".

If you are familiar with tests of statistical significance, you will know that the usual guidelines suggest that the null hypothesis will be rejected only when we find an F Ratio whose "significance" is .05, .01, or even less. Here our value, referred to as "(sig.)", is .363, well above the .05 level -- if we were really testing the null hypothesis, we would not reject it in this case.

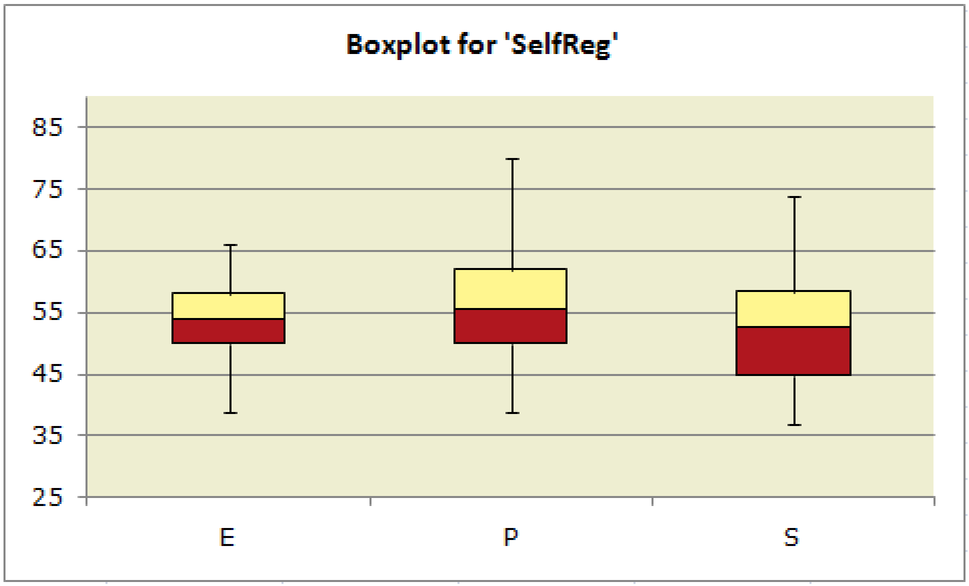
The problem with the F Ratio, and its "significance", is that very small differences in means will sometimes be referred to as being "significant" even when the differences are meaningless; this is prone to happen when sample sizes are large. To circumvent this now well known, widely acknowledged problem, a recommended procedure is to carry along an effect size estimator, such as eta squared: if we find a "significant" F, is it confirmed by a useful effect size (say, for example, at least .10 for eta squared)?

Refer to Thompson (2006), or Pedhazur and Schmelkin (1991), for more readings in this very significant area. Thompson's text is particularly strong on the use of effect-size estimators, and is certainly one of the most compelling sources when it comes to discussing the limitations of tests of statistical significance.

4.4.7.1 Box and whiskers

Once you've used the Run menu to "Breakout scores by groups", you'll have a Breaks report, a worksheet whose name begins with the word "Breaks".

And, once you have a Breaks report, you can use the "Box and whiskers" option to get a graph which looks like the following (this option is found by clicking on the Run menu's "+ More" section, right below "$ Elmillon"):



The data plotted above are from the "MSLQ" study mentioned in Chapter 9 of your favorite read, the Lertap manual.

Results from one of the MSLQ scales, "SelfReg", have been plotted for three groups of student teachers: ECE (Early Childhood Education), Primary, and Secondary.

The top of each box corresponds to "Q3", the 75th percentile; the bottom of each box corresponds to "Q1", the 25th percentile. The line in the middle of each box represents the position of "Q2", the median, the 50th percentile.

Shouldn't the median be halfway between the 75th and 25th percentiles? If the distribution of scores is symmetric, yes, but otherwise no. ("Otherwise" is the usual case as scores are not often exactly symmetric about the median.)

The lines sprouting from the top and bottom of each box are the "whiskers". The top whisker extends from the 75th percentile, Q3, to the highest score which is not an "outlier". Similarly, the bottom whisker extends from the 25th percentile, Q1, down to the lowest score which is not an "outlier". Outliers are discussed below.

A plot such as the one above has a lot of information. The Primary students tended to have higher SelfReg scores, and the range of their scores was the greatest: the P group's whiskers extend a bit further than those for the other groups.

The highest median score is found in the P group.

The SelfReg scores of the Secondary students are, by and large, the weakest: their median score and their "Q1" score are lowest of the lot.

How it works

The "Box and whiskers" option works by first making a copy of a Breaks report. If you were looking at a Breaks1 report, for example, then you'll get a new report called "Breaks1bw", with "bw" standing for box and whiskers.

This new report will have a section at the bottom with summary score data organised in the fashion shown here:

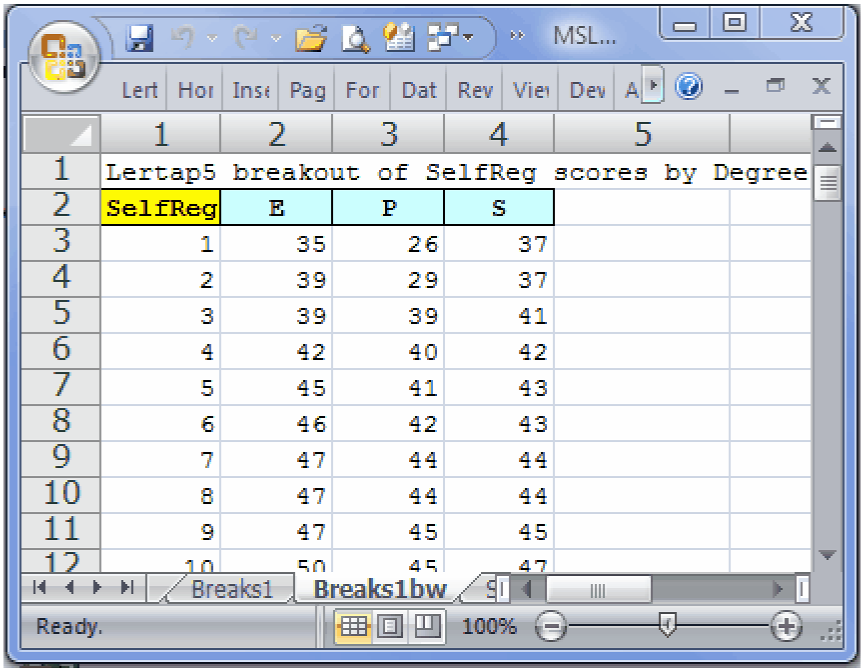
SelfReg	E	P	S
n	41	66	32
Mean	53.51	55.14	52.53
s.d.	7.22	9.63	9.07
Median	54.00	55.50	52.50
Q1	50.00	50.00	44.75
Q3	58.00	62.00	58.25
Minimum	35.00	26.00	37.00
Maximum	66.00	80.00	74.00
25th Pct	50.00	50.00	44.75
50th Pct	4.00	5.50	7.75
75th Pct	4.00	6.50	5.75
Bottom whisk	11.00	11.00	7.75
Top whisker	8.00	18.00	15.75
Lowliers	1	2	0
Highliers	0	0	0

Boxplots, also known as box-and-whisker plots, are an invention of Tukey (1977). They're a very useful way to visually compare group scores.

Lertap uses Excel's stacked-column chart as the basis for its boxplots. The core information for the chart is found in the three "Pct" rows: 25th Pct, 50th Pct, and 75th Pct.

The length of each whisker is determined by finding the highest and lowest scores in each group which are not "outliers", that is, not extreme scores. Tukey defined the limits for outliers as 1.5 times IQR, the inter-quartile range, Q3-Q1. Looking at the "P" group in the table above, IQR = 62 - 50, or 12. Multiplying this by 1.5 gives 18; scores above Q3 + 18, and below Q1 - 18, are Tukey's outliers. In the case of the P group, any score above 80 (Q3 + 18) and below 32 (Q1 - 18) will be deemed an outlier.

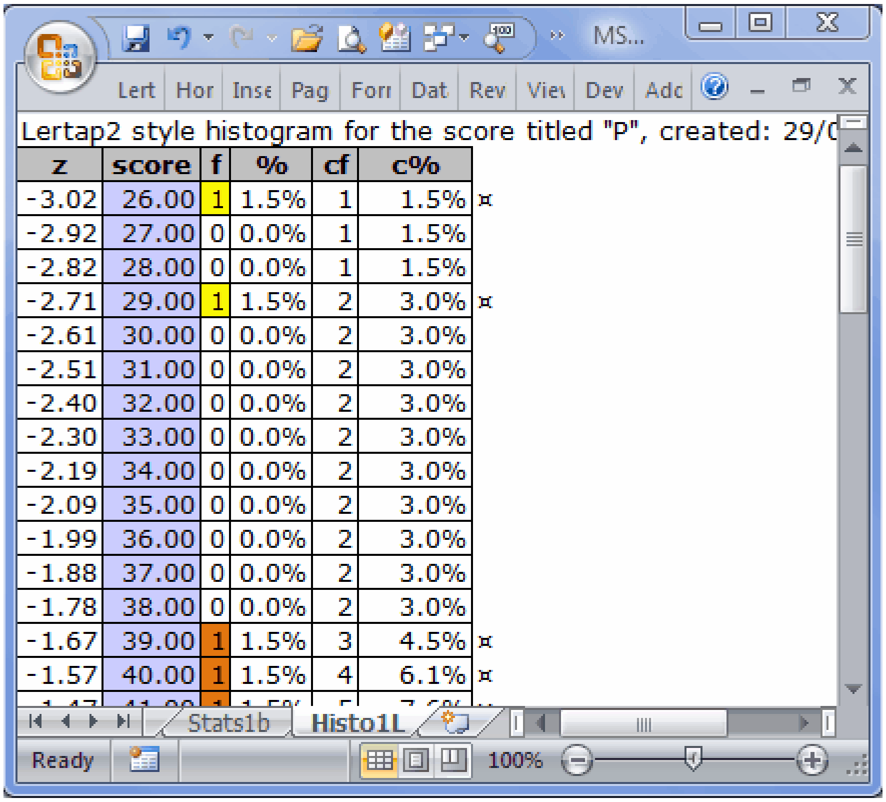
Many boxplot routines, such as that found in SPSS, indicate the presence of outliers by showing asterisks above and/or below the whiskers. Lertap does something different: if there are outliers, the number of them is shown in the "Lowliers" and "Highliers" rows. The wee table above says (for example), that there are 2 "Lowliers" in the P group. You can see them if you scroll up to the top of the worksheet, as exemplified here:



	1	2	3	4
1	Lertap5 breakout of SelfReg scores by Degree			
2	SelfReg	E	P	S
3	1	35	26	37
4	2	39	29	37
5	3	39	39	41
6	4	42	40	42
7	5	45	41	43
8	6	46	42	43
9	7	47	44	44
10	8	47	44	44
11	9	47	45	45
12	10	50	45	47

The scores of 26 and 29 are the two outliers for group P. The effective lowest score for this group is 39, which is 11 points below group P's Q1. These 11 points are the length of the "Bottom whisker" for group P.

A histogram would be another way to look at how outlying a Tukey outlier is. If you switch over to the Breaks1 report, and then take the "Histograms" option, using column 3, the column with group P's scores, you'll see something like the following:



z	score	f	%	cf	c%	
-3.02	26.00	1	1.5%	1	1.5%	¤
-2.92	27.00	0	0.0%	1	1.5%	
-2.82	28.00	0	0.0%	1	1.5%	
-2.71	29.00	1	1.5%	2	3.0%	¤
-2.61	30.00	0	0.0%	2	3.0%	
-2.51	31.00	0	0.0%	2	3.0%	
-2.40	32.00	0	0.0%	2	3.0%	
-2.30	33.00	0	0.0%	2	3.0%	
-2.19	34.00	0	0.0%	2	3.0%	
-2.09	35.00	0	0.0%	2	3.0%	
-1.99	36.00	0	0.0%	2	3.0%	
-1.88	37.00	0	0.0%	2	3.0%	
-1.78	38.00	0	0.0%	2	3.0%	
-1.67	39.00	1	1.5%	3	4.5%	¤
-1.57	40.00	1	1.5%	4	6.1%	¤

The two Tukey outliers have been highlighted in yellow above, and it's now possible to get another idea of how extreme these scores are: notice the distance between them and the score of 39 (highlighted in orange). These scores are so extreme that they caused the score distribution to have negative skewing (-0.37). However, the boxplot indicates that, once these scores are removed from the scene, the skewing actually appears to be towards the high end of the distribution; this is so as the top whisker in the boxplot is substantially longer than the bottom one.

> *Note*: a common cause of Lertap boxplot failure relates to the codes / names used to label the groups. In this example, the labels are E, P, and S. Failure is likely when the codes are numbers, or single digits, such as 1, 2, 3. And failure is pretty much guaranteed if one of the labels is blank. If you get a message from Lertap or Excel concerning this matter, go back to the corresponding Breaks report, such as Breaks1, and change the codes. Note that it's possible to still use digits if you convert them to text -- an easy way to do this is to simply place an apostrophe before the digits: '1, '2, '3 and so forth.

Flexibility

Once you have one of these boxplots on the screen, it's very easy to alter its appearance. Right-click on the chart, and Excel will open up numerous options. To set the score range used in the plot, right-click on one of the scores seen along the vertical axis, and then take the "Format Axis" options.

Note that you can delete whole columns in the Breaks1bw report, and the corresponding box and whisker set will disappear from the plot. This is useful when you've got too many groups, and want the boxplot to highlight only some of them. For example, if above we eliminated column 2 from the Breaks1bw report, the resultant boxplot would contain just box-whisker sets for the P and S groups.

Related tidbits:

More about the use of Excel stacked column charts for boxplots may be found in work by Jon Peltier (Lertap's boxplots are largely based on Peltier's suggestions):

http://peltiertech.com/Excel/Charts/BoxWhisker.html

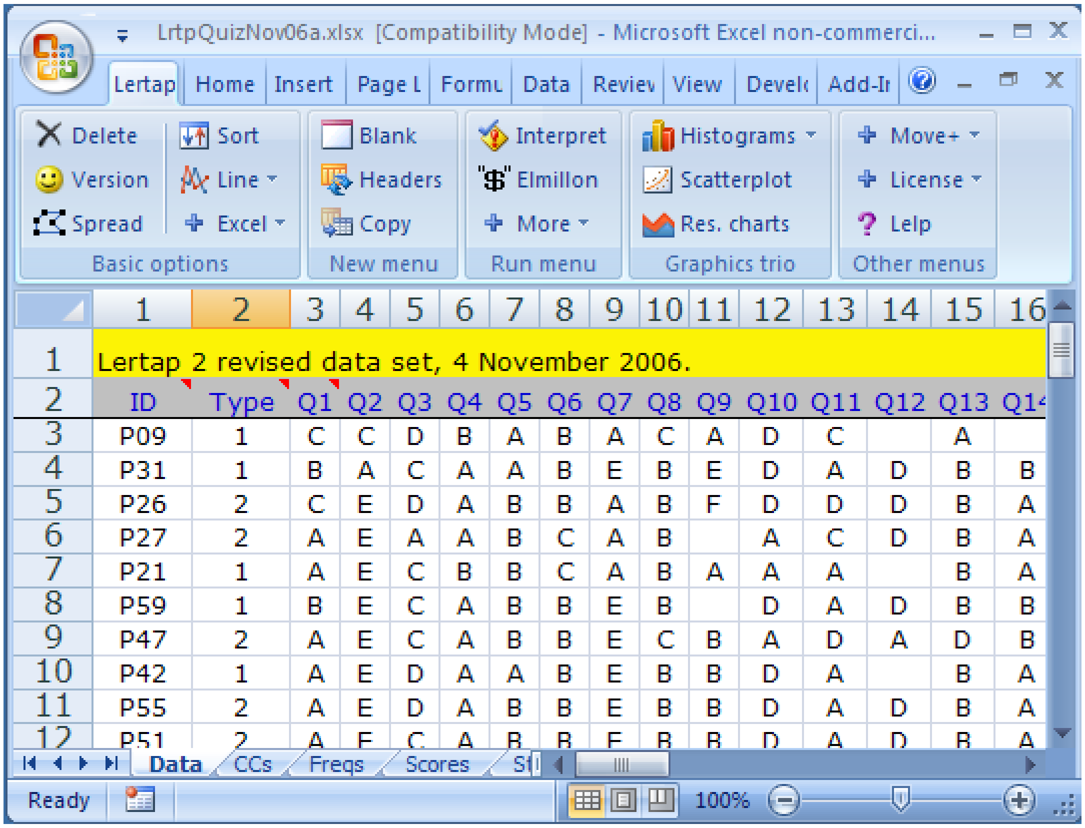
Much more about boxplots, or box and whisker plots, may be found on the internet, and in many statistics texts (such as Thompson (2006)).

4.4.8 Ibreaks

"Ibreaks" refers to the process of creating a report or two which summarizes the way groups have responded to test items.

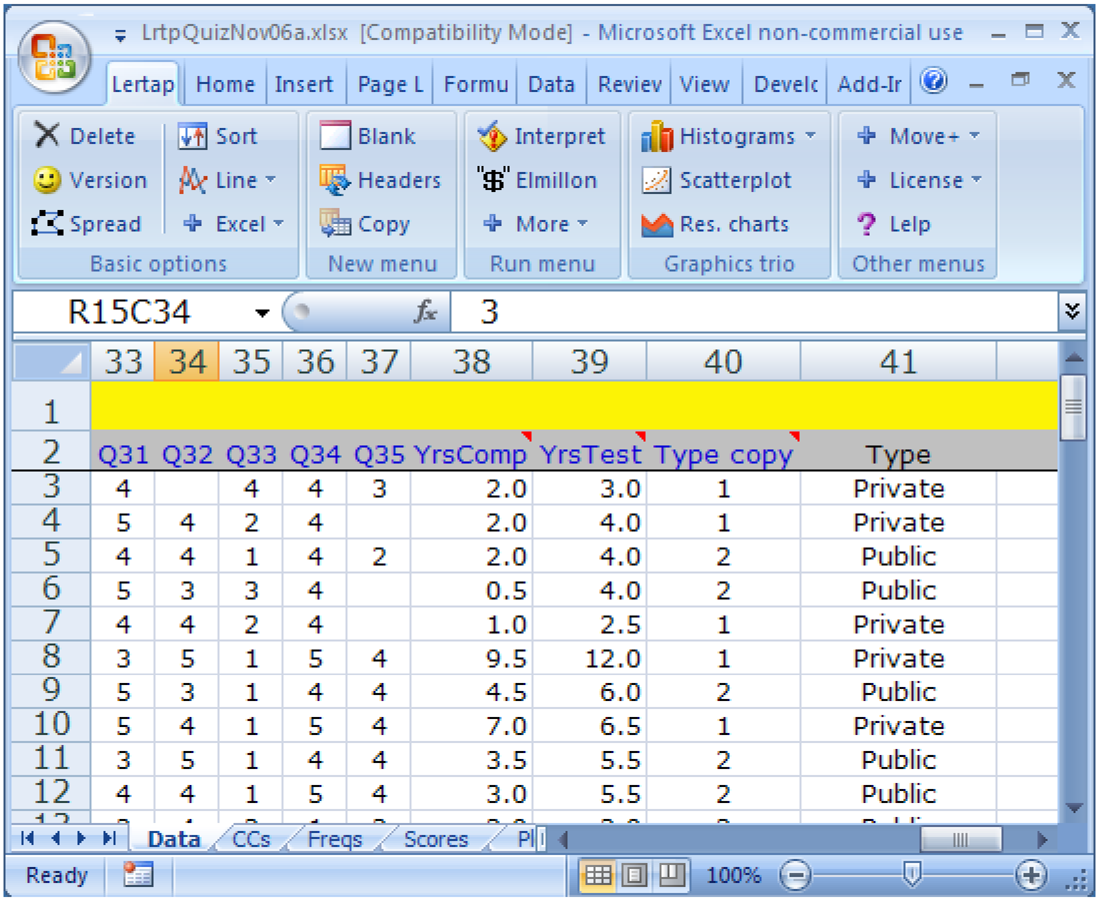
Ibreaks requires two things: a column in the Data worksheet which identifies groups, and a Scores worksheet which it can interrogate as needed. (Scores worksheets are created by the Elmillon option, while Data worksheets are created by you -- click here if you need to be reminded about Data worksheets and their proper format.)

Caveat: The codes used to identify groups must begin with a letter. If, for example, gender has been coded as 1 or 2, Lertap5's recoder should be used to create a new column with new gender codes of, for example, F and M.



In the sample above, a group code, "Type", is seen in column 2. In this case, a "1" was used to denote workshop participants who were from private schools, with "2" used to code participants from public schools.

These codes are *not* optimal. As mentioned in the caveat above, they should start with a letter. The data recoder may be used to quickly change the codes. The screen snapshot below shows recoded Type values in column 41:

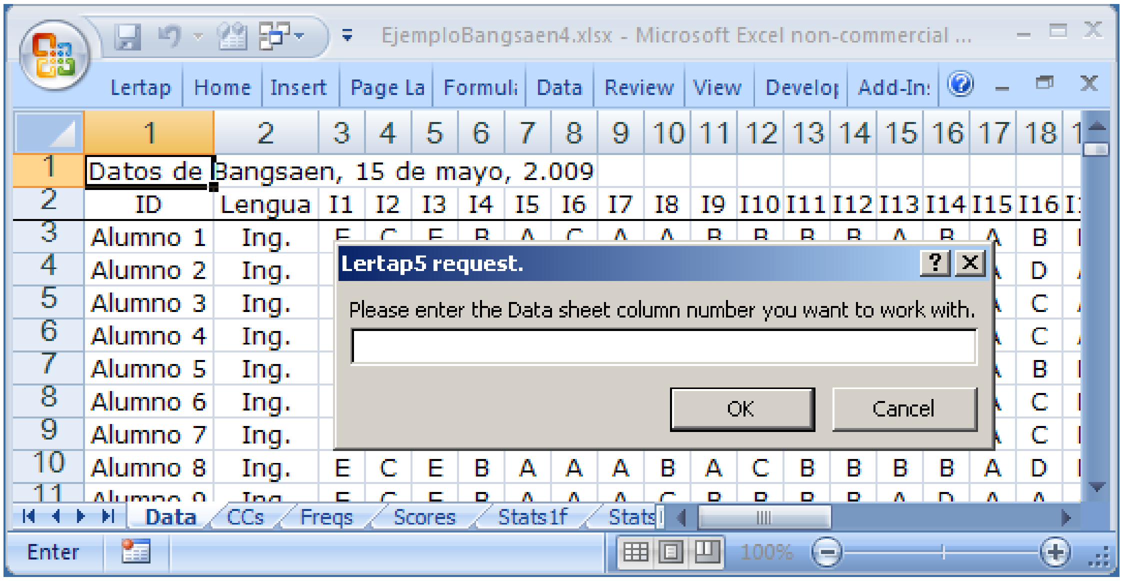


In this example, an original Type code of "1" has been recoded as "Private". The new code can be anything, as long as it begins with a letter. The new codes may have any length, but short codes result in Ibreaks tables and charts which tend to look a bit better. What's 'short'? -- say 8 characters or less.

Note: the recoder, also known as the "Recode macro", is available via the Move+ Menu. It is also possible to exclude certain cases from Ibreaks, such as, for example, cases with missing data. Click here to read more.

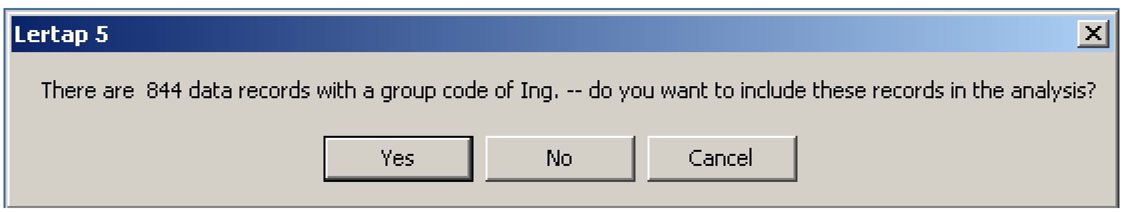
What might happen if group codes are numeric? The sky will fall; your bicycle tires will go flat. A more substantial outcome: the charts made by Excel will be wonky (incorrect).

Okay then, all's in order for an Ibreaks run. Ready to answer a batch of questions? Ibreaks will present several as it strives to do what you require. We'll be looking at an example of a cognitive test with 40 multiple-choice items scored on a right/wrong basis. A long-used test presented for years in a country's native language was translated to English. One group of high-school students sat the traditional native-language version of the test, while another group sat the English-language version. The two test forms were judged to be equivalent, differing only in the language used.



The first question Ibreaks asks concerns the location of the column with the group code. In this example, groups are coded by "Lengua" (tongue, or language) in column 2. A code of "Ing." means English, while a code of "Nat." means native.

In this case, the answer to the question posed is 2 (enter 2, then press the <Enter> or <Return> key on the keyboard).

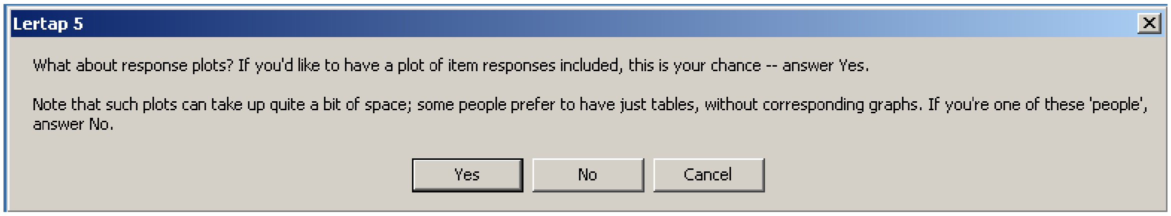


Ibreaks will step through all the distinct codes found in the Lengua column, presenting a question such as this one each time it encounters a new group code. In this example there are only two group codes, but often there will be more. Groups are included or excluded from the analysis according to the answer you give each time this questions appears.

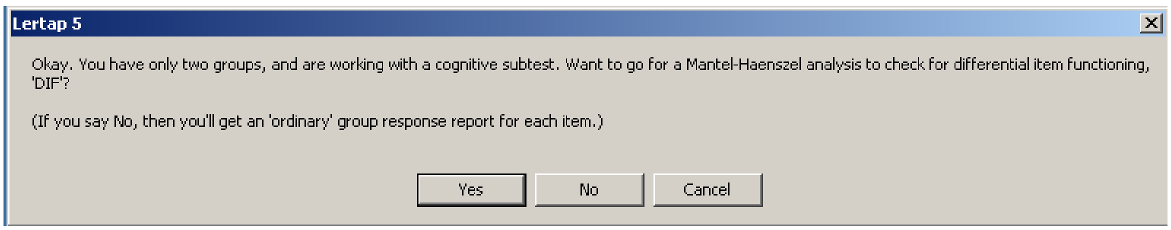
At this point, a click on the Yes button will make sure that the responses of the Ing. group are included in the analysis. Ibreaks then presents the same question, this time for the Nat. group. Another click on the Yes button, and Ibreaks moves along to its next question.

If the Scores worksheet contains more than one score, the "next question" is one which asks you to pick out the score you want to work with. The score you select may be from either a cognitive or an affective subtest.

Next:

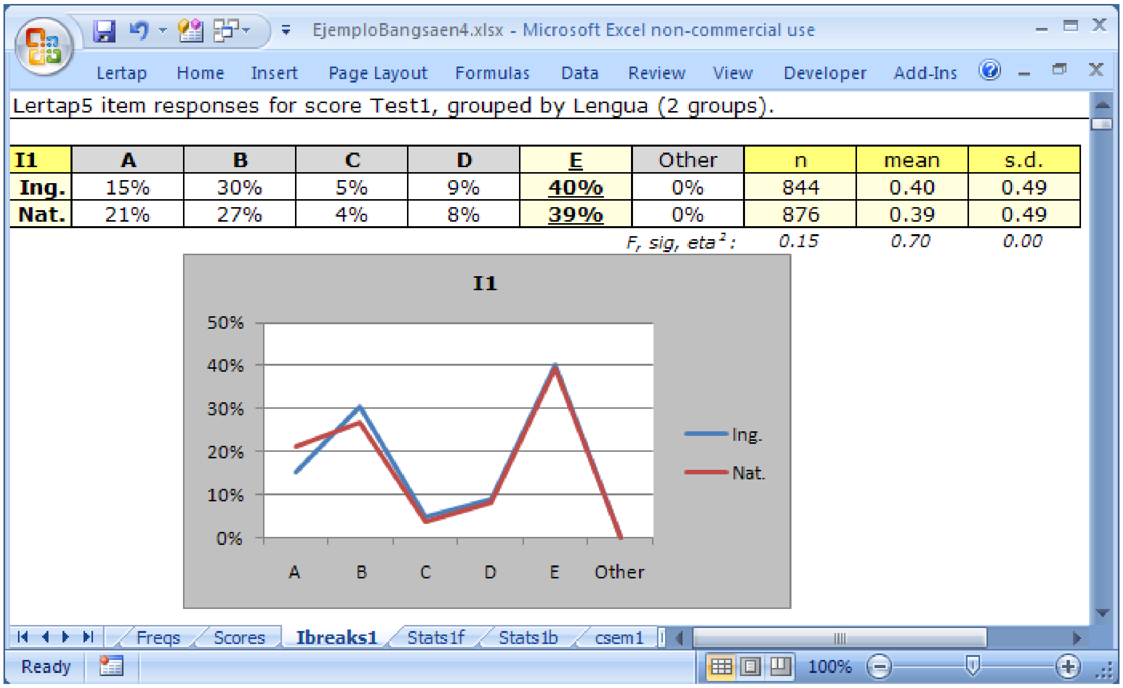


Over the years, and versions, Excel has at times placed a limit on the number of charts (or plots) which it will allow to be held in computer memory at any given time. This would be one possible reason to answer No to this question, but in this example a Yes answer is most appropriate so that you can see what the group response charts look like.



Here the answer will be No for the moment. This matter of "DIF" opens up a whole new topic, one which will be covered later.

Once you've answered its questions, Ibreaks gets down to work. It'll make a table of results for each item, and a response chart too, if you asked for charts.



I1	A	B	C	D	E	Other	n	mean	s.d.
Ing.	15%	30%	5%	9%	40%	0%	844	0.40	0.49
Nat.	21%	27%	4%	8%	39%	0%	876	0.39	0.49
						F, sig, eta²:	0.15	0.70	0.00

The correct answer to item I1 was E. Ibreaks denotes the correct answer by underlining, and by different shading.

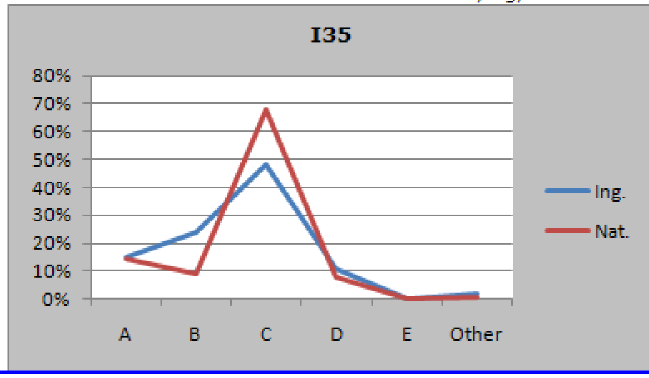
The table and the graph both indicate that the two groups had rather similar response patterns for this item, especially when it came to the correct answer.

The short row of statistics below the table provides quantitative measures of how dissimilar the groups were in their response to the correct option. Ibreaks gets Lertap to undertake an analysis of variance, resulting in an F ratio to provide a test of the null hypothesis: "The means of the two populations of test takers who have been represented by the Ing. and Nat. samples of this data set are equal". The statistical significance of F is given by "sig".

In this case, F is 0.15, and sig. is 0.70. Not statistically significant: we usually want sig. to be .05 or less *if* we're interested in the null hypothesis and statistical significance.

Any analysis of variance worth its salt will provide an index of practical significance as well as F and sig. Lertap uses the correlation ratio for this, referred to as "eta^2". It is thought, by some, that an eta^2 value of .10 or more may be interpreted as indicating that the difference in sample means is practically significant, possibly pointing to a difference we might refer to as "important", or "meaningful". (For more comments on these statistics, and on the matter of "significance", see the 'Analysis of variance table' discussion at the end of the Breakouts topic.)

I35	A	B	**C**	D	E	Other	n	mean	s.d.
Ing.	15%	24%	**48%**	11%	0%	2%	844	0.48	0.50
Nat.	14%	9%	**68%**	8%	0%	1%	876	0.68	0.47
						F, sig, eta²:	*71.39*	*0.00*	*0.04*



On this item, I35, 68% of the students in the Nat. group apparently knew the correct answer, compared to 48% in the Ing. group. The sig. value of F ratio is so small that, when rounded, it shows as *0.00*. The difference in item I35 means for the two groups is statistically significant. Eta², at 0.04, fails to reach great heights -- if we applied standard guidelines for measures of practical significance, we'd conclude that the difference in means is not what many would term "meaningful".

But here we would be almost foolish to think that there's no real difference. Analysis of variance methods are based, not surprisingly, on partitioning score variances. With cognitive items scored on a right / wrong basis, there may well not be much score variance to partition. There *is* a difference between the groups on their answers to I35; even "Blind Freddy" would be likely to see it. And it's a meaningful difference, too. An advantage of 20 "points" on an item difficulty scale *is* substantial.

It could be that the Nat. students are more capable, full stop. If so, the response differences seen in our little I35 plot might not surprise us at all. However, this wasn't the case. There was solid evidence to suggest that the Nat. and Ing. groups may have been equally proficient on the subject matter covered by the test. The stage is set, then, for a "DIF" analysis, differential item functioning. Jump to the DIF topic now, if you'd like to.

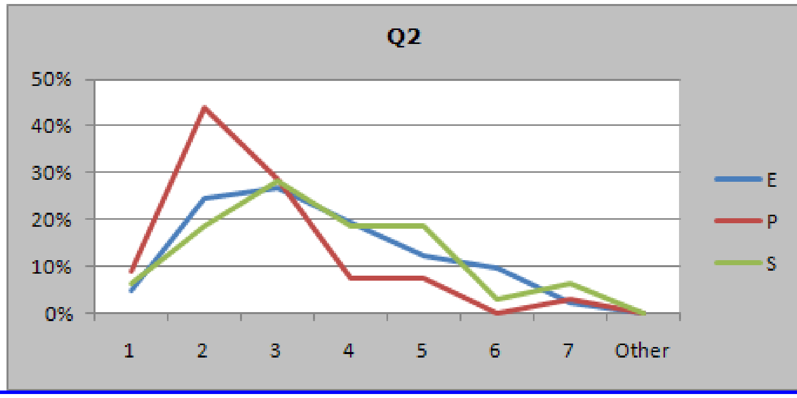
Colouring. The sig. value for I35 of 0.00 has been shaded above. Lertap's Ibreaks routine is trained to colour sig. values, and also eta² values, whenever they equal or fall below a cutoff minimum set in Lertap's System worksheet. Refer to the System settings topic, if you please.

4.4.8.1 Ibreaks charts

The charts, or plots, seen in the previous topic were based on right / wrong cognitive items, and only two groups.

The plot below is from the dataset used in the Breakouts topic. It used affective scales, many with items having seven response options. Three groups were involved.

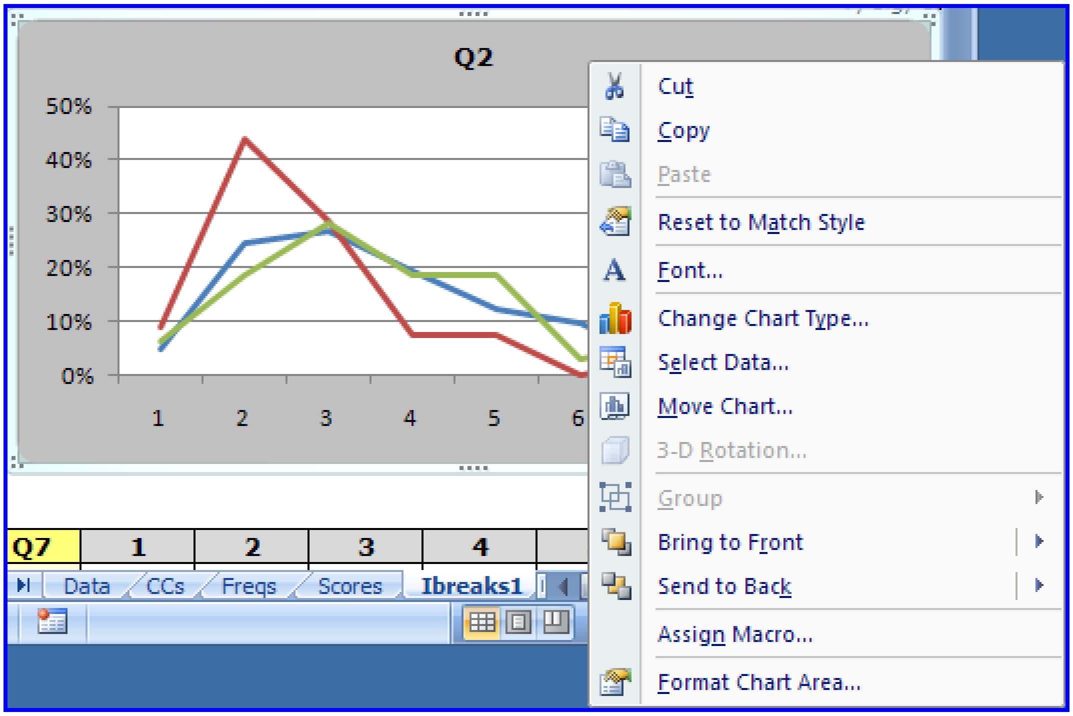
Q2	1	2	3	4	5	6	7	Other	n	mean	s.d.
E	5%	24%	27%	20%	12%	10%	2%	0%	41	4.51	1.47
P	9%	44%	29%	8%	8%	0%	3%	0%	66	5.27	1.26
S	6%	19%	28%	19%	19%	3%	6%	0%	32	4.41	1.52
								F, sig, eta²:	*5.80*	*0.00*	***0.08***



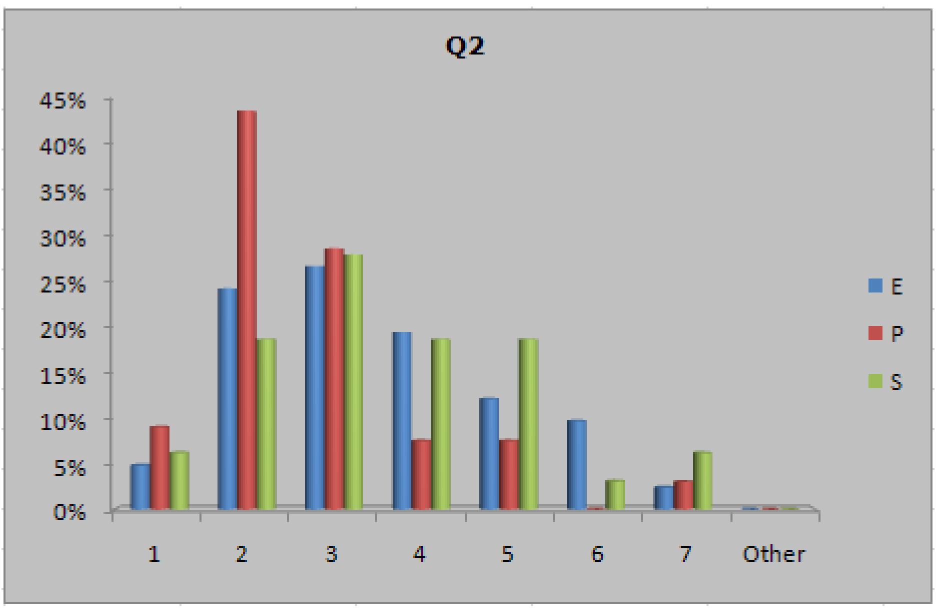
Q2 is an affective item from the 'MSLQ' instrument developed at the University of Michigan. MSLQ items typically ask students how often they had used a particular learning or study strategy during a given time period. For example, the item stem might be "*During the past week, how often did you meet with other students to go over material presented in lectures?*" Responses to some MSLQ items are gathered with a 7-point scale, ranging (for example) from 'Every day' to 'Never'.

The graph for Q2 seen above suggests that the 'P' (Primary) group made more use of the strategy than did students in the other groups.

It is easy to change Excel charts. Easy! Right-click on one, somewhere around the title (Q2 above):



Click on "Change Chart Type...", and say goodbye to the next two hours. Below is a rather crude chart change -- it's not difficult to make something which is more sophisticated.



Let Ibreaks create its usual response charts, which are, in Excel chart parlance, "Line" charts (but see this topic for an update: the default chart is now a "clustered columns" type). Then select any chart, and modify it to your heart's content. Save the result as a chart template. The, click on any other chart, and apply the template to it.

Or: use the *ChartChanger1* macro to modify the chart type in one go -- with this special macro, all you do is tizz up the first chart, apply the macro, and all following charts are tizzed. See the following topic.

Move 'em, Dan-O!

Alright, you may not be a fan of *Hawaii 5-O*, but here's another note, one which will be a bit more useful when we get into the DIF topic.

If you might want to move charts to a worksheet of their own, create a new worksheet, then use the "Move Chart..." option (seen above) to shift a selected chart to the new worksheet. Move one, two, or how ever many you want, and (in theory) they will line up well in the new worksheet.

Related tidbits:

One of our videos shows how to use many of the lbreaks features -- see Video (5) at this webpage.

There are many books on using Excel charts. See, for example, John Walkenbach (2007): *Excel 2007 Charts*, Wiley Publishing, ISBN 978-0-470-04400-1.

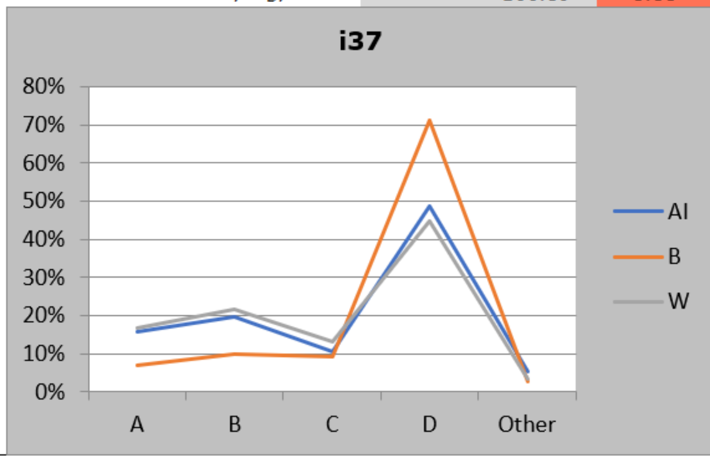
4.4.8.2 lbreaks chart type

The response plots displayed on the previous page use a chart type that's referred to in Excel as a "line" chart.

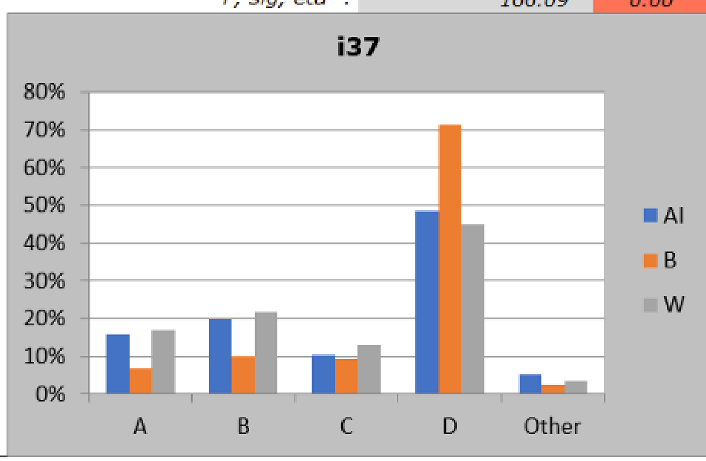
Versions of Lertap5 dated 26 August 2019 or later use "clustered columns" charts as their default chart type.

The two graphic images below display response patterns for item "i37" from a multiple-choice exam, "Exam1", used as an example in a text by Meyer (2014). Results from this exam are also discussed in this paper (page 15 or thereabout).

i37	A	B	C	D	Other	n	mean	s.d.
AI	16%	20%	11%	**49%**	5%	76	0.49	0.50
B	7%	10%	9%	**71%**	3%	1659	0.71	0.45
W	17%	22%	13%	**45%**	3%	3338	0.45	0.50
				F, sig, eta^2:		*166.09*	*0.00*	***0.06***



i37	A	B	C	D	Other	n	mean	s.d.
AI	16%	20%	11%	**49%**	5%	76	0.49	0.50
B	7%	10%	9%	**71%**	3%	1659	0.71	0.45
W	17%	22%	13%	**45%**	3%	3338	0.45	0.50
				F, sig, eta^2:		*166.09*	*0.00*	***0.06***



Of these two charts, the first (top) is referred to in Excel as a "line" chart, while the second (bottom) is called a "clustered columns" chart.

The default chart type used by Lertap5 is set in Row 56 of the System worksheet -- please refer to this topic.

The next topic mentions another way to get Excel change its chart type.

4.4.8.3 ChartChanger1

When you installed Lertap, several files were set up on your computer.

One of them is the Lertap5MacroSetA.xlam file. This file has some special-purpose 'macros', little collections of computer code designed to meet special needs. Read all about this with a clickity-click here.

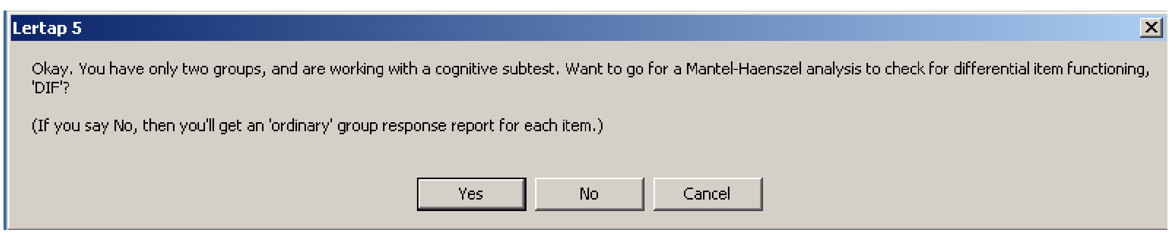
There's a fair chance your copy of the Lertap5MacroSetA.xlam file has our *'ChartChanger1'* macro in it. If it does, when you follow the instructions found in the clickity-click link above, you will see *ChartChanger1* listed as an available macro. If it doesn't, write to us at larry@lertap.com and complain -- we'll send out the right version of the file (it's free).

The *ChartChanger1* macro operates in this way: you make changes to the first chart seen on a worksheet, such as Ibreaks, and then run *ChartChanger1*. Bingo! All other charts on the worksheet are changed so that they match the type and style of the first chart.

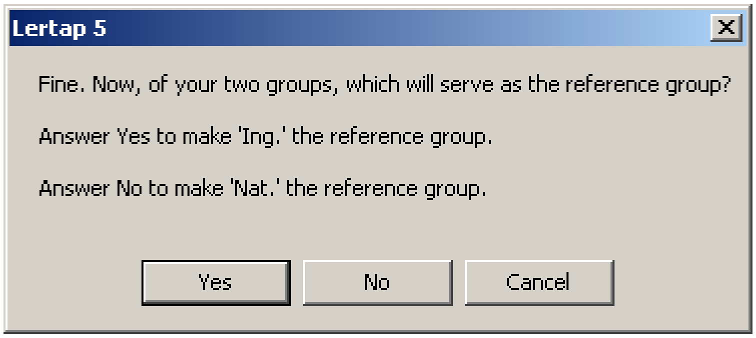
How to change the first chart? Select it by clicking on it. Then use Excel's Chart Tools, a suite of options which lets you change the design, layout, and format of a chart.

4.4.8.4 DIF

A couple of topics back, at "Ibreaks", an answer of No was entered when the following question appeared:



Now an answer of Yes will be entered, something which prompts the Ibreaks routine to ask one more question:



For this example, an authentic one to be sure, a professionally-developed 40-item achievement test had been used for years as an important assessment tool in science education. It had been presented in the country's native language, that is, the one most used in the general population. However, a push to promote the wider use of English, a strong, widely used second language, eventually resulted in high school science and mathematics instruction switching to English. Selected assessment instruments were carefully translated to English; cycles of forward and back translations were used to control the process, and, after trials stretching over three years, many of these instruments came into main-stream application.

In this example the Nat. group served as the reference group. The native-language version of the test was seen as the "gold standard"; some educators thought that the switch to English was disadvantageous -- the focus was on students sitting the English version of the test, and the question was: was it a fair test? Was there any evidence to suggest that the English version of the test worked against students, putting them at a disadvantage when compared to those who might still get to sit the native-language version of the test?

So it was that the answer to this Ibreaks question was No. This served to define the Nat. group as the reference group. The Ing. group became the "focal" group.

Once Ibreaks has an answer, it starts to churn out results, placing them in a report (worksheet) named "IbreaksMH1".

	1	2	3	4	5	6	7	8	9	10	11	12
1	Lertap5 Mantel-Haenszel results based on score levels from Test1, grouped by Lengua.											
2		*Score levels->*	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3		**Nat. (r)**	1	2	4	3	13	11	24	23	37	40
4		**Ing. (f)**	1	2	4	3	15	16	23	34	32	41
5	**I1**											
6		***Nat. diff***	*.00*	*.50*	*.00*	*.00*	*.15*	*.09*	*.17*	*.09*	*.11*	*.35*
7		***Ing. diff***	*.00*	*.00*	*.00*	*.00*	*.27*	*.31*	*.13*	*.15*	*.19*	*.22*
8		*odds ratio->*	.00	.00	.00	.00	.50	.22	1.33	.55	.53	1.91
9			*MH chi-sq:* .06		*Prob:* .81		*MH alpha:* .97		*MH D-DIF:* .08		*ETS level:* A (neg.)	

	1	2	29	30	31	32	33	34	35	36	37	38	39
1	Lertap5 Mantel-												
2		*Score levels->*	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	n total
3		**Nat. (r)**	26	23	11	12	13	6	4	0	0	1	876
4		**Ing. (f)**	16	16	9	14	7	13	2	3	2	1	844
5	**I1**												
6		**Nat. *diff***	*.65*	*.74*	*.91*	*.83*	*.85*	*.83*	*.75*	*.00*	*.00*	*1.00*	***.39***
7		**Ing. *diff***	*.69*	*.69*	*.67*	*1.00*	*.86*	*.85*	*1.00*	*1.00*	*1.00*	*1.00*	***.40***
8		*odds ratio->*	.86	1.29	5.00	.00	.92	.91	.00	.00	.00	.00	

When a DIF analysis has been requested, Lertap's Ibreaks routine creates two new reports, or worksheets. The snapshots above are from one of these, the "IbrakesMH1" report. MH stands for Mantel-Haenszel, the method Ibreaks uses for its DIF analysis.

The *Score levels* row starts at the lowest test score found, 4, and continues, in steps of 1 (one) to the highest score, which in this case was 39. These tables can become quite wide. The screen snapshots above have captured the first ten score levels, and the last ten.

Rows 3 and 4 give the number of students in each group at each score level. These numbers start to become relatively "substantial" at score level 8 (column 7), with 13 students in the (r)eference group, and 15 in the (f)ocal group.

The two *diff* rows give the proportion of students in each group who answered the item correctly. The *odds ratio* is a relative measure of how likely it is that a student in the reference group will get an item correct when compared to a student in the focal group. Greater than one, and the odds favour members of the reference group as being more likely to return the right answer. Less than one, and the focal group has the advantage. Equal to one, and it's even-steven.

The final column of the table gives the number of students in each group (876 and 844), and then, for each item, the proportion of correct responses for each group, over all score levels. Because this version of Lertap Ibreaks looks only at items which have been scored on a right / wrong basis, these proportions are equivalent to classical item difficulty figures. Thus, for item I1, the difficulty was .39 in the reference group, and .40 in the focal group.

Row 9 above has MH statistics for I1. *MH alpha* is the "common odds ratio", a figure derived by forming an average of the odds ratios over all score levels, weighted by the number in each group at each score level. In this example, an MH *alpha* of .97 indicates that the odds favour the focal group, but, since a value of 1.00 indicates equal odds, it's not a big favour.

MH chi-sq. is used to test the hypothesis that *MH alpha* is equal to one in the population from which the two samples of students have been drawn. *Prob:* is used to judge the statistical significance of *MH chi-sq*. We'd generally say that *MH chi-sq.* is statistically significant when *Prob:* is equal to or less than .05, and, this being the case, we would be tempted to say that *MH alpha* is, in fact, not equal to one; in turn, if this is indeed so, then we have evidence suggesting that group membership makes a difference: the chances of us observing a correct answer from a student might be said to depend on which group s/he is a member of.

Now, if it can be shown that the students in each group are of equal proficiency, or ability, then what might it be which would make it more likely for one of the groups to get an item correct? Perhaps a difference in the two versions of the item? Perhaps we have evidence of DIF, differential item functioning. This is often unwanted. If the objective is to have a "fair" test, test developers will generally weed out items which have such behavior (see tidbits references below; Angoff has a discussion of when DIF might be tolerated, even expected).

MH D-DIF is a statistic which results from converting *MH alpha* to a different scale: *MH D-DIF* = -2.35 times the natural logarithm of *MH alpha*. The D in D-DIF stands for delta. The delta metric is used by ETS, the Educational Testing Service, to express item difficulty.

Items with positive *MH D-DIF* favour the focal group, negative *MH D-DIF* favours the reference group.

The *ETS level* for an item will be A, B, or C. It will be A if *MH D-DIF* is between negative one and plus one (-1.00 < *MH D-DIF* < +1.00), or if *Prob.* is greater than .05 (*MH chi-sq.* is *not* statistically significant). A-level items are said to indicate negligible DIF; the IbreaksMH tables show this as A (neg.). *Note*: yes, it is possible to have a statistically-significant *MH chi-sq.*, but still have the item falling into the ETS A level if the magnitude of *MH D-Dif* is less than 1.00.

An item with substantial DIF, the C level on the ETS "scale", has an *MH D-DIF* value with a magnitude of at least 1.5 (that is, at or above 1.5, or at or below -1.5), with *MH D-DIF* significantly greater than 1.0 in magnitude. These items are denoted in IbreaksMH tables as *C (large)*. *Note*: not shown in the examples on this page is another statistic which appears to the right of the ETS level, "s.e.", the standard error of *MH D-DIF.* (A concise reference for the calculations used in this part of Lertap is Michaelides (2008); see Dorands & Kulick (2006) for a practical application of MH statistics and discussion of the ETS "scale" -- see "Related tidbits" below.)

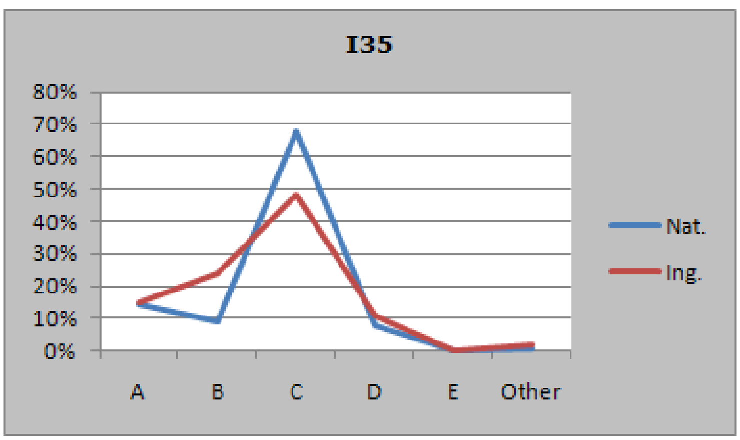
Bet you can't guess what an ETS B-level item is? It's one that is not in one of the other two levels. It exhibits some DIF, moderate DIF is the correct ETS term, and IbreaksMH tables show this as B (mod.).

1	Lertap5 Mantel-Haenszel results based on score levels from Test1, grouped by Lengua.										
2	*Score levels->*	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	**Nat. (r)**	1	2	4	3	13	11	24	23	37	40
4	**Ing. (f)**	1	2	4	3	15	16	23	34	32	41
209	**I35**										
210	***Nat. diff***	*1.00*	*.00*	*.50*	*.33*	*.08*	*.64*	*.38*	*.52*	*.49*	*.40*
211	***Ing. diff***	*.00*	*.00*	*.00*	*.33*	*.27*	*.38*	*.22*	*.12*	*.16*	*.44*
212	*odds ratio->*	.00	.00	.00	1.00	.23	2.92	2.16	8.18	5.12	.85
213		*MH chi-sq:* 73.82		*Prob:*	**.00**	*MH alpha:* 2.47		*MH D-DIF:* -2.12		*ETS level:*	***C (large)***
214											
215	**I36**										
216	***Nat. diff***	*1.00*	*.00*	*.25*	*.33*	*.54*	*.36*	*.54*	*.78*	*.59*	*.80*
217	***Ing. diff***	*.00*	*.00*	*.25*	*.33*	*.33*	*.38*	*.48*	*.47*	*.31*	*.59*
218	*odds ratio->*	.00	.00	1.00	1.00	2.33	.95	1.29	4.05	3.23	2.83
219		*MH chi-sq:* 10.37		*Prob:*	**.00**	*MH alpha:* 1.68		*MH D-DIF:* -1.21		*ETS level:*	**B (mod.)**

Here we have examples of two items which favour the reference group, Nat. The MH D-DIF figure for both of these items, I35 and I36, is high. Whenever MH D-DIF is outside of the range +1 to -1, it can be useful to get a picture which encapsulates some of the information in an IbreaksMH1 table -- the next topic has an example.

Plots. Above it says that asking for a DIF analysis gives "two new reports". IbreaksMH is one. The other is a version of the usual Ibreaks reports; this time its charts will come with DIF data included:

I35	**A**	**B**	**<u>C</u>**	**D**	**E**	Other	n	mean	s.d.
Nat.	14%	9%	**<u>68%</u>**	8%	0%	1%	876	0.68	0.47
Ing.	15%	24%	**<u>48%</u>**	11%	0%	2%	844	0.48	0.50
					*DIF sig, alpha, **delta**, ETS:*	*.00*	*2.47*	***-2.12***	*C (large)*
						F, sig, eta²:	*71.39*	*0.00*	*0.04*



The line of DIF stats has four fields imported from IbreaksMH tables: *Prob.* is here called *sig*, *MH alpha* is just called *alpha*, *MH D-DIF* becomes *delta*, and *ETS* is unaltered.

More plots. IbreaksMH tables given ample opportunity to make additional Excel charts. Examples are given in the next topic, Make M-H charts.

Related tidbits:

There's a paper on the Lertap website with more about DIF, especially as implemented in Ibreaks. Please see: http://www.lertap5.com/Documentation/GimmeABreak1.pdf.

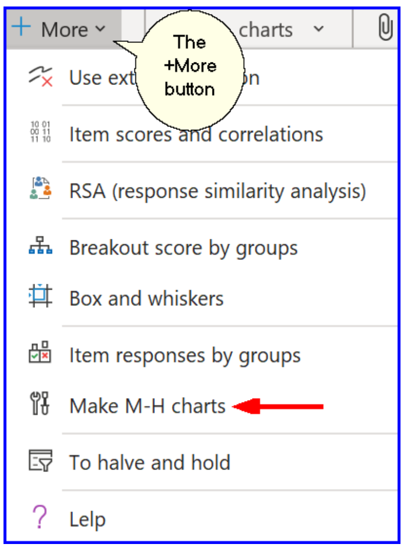
One of our videos shows how to use Lertap 5's DIF features -- see Video (5) at this webpage.

For DIF references, see Angoff (1993), Camilli & Shepard (1994), Clauser & Mazor (1998), Dorans & Holland (1993), Dorans & Kulick (1986), Dorans & Kulick (2006), Michaelides (2008), Zieky (2003), and Zwick (2012). (Refer to this list of references.) Note: DIF is sometimes also referred to as "item bias", now regarded by many as an outmoded term.

4.4.8.5 Make M-H charts

The "Make M-H charts" option on the Run menu is used to get "Lertap 5 DIF plots". It's used once an IbreaksMH1 report/worksheet has been made as discussed in the previous topic.

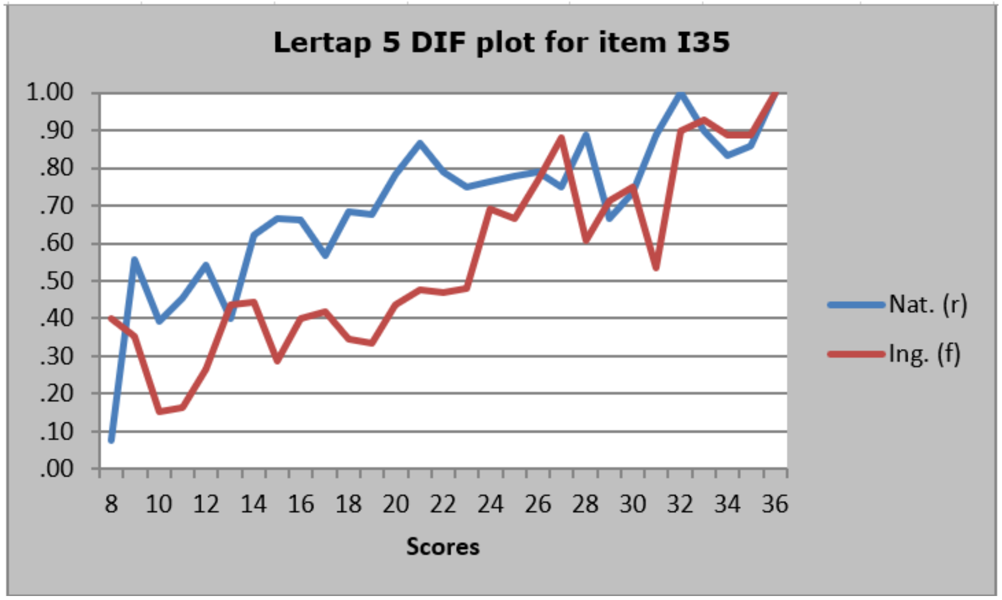
This screen snapshot is what will be seen when the "+More" drop-down option is clicked on:



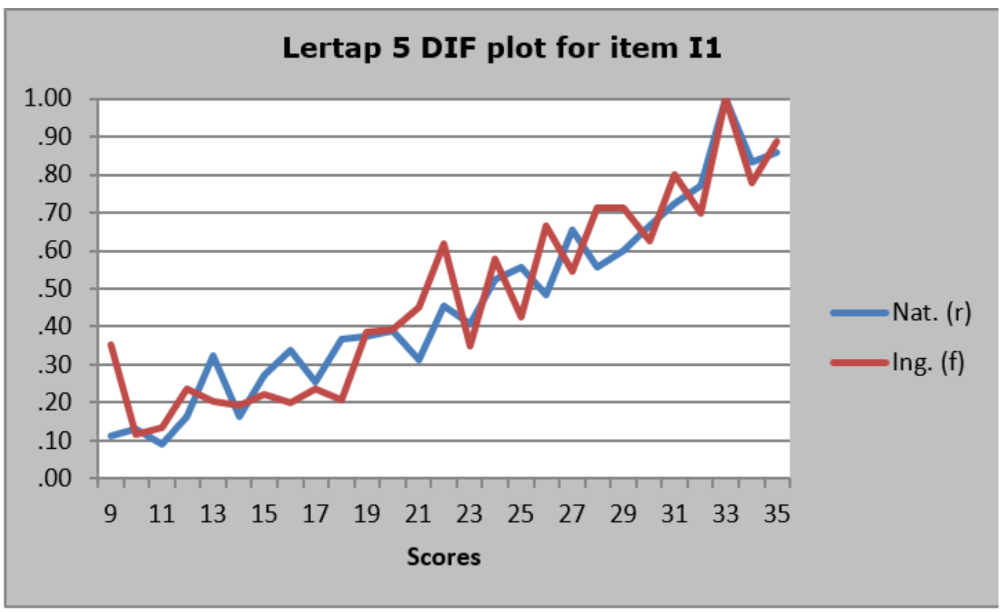
As this option starts up, it requests users to designate the score range for the DIF plots it be making. This is done using the "*Score levels ->*" seen in Row 2 of the IbreaksMH1 report. Very often there will be a small number of students at the extremes of the score levels, and the resultant DIF plots will be quite jagged at their edges. Using the score levels on the previous page as an example, one might pick column 7 (score level of 8) to start the plot, possibly ending with column 32 (score level of 33).

Once the score range has been designated, a copy is made of IbreaksMH1 so that users can use it (the copy) if they'd like to apply the option again using a different score range. The copy will be labeled IbreaksMH1(2).

A "DIF plot" for each test item will then be created, *similar* to the one shown here:



The chart above, for I35, reflects an item that meets the ETS criteria for *C (large)*. There's quite a range of scores where the Nat. group consistently out-performed the Ing. group. This is a picture which may well suggest the presence of 'DIF', differential item functioning.



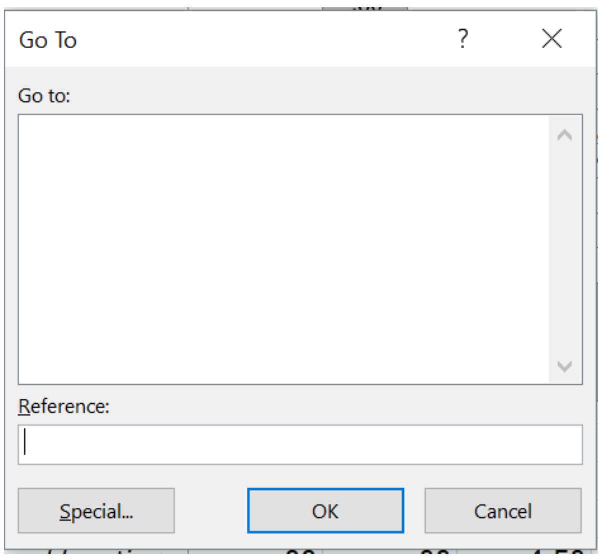
The chart above, for item I1, reflects no DIF, ETS level A (neg.).

The first time the option is used, the IbreaksMH1 worksheet will change - it will have DIF plots, similar to the two above. What was IbreaksMH1, the original worksheet without plots, becomes IbreaksMH1(2). IbreaksMH1(2) will not have plots.

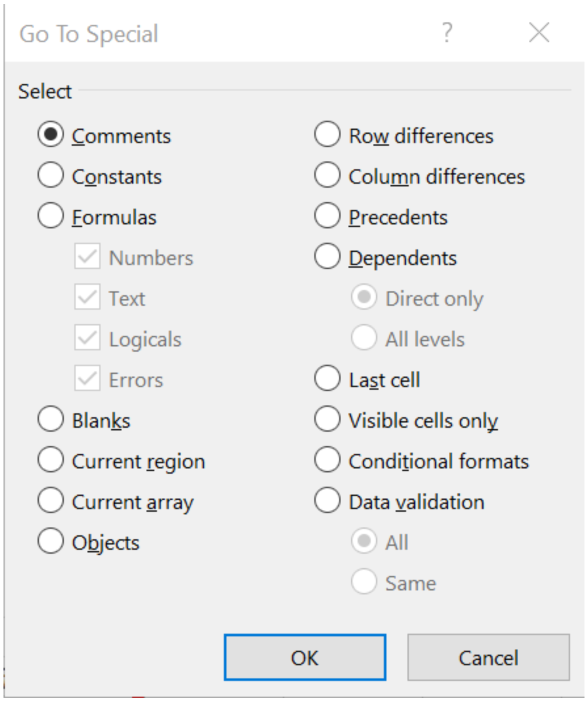
To plot over a new score range, the worksheet without plots, IbreaksMH1(2), would be selected and the option would be called on once again. As before, a range of scores would be entered and a new IbreaksMH1 sheet will result, with plots. The worksheet with the previous previous plots will then be in IbreaksMH1(2), with IbreaksMH1(3) now becoming the original IbreaksMH1 (no plots).

This process may continue; each time the option is taken it should be used with the worksheet that has no plots - this will always be the last of the workbook's IbreaksMH1(x) worksheets, where "x" is the highest number of all the IbreaksMH1() worksheets.

Note that there is a handy way to select all of the "Lertap 5 DIF plots" in a worksheet and carry them over to a Word document. While looking at an IbreaksMH worksheet, hold down the <Ctrl> key on the keyboard and at the same time press the G key. A "Go To" box such as the following should appear:



Click on "Special", the button seen at the lower left of the box.



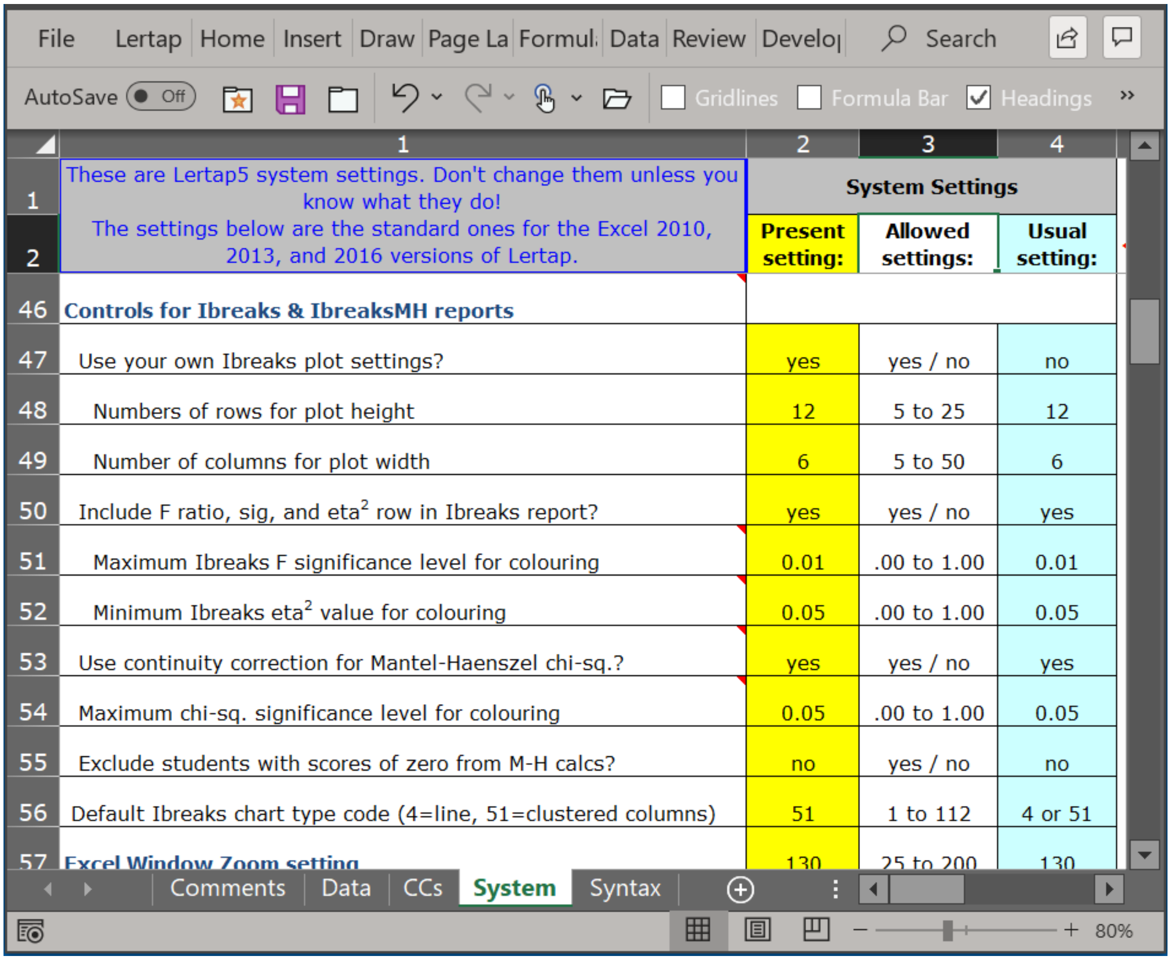
Select the "Objects" circle (also known as a "button"), click OK, and all of the plots (known more generally as 'charts' in Excel) will be copied to the clipboard. They may then be pasted into a Word document. If there are tens of plots, when they are pasted page breaks will be correctly observed so that any single plot will not split over two pages.

Related tidbits:

For other pictures of DIF, see, for example, Figure 4.4 in Camilli & Shepard (1994), and FIGs. 3.1 and 3.3 in Dorans & Holland (1993). (Refer to: http://www.lertap5.com/References.htm.)

4.4.8.6 System settings

Lertap's main options are set in the System worksheet of the Lertap5.xlsm workbook. The screen snapshot below displays some of the settings as seen in mid 2019.



The options which pertain to Ibreaks and IbreaksMH reports are seen in rows 47 to 56.

Ibreaks attempts to determine the optimal size for its charts. However, you can set your own by using the settings in rows 47, 48, and 49.

Rows 51 and 52 determine when Excel will use colours to highlight the *sig* and *eta*^{2} values in Ibreaks charts. Row 54 does likewise for the MH Chi sq *prob* cell in an IbreaksMH table.

Row 53 turns Yates' chi -square continuity correction on or off. In this snapshot it's set to No, which means Off.

Row 55's option is presently ineffective; Lertap doesn't use it. If scores of zero are encountered, they're converted to a score of one automatically. If this is done, Lertap will display an informative message.

Row 56's setting controls the type of chart that will be used as the default when Lertap5 produces Ibreaks summaries. Each Excel chart type has a corresponding numeric code. A list of the codes was available at this website in August 2019.

Printing

The January 2015 release of version 5.10.5 was the first to introduce System settings which provide for some control over page margins and the page breaks inserted whenever Excel's print options are used.

More about this is found towards the end of the pdf file available here (see the heading "Ibreaks charts have not been left out" on page 14).

4.4.9 To Halve and Hold

This option is used to create two random samples of data records, dividing a data set into halves on a random basis.

How does it do it? It begins by making a copy of the original Data and CCs worksheets, placing them in a new workbook. For convenience, assume that Excel calls this new workbook "Book1".

Then Halve&Hold uses two standard Excel functions to generate a set of random numbers between 1 and the number of data records in the original Data worksheet, denoted as "ArraySize" below:

```
Randomize
{... more code ...}
RandomValue = Int((ArraySize * Rnd) + 1)
{... more code ...}
```

The Randomize function provides a seed to Excel's Rnd routine. It uses the computer's clock to do this, guaranteeing that the random numbers generated will differ each time Halve&Hold is run.

Random numbers are generated until half of the original data records have been fingered (that is, identified). The unfingered records are then deleted from Book1's Data worksheet.

Then another copy of the original Data and CCs worksheets is made, and placed in a second new workbook, which we may call "Book2" for purposes of this discussion.

Next, the data records known to reside in Book1's Data worksheet are deleted from Book2's Data worksheet, and we end up with two essentially random samples of the original data, leaving the original untouched.

When the number of data records in the original Data worksheet is not an even number, Book1 will have one more data record in it than Book2.

How to generate a smaller random sample of data records? Halve&Hold always creates halves, workbooks whose Data worksheets have 50% of the records in the original Data worksheet. To get a sample with 25%, run Halve&Hold again, using one of the 50% samples -- for example, if Book1 contains 50% of the original Data records, run Halve&Hold with Book1 to get two new random samples, each with 25% of the original Data records.

Who uses Halve&Hold? Researchers and teachers, often people who are going on to undertake some sort of IRT analysis. At times one wants to have two samples of the original data; one of these might be used to calibrate an IRT model, with the second sample then used to validate the calibration.

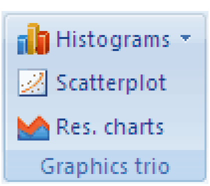
Teachers might use Halve&Hold to demonstrate sampling variance -- how do Lertap's scores and item statistics vary as we compare one of the samples with the other?

Time trials, September 2003, on a Pentium 4 running at 2 GHz: with 3,000 original records, the two halves were created in 18.8 seconds. With a bit over 11,000 original records, the two halves were ready in 4 minutes 18.4 seconds.

4.5 Graphics trio

There are numerous areas in Lertap where charts appear, or where you'll find an option to make a chart.

There are three chart types which are used so often that they warrant a section of their own on the Lertap tab, just to make it easier to get to them:



Go ahead and click on the options above, or simply page ahead to read about 'em.

4.5.1 Histogram

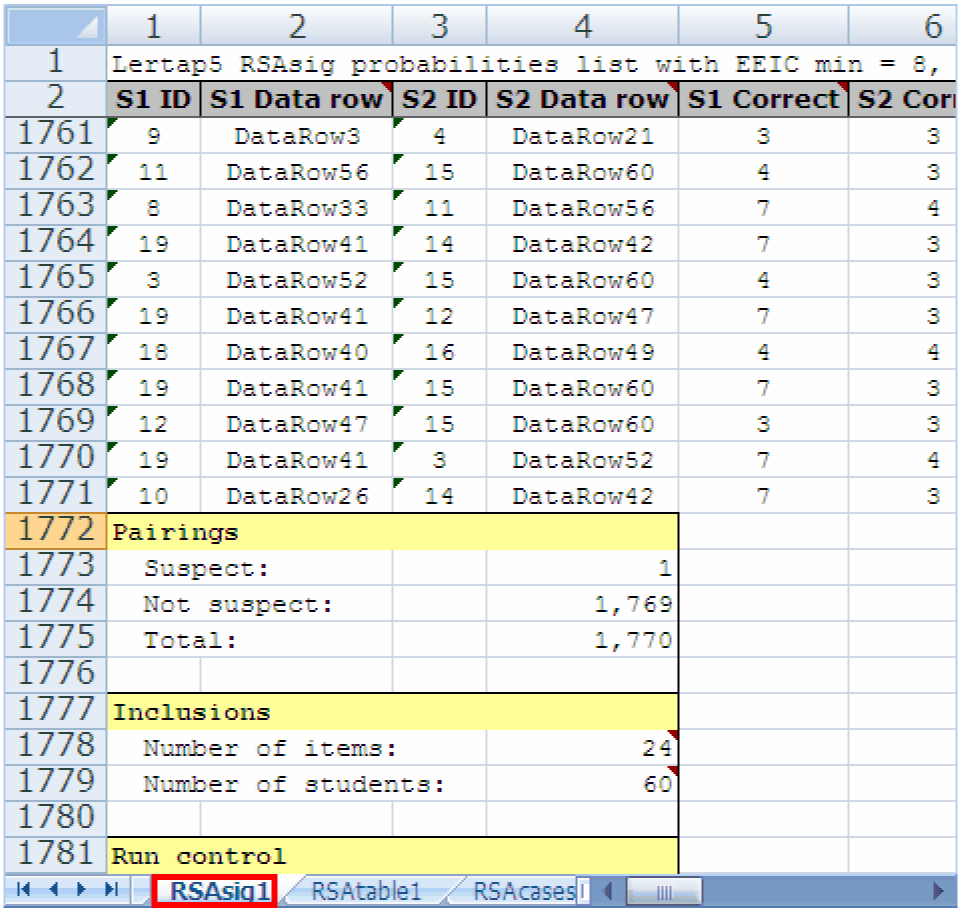
 Lertap's histograms are made from the columns of one of three reports worksheets: Scores, Breaks, and RSAsig. An exemplary snapshot of each of these worksheets, or 'reports', is showing here:

	1	2	3	4
1	Lertap5 Scores worksheet, last updated on: 27/07/2007			
2	**ID**	**Knwldge**	**Comfort**	**MDOCmfrt**
52	3	4.00	31.00	31.00
53	53	19.00	39.00	39.00
54	37	8.00	33.00	30.00
55	38	11.00	37.00	37.00
56	11	4.00	31.00	31.00
57	39	16.00	32.00	32.00
58	60	21.00	40.00	40.00
59	56	19.00	43.00	43.00
60	15	3.00	33.00	33.00
61	40	14.00	36.00	36.00
62	46	18.00	40.00	40.00
63	n	60	60	60
64	Min	1.00	26.00	26.00
65	Median	12.50	33.00	32.50
66	Mean	12.63	34.48	33.73
67	Max	24.00	43.00	43.00

Data | CCs | Freqs | **Scores**

	1	2	3
1	Lertap5 breakout of Knwldge score		
2	Knwldge	1	2
38	n	25	35
39	Min	1.00	3.00
40	Median	14.00	12.00
41	Mean	13.40	12.09
42	Max	23.00	24.00
43	s.d.	7.45	6.51
44	var.	55.52	42.36
45	Range	22.00	21.00
46	IQRange	15.00	9.50
47	Skewness	-0.33	0.31
48	Kurtosis	-1.52	-1.05
49	MinPos	0.00	0.00
50	MaxPos	25.00	25.00

Breaks1 | Stats1f | Stats1b



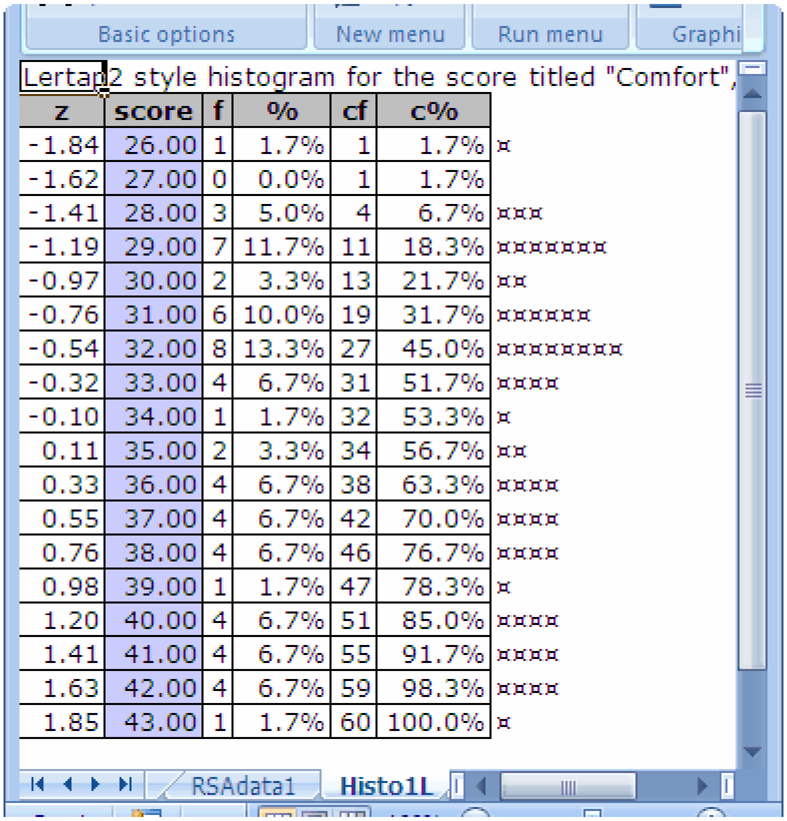
	1	2	3	4	5	6
1	Lertap5 RSAsig probabilities list with EEIC min = 8,					
2	S1 ID	S1 Data row	S2 ID	S2 Data row	S1 Correct	S2 Cor
1761	9	DataRow3	4	DataRow21	3	3
1762	11	DataRow56	15	DataRow60	4	3
1763	8	DataRow33	11	DataRow56	7	4
1764	19	DataRow41	14	DataRow42	7	3
1765	3	DataRow52	15	DataRow60	4	3
1766	19	DataRow41	12	DataRow47	7	3
1767	18	DataRow40	16	DataRow49	4	4
1768	19	DataRow41	15	DataRow60	7	3
1769	12	DataRow47	15	DataRow60	3	3
1770	19	DataRow41	3	DataRow52	7	4
1771	10	DataRow26	14	DataRow42	7	3
1772	Pairings					
1773	Suspect:			1		
1774	Not suspect:			1,769		
1775	Total:			1,770		
1776						
1777	Inclusions					
1778	Number of items:			24		
1779	Number of students:			60		
1780						
1781	Run control					

RSAsig1 RSAtable1 RSAcases

If you're looking at a Scores, Breaks, or RSAsig report, the histogrammer will swing into action immediately after you've clicked the histograms icon. In the case of Scores and Breaks, it will usually ask you to indicate which column you'd like to 'gram (the question is not required when you're on an RSAsig report as in that case there will be only one column, Log(PROB), to plot).

If you click on the histograms icon whilst viewing another type of report, such as Stats1b, for example, Lertap will take you to the Scores report, and wait for you to click on the icon again. If you didn't want to plot from Scores, you have the chance to click on Breaks or RSAsig instead, after which you have to click on the histograms icon yet again.

The 'histogram' created by the Histograms option is the standard Lertap 5 histogram, dating back to Lertap 2. Its style is exemplified below:



Basic options | New menu | Run menu | Graphi

Lertap2 style histogram for the score titled "Comfort",

z	score	f	%	cf	c%	
-1.84	26.00	1	1.7%	1	1.7%	¤
-1.62	27.00	0	0.0%	1	1.7%	
-1.41	28.00	3	5.0%	4	6.7%	¤¤¤
-1.19	29.00	7	11.7%	11	18.3%	¤¤¤¤¤¤¤
-0.97	30.00	2	3.3%	13	21.7%	¤¤
-0.76	31.00	6	10.0%	19	31.7%	¤¤¤¤¤¤
-0.54	32.00	8	13.3%	27	45.0%	¤¤¤¤¤¤¤¤
-0.32	33.00	4	6.7%	31	51.7%	¤¤¤¤
-0.10	34.00	1	1.7%	32	53.3%	¤
0.11	35.00	2	3.3%	34	56.7%	¤¤
0.33	36.00	4	6.7%	38	63.3%	¤¤¤¤
0.55	37.00	4	6.7%	42	70.0%	¤¤¤¤
0.76	38.00	4	6.7%	46	76.7%	¤¤¤¤
0.98	39.00	1	1.7%	47	78.3%	¤
1.20	40.00	4	6.7%	51	85.0%	¤¤¤¤
1.41	41.00	4	6.7%	55	91.7%	¤¤¤¤
1.63	42.00	4	6.7%	59	98.3%	¤¤¤¤
1.85	43.00	1	1.7%	60	100.0%	¤

RSAdata1 | Histo1L

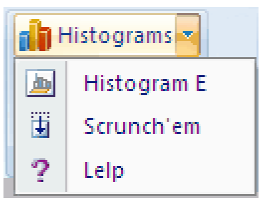
There may not be many columns in this "histogram", but there's a wealth of information:

z	The z-score corresponding to score. Found by subtracting the overall score mean from the score, and then dividing by the overall score standard deviation (s.d.). The overall mean and s.d. values are found in the Scores worksheet. In this case, the overall mean was 34.48, while s.d. was 4.61. Thus, for the first score of 26.00, z = (26.00-34.48)/4.61 = -1.84. z-scores will be zero when score equals the overall mean. (More z stuff here.)
score	These usually start at the lowest score found and end at the highest. However, when the range of scores to be processed exceeds 300, scores are collapsed into intervals and, in this case, the "score" displayed is the center of the interval.
f	The frequency of the score. In this case, there was one score of 26.00, no scores of 27.00, and so on. The most frequent score in this case (referred to as the "mode" in the literature), was 32.00, with f=8. When the range of

	scores is greater than 300, the value displayed for f is the number of scores in the interval mentioned immediately above, under "score".
%	The percentage corresponding to f, found by dividing f by the total number of scores (60 in this case), and multiplying by 100.
cf	The cumulative frequency, that is, the number of scores at and below score.
c%	The percentage corresponding to cf, found by dividing cf by the total number of scores (60 in this case), and multiplying by 100.

Histograms in this old format are saved in worksheets with names such as "Histo1L", "Histo2L", and so on -- the L is used to refer to the Lertap-2 style.

But wait. There's more to the Histograms option. A click on the arrowhead to the right will get a drop-down box to show, with more options. See:

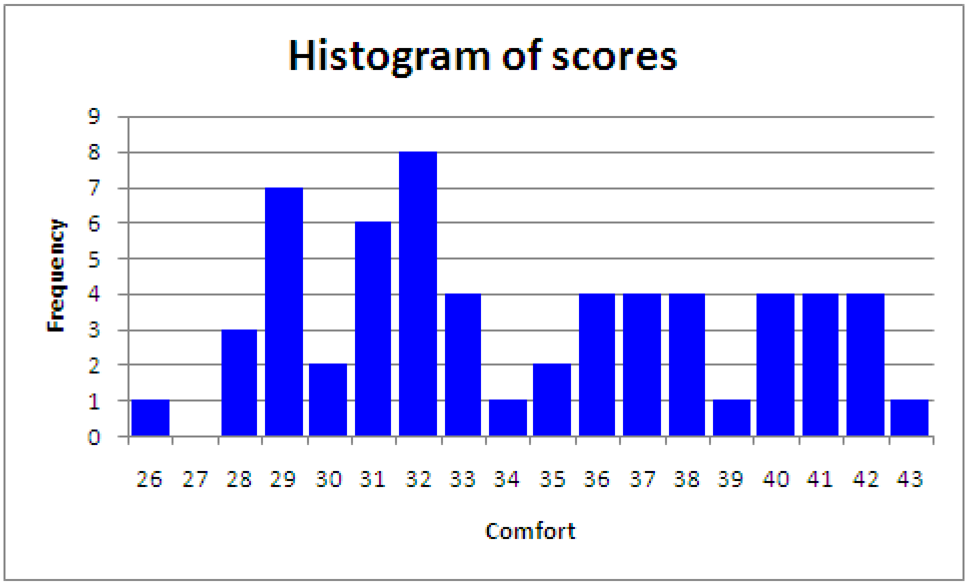


The Histogram E option will get Excel to look at an L-type histogram, such as found on a Histo1L worksheet, and make a fancier chart, one which may strike you as what folks traditionally expect to see in a histogram.

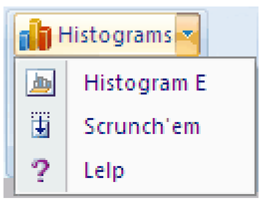
Note: some versions of Lertap automatically append the "fancier chart" to the output, making it unnecessary to use this option.

Go to a Histo1L or Histo2L (or so on) worksheet, and then click on Histogram E.

You should see a "chart" like this:



This is an Excel chart, of course, and as such you can do what you want with it. It's real easy to change titles, and colours, and to reformat the values seen along the x and y axes (x=Comfort and y=Frequency in this case).



Scrunch'em? Sounds like it could be painful, eh? Not so, read on ....

It is often the case that a HistoL histogram will have too many intervals. There'll be one interval for each possible score, starting with the minimum score and ranging up to the maximum score. If the minimum score came out to be 4, for example, and the maximum 100, there would be ninety-seven (97) intervals in the HistoL report. Oft times some of the intervals will have no entries; for example, if no-one got a test score of 13, there will nonetheless be an interval in the HistoL report for 13.

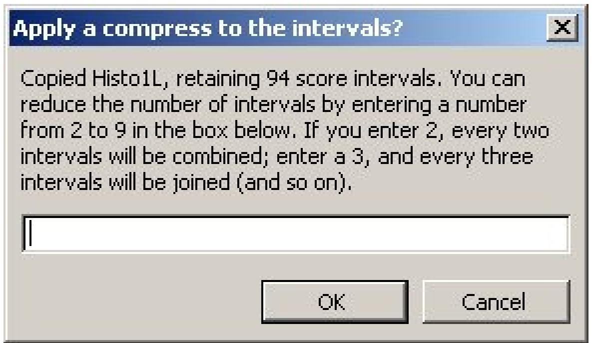
The "Scrunch'em" option lets you reduce the number of intervals in a HistoL-type report.

To use it, start by going to a HistoL report, such as, say, "Histo1L".

Then click on Scrunch'em.

Lertap will make a copy of the Histo1L report, calling the copy Histo1LCa. The "LCa" letters mean copy "a" of a compressed L-type histogram (if you use this option again with Histo1L, the next copy will be called Histo1LCb).

You'll then see a little dialog box such as the one seen below:



At this point you'll enter a compression factor, an integer, a number from 2 to 9. Let's see -- Lertap says that my Histo1L report presently has 94 intervals. Were I to make an Excel chart from the Histo1L report, using all 94 intervals, I know, from experience, that the chart would either be too large, or, if I've re-sized it to make it small enough to print on a standard page size, it'll be too "busy", too hard to read.

So I'll enter 4 in the box, and click on OK.

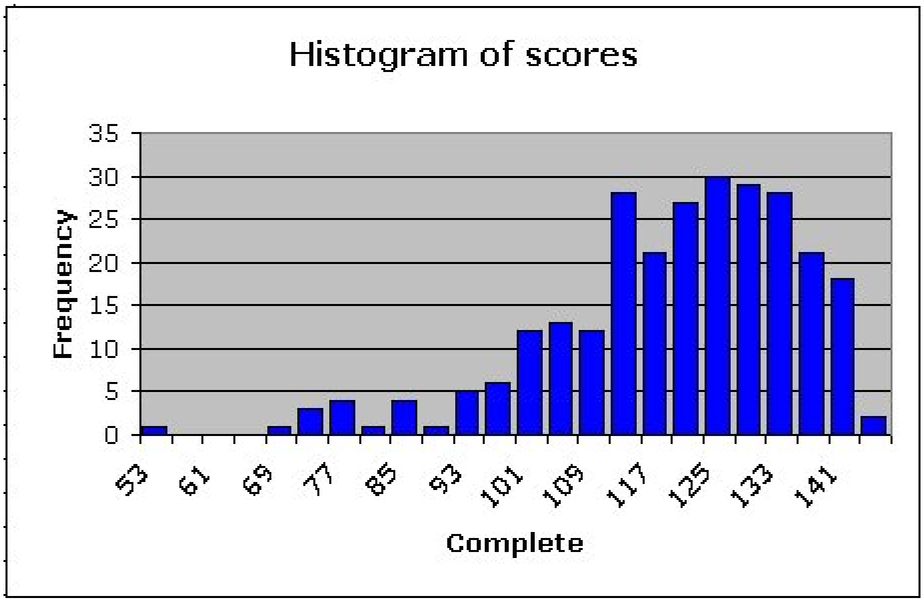
Before you can recite the birth dates of all the ancestors in your family, Lertap will come forth with a nifty table which looks something like this:

New Run Shorts

Lertap 5 HistoL compressed table for "C

z	interval	f	%	cf	c%
	52.50	1			
	56.50	0			
	60.50	0			
	64.50	0			
	68.50	1			
	72.50	3			
	76.50	4			
	80.50	1			
	84.50	4			
	88.50	1			
	92.50	5			
	96.50	6			
	100.50	12			
	104.50	13			
	108.50	12			
	112.50	28			
	116.50	21			
	120.50	27			
	124.50	30			
	128.50	29			
	132.50	28			
	136.50	21			
	140.50	18			
	143.50	2			

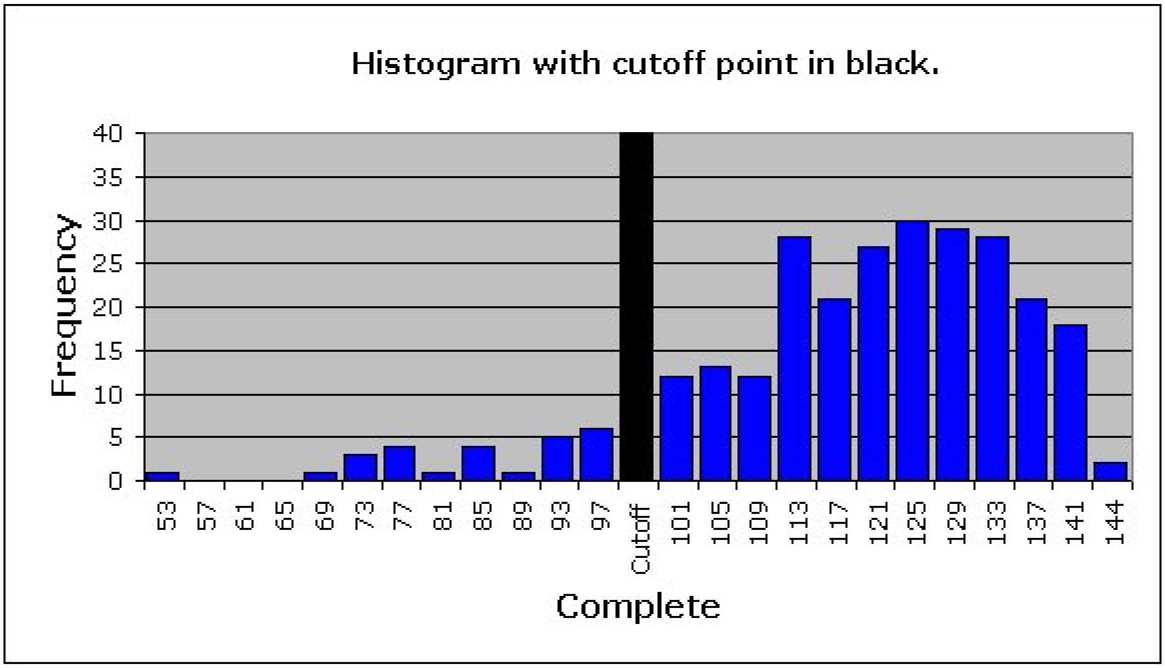
If your luck holds good, there will also be a spiffy little chart, a true wonder, ready for inclusion in that report you wanted to have ready for the school board meeting on Monday:



Now, you can do all sorts of things with both the table and the chart shown above.

If you change the entries in the table, the chart will change too, immediately. If you wanted, you could delete the chart, make some changes to the table, and then take the Histogram E option to get the chart back. For example, you could maybe insert a new row in the table to indicate a cutoff point, and then make the histogram chart again:

	1	2	3	4	5	6	7
1	Lertap 5 HistoL compressed table for "Complete".						
2	z	interval	f	%	cf	c%	
3		53	1				
4		57	0				
5		61	0				
6		65	0				
7		69	1				
8		73	3				
9		77	4				
10		81	1				
11		85	4				
12		89	1				
13		93	5				
14		97	6				
15		Cutoff	40				
16		101	12				
17		105	13				
18		109	12				
19		113	28				
20		117	21				
21		121	27				
22		125	30				
23		129	29				
24		133	28				
25		137	21				
26		141	18				
27		144	2				



Note that while you were still trying to remember all those birth dates, I got into the chart and fiddled around with it. I changed the title, and the font size of the values along the axes. How'd I get just one bar to be black? By clicking once or twice on the bar, until it was the only selected one. When it had obviously been selected, I then right-clicked on it, and made use of the options which Excel opened for me.

Complete?, you ask. Why is the title on the x-axis "Complete"? Because that's what this test score was called, all the way back in the respective CCs sheet:

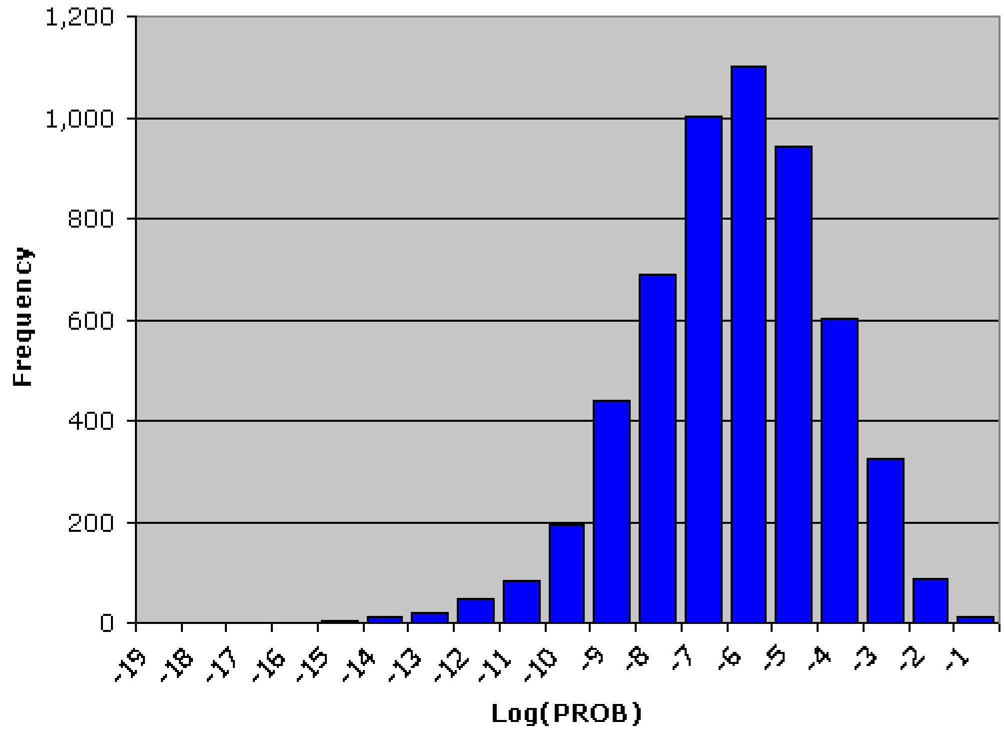
1	*col (c5-c154)
2	*sub name=(The complete test), title=(Complete), per

Note that the title of this test, Complete, appears in quotes at the top of the table, in the first row. You can change the title by typing over it, or, more simply, by just changing the title in the chart itself.

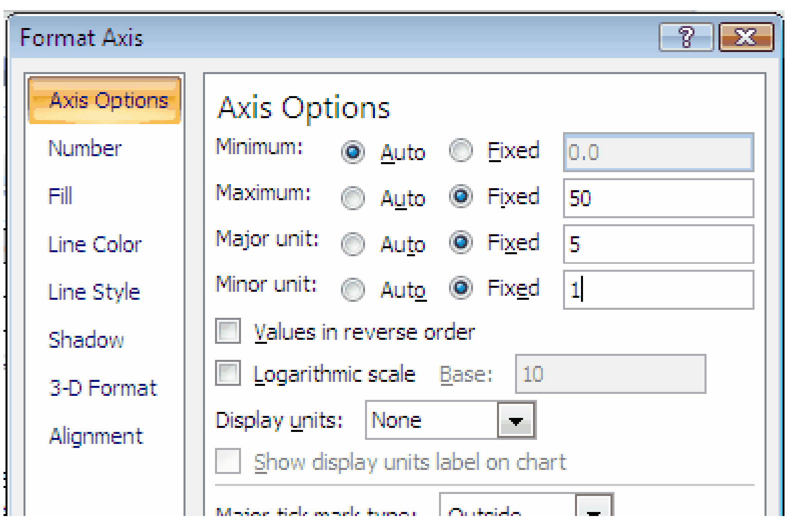
4.5.1.1 Scaling RSAsig

It was mentioned in the previous topic that the Histograms option will work with a Scores worksheet, with a Breaks report, and with an RSAsig report. In this topic you'll find some special comments on histogramming the results from an RSAsig worksheet.

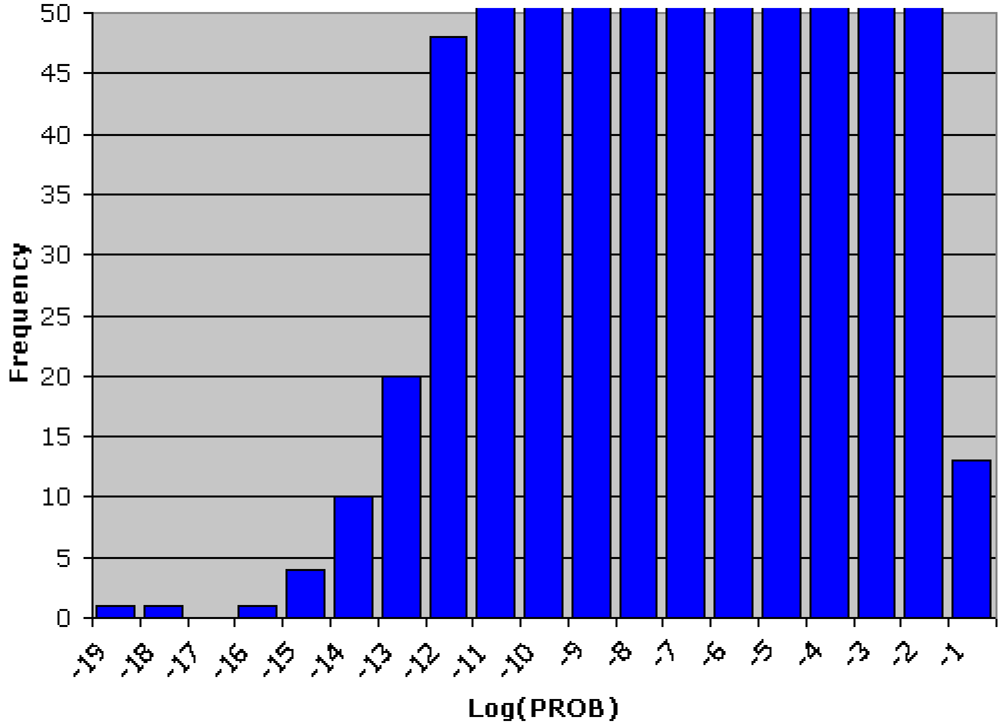
The chart below was made by first using the Histograms option when looking at an RSAsig report, followed by use of the Histogram E option:



It's often of considerable interest to have a more precise display of what's going on in the left tail of a Log(PROB) plot. To do so, right-click on one of values on the Frequency axis, and take the Format Axis option:



Above the Maximum has been set at 50 (the former value was 1200). Click OK, and: You called tails?:



Changing the scale of the (Y) axis has made the left-most Log(PROB) outliers visible. Note that these outliers are also sort of visible in the corresponding HistoL sheet:

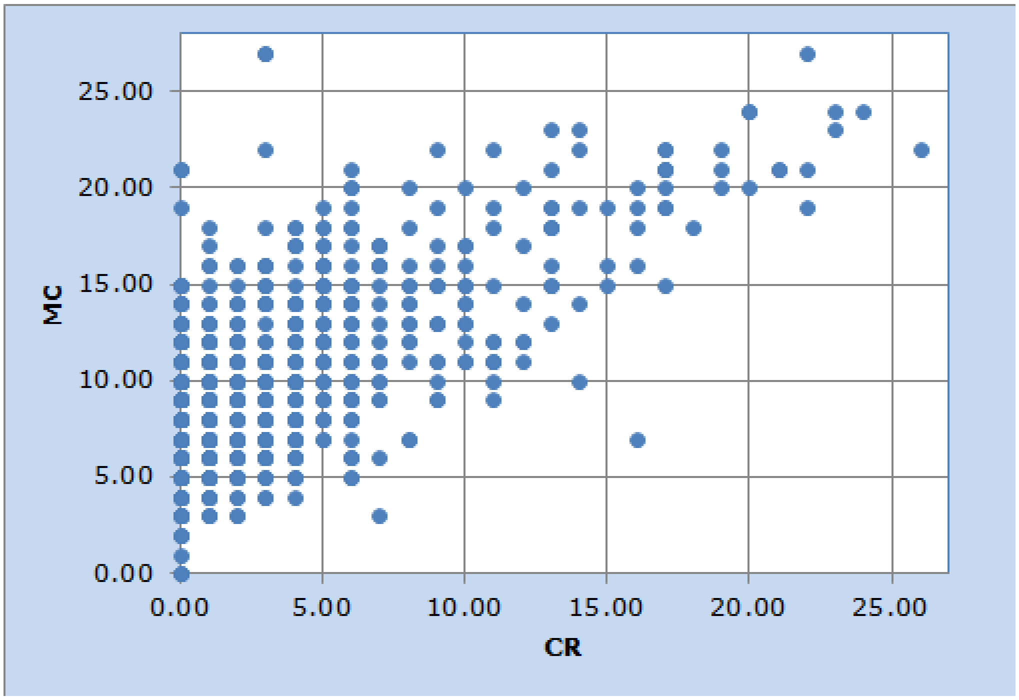
Distribution of "Log(PROB)", as at 25/01/2006. Eac

<table>
<tr><th>z</th><th>score</th><th>f</th><th>%</th><th>cf</th><th>c%</th><th></th></tr>
<tr><td>-6.12</td><td>-19.00</td><td>1</td><td>0.0%</td><td>1</td><td>0.0%</td><td></td></tr>
<tr><td>-5.63</td><td>-18.00</td><td>1</td><td>0.0%</td><td>2</td><td>0.0%</td><td></td></tr>
<tr><td>-5.15</td><td>-17.00</td><td>0</td><td>0.0%</td><td>2</td><td>0.0%</td><td></td></tr>
<tr><td>-4.67</td><td>-16.00</td><td>1</td><td>0.0%</td><td>3</td><td>0.1%</td><td></td></tr>
<tr><td>-4.18</td><td>-15.00</td><td>4</td><td>0.1%</td><td>7</td><td>0.1%</td><td>|</td></tr>
<tr><td>-3.70</td><td>-14.00</td><td>10</td><td>0.2%</td><td>17</td><td>0.3%</td><td>||</td></tr>
<tr><td>-3.22</td><td>-13.00</td><td>20</td><td>0.4%</td><td>37</td><td>0.7%</td><td>|||</td></tr>
<tr><td>-2.73</td><td>-12.00</td><td>48</td><td>0.9%</td><td>85</td><td>1.5%</td><td>|||||||</td></tr>
<tr><td>-2.25</td><td>-11.00</td><td>84</td><td>1.5%</td><td>169</td><td>3.0%</td><td>|||||||</td></tr>
<tr><td>-1.77</td><td>-10.00</td><td>195</td><td>3.5%</td><td>364</td><td>6.5%</td><td>|||||||</td></tr>
<tr><td>-1.28</td><td>-9.00</td><td>440</td><td>7.9%</td><td>804</td><td>14.5%</td><td>|||||||</td></tr>
<tr><td>-0.80</td><td>-8.00</td><td>689</td><td>12.4%</td><td>1,493</td><td>26.8%</td><td>|||||||</td></tr>
<tr><td>-0.32</td><td>-7.00</td><td>1,003</td><td>18.0%</td><td>2,496</td><td>44.9%</td><td>|||||||</td></tr>
<tr><td>0.17</td><td>-6.00</td><td>1,100</td><td>19.8%</td><td>3,596</td><td>64.6%</td><td>|||||||</td></tr>
<tr><td>0.65</td><td>-5.00</td><td>943</td><td>16.9%</td><td>4,539</td><td>81.6%</td><td>|||||||</td></tr>
<tr><td>1.13</td><td>-4.00</td><td>602</td><td>10.8%</td><td>5,141</td><td>92.4%</td><td>|||||||</td></tr>
<tr><td>1.62</td><td>-3.00</td><td>324</td><td>5.8%</td><td>5,465</td><td>98.2%</td><td>|||||||</td></tr>
<tr><td>2.10</td><td>-2.00</td><td>86</td><td>1.5%</td><td>5,551</td><td>99.8%</td><td>|||||||</td></tr>
<tr><td>2.58</td><td>-1.00</td><td>13</td><td>0.2%</td><td>5,564</td><td>100.0%</td><td>||</td></tr>
</table>

Look carefully at the f column above, and you'll spot the outliers below the score of -14. There aren't many, but the presence of just a few can influence the interpretation of an "RSA", that is, a response similarities analysis.

4.5.2 Scatterplot

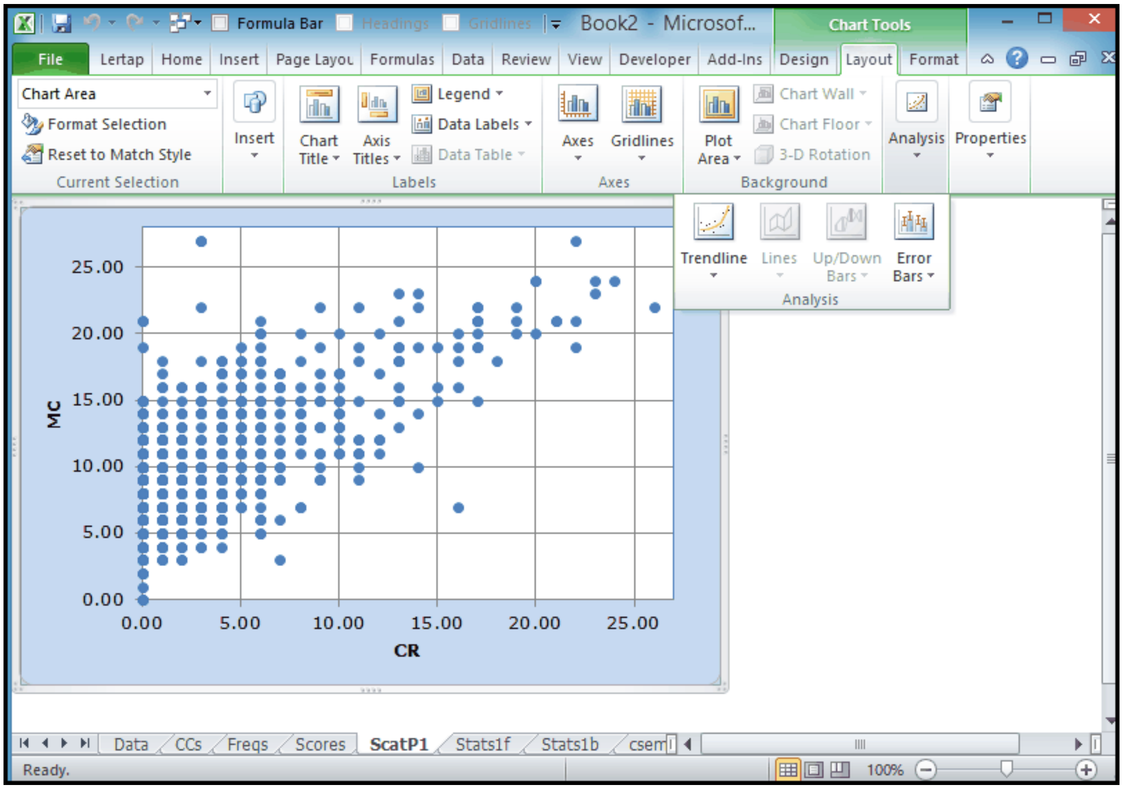
This icon makes it possible to compare any two columns from the Scores worksheet in a traditional scatterplot, also know as a scattergraph, or a scattergram. A typical plot is shown below, indicating the relationship between the multiple-choice and constructed-response sections of an exam taken by 950 high-school students.



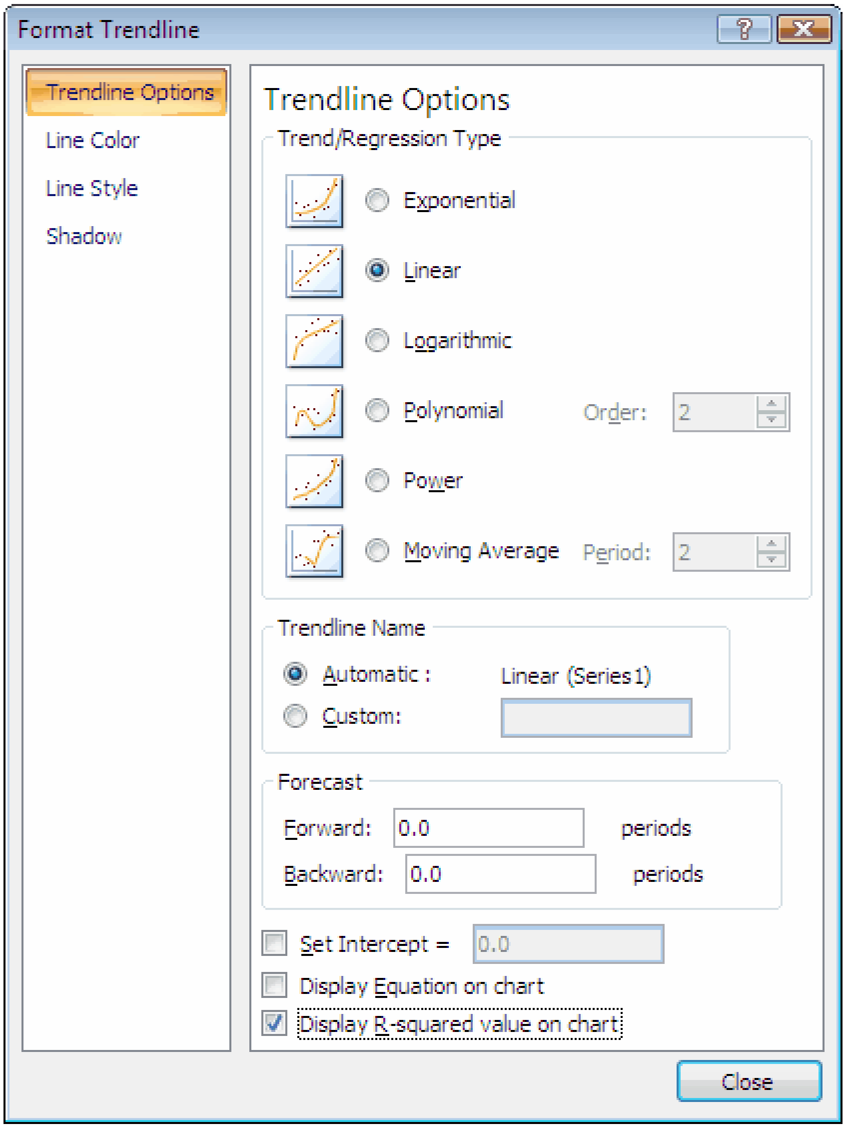
In this case, students found the constructed-response items to be difficult. More than 80% of the students had a score of 5 or less on the "CR" items (out of a possible maximum score of 29; the maximum possible on the "MC" items was also 29 in this example).

Lertap's scatterplots are Excel charts. They may be extensively reformatted. The axes may be expanded; the colours may be changed; even different symbols may be selected instead of the blue circles seen above. How? Double-click or right-click on various areas of the chart when viewing it in its worksheet. Or, look for help on "About charts" by using Excel Help.

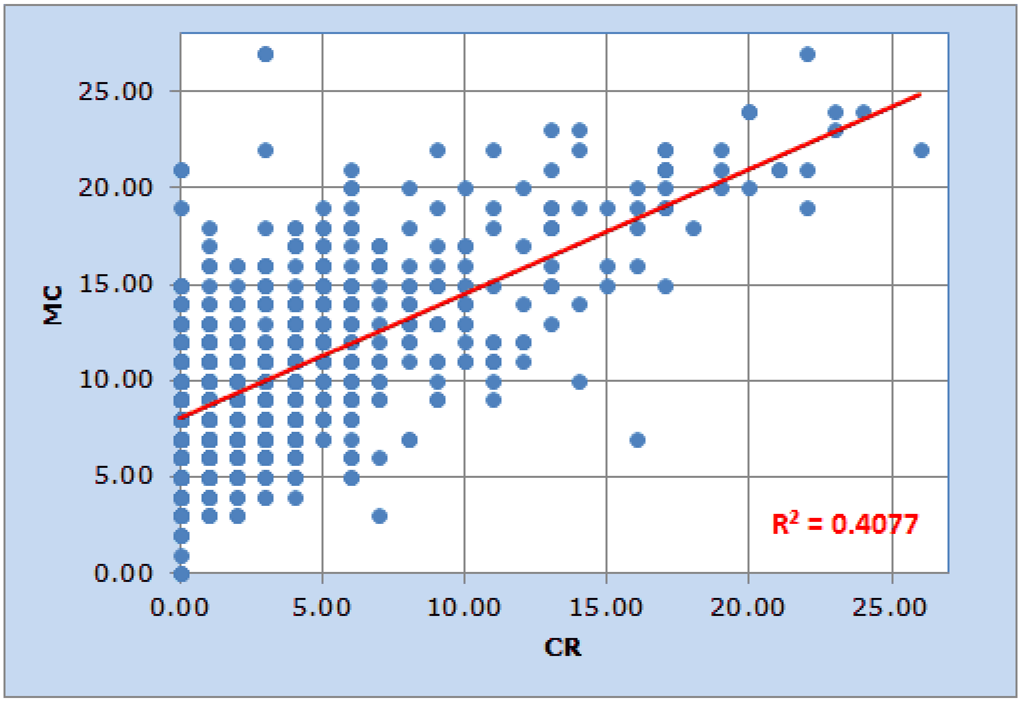
Excel also has some standard data analysis tools which may be quickly applied to scatterplots. How about trying to superimpose a linear trendline over the chart, and getting the value of R-squared to display? In Excel 2010, a click on the scatterplot will activate the "Chart Tools" options, as seen below:



In Excel 2010, the Trendline option is under the Analysis drop-down menu; clicking on it will reveal a box of options similar to this one:



Quicker than you can spread cold butter on cold toast you'll see the result:



(Note: some formatting was applied to the line and the R^2 text box in order to get the display seen here.)

Chapter 10 of the manual has a small section on Scatterplots.

4.5.3 Response charts

 Item response charts are made by clicking on this icon.

Two or three styles of charts are made, depending on the type of Lertap report worksheet active when the icon is clicked on.

How to make a worksheet "active"? Just get it to show -- this is best done by clicking on its tab (for an example, click here).

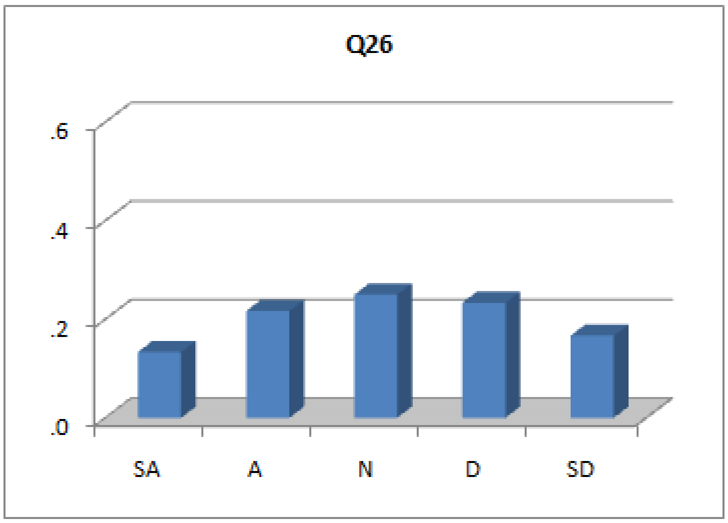
If a worksheet of the Stats1b style is active, then simple item response charts are made for all the items summarised in the Stats1b report (this applies to all reports of the "b" style, such as Stats2b, Stats3b, and so on). The items may be either cognitive or affective.

If a worksheet of the Stats1ul style is active, then quintile-style plots are made for all the items summarised in the Stats1ul report (this applies to all reports of the "ul" style, such as Stats2ul, Stats3ul, and so on). These 'quintile' plots can greatly assist with the process of identifying how well cognitive items perform.

Page forward to plot your future, topic by topic, or click here if you'd like to leap frog ahead to take in the special topic on chart "toggles".

4.5.3.1 Unidimensional response charts

 If you have a Stats1b-style report active, clicking on this icon will produce charts such as the one pictured below:



Lertap's item response charts are made by Excel. They may be copied and pasted to any other application, such as Word. They may also be extensively reformatted -- for help on this, please refer to Excel Help.

The response "labels" showing at the bottom of the chart are as found at the *top* of a Stats"b" worksheet. Here's a snapshot showing the top of a typical Stats"b" worksheet:

Lertap5 brief item stats for "Comfort with usin

Res =	1	2	3	4	5	oth
Q26	13%	22%	25%	23%	17%	

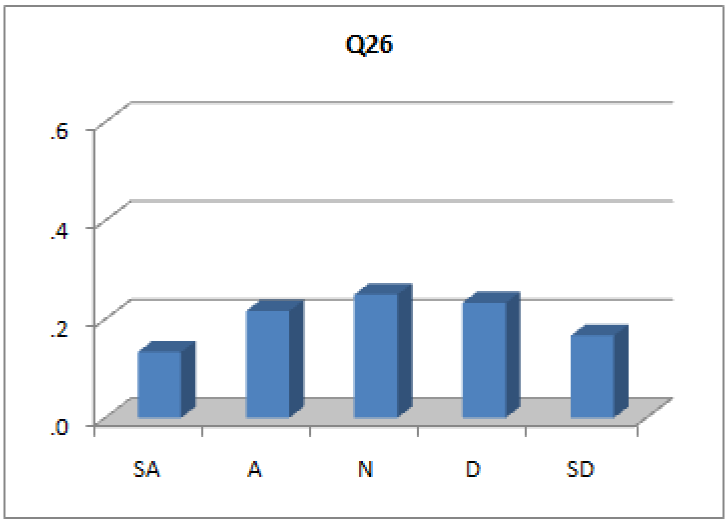
These response labels may be changed. If they are, the change will carry through to the item response charts.

For example:

Lertap5 brief item stats for "Comfort with usin

Res =	SA	A	N	D	SD	otl
Q26	13%	22%	25%	23%	17%	

Note how the original response labels of {1 2 3 4 5} have been changed to {SA A N D SD}? Look at what happens when item response charts are requested now:

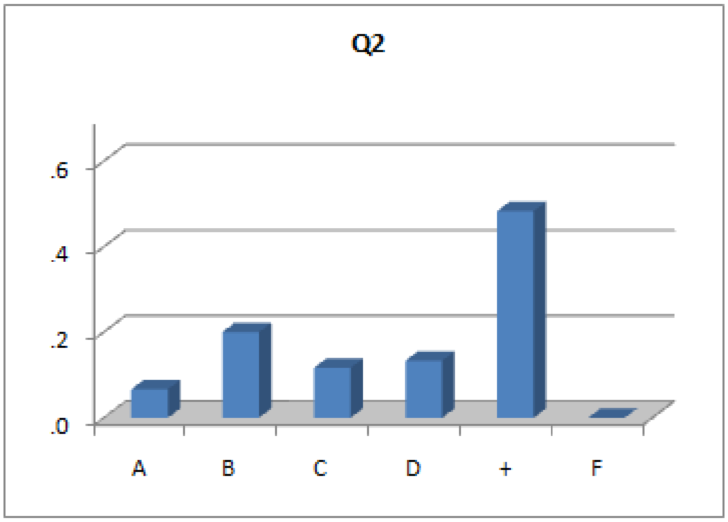


Lertap places item response charts in new worksheets, having names such as Stats1bCht and Stats2bCht.

Item response charts are briefly discussed in Chapter 10 of the manual (page 172 in the printed manual). An example suggesting how item response charts for different groups of test takers may be obtained is provided in Chapter 8 of the manual (pp. 130-133 in the printed manual).

Response charts of this type differ a bit, depending on the nature of the responses; if they are from an affective instrument with Likert-style items, the charts will be similar to the ones seen above. (Likert items typically use strongly agree to strongly disagree response options.)

For cognitive items, response charts will resemble the one shown below:

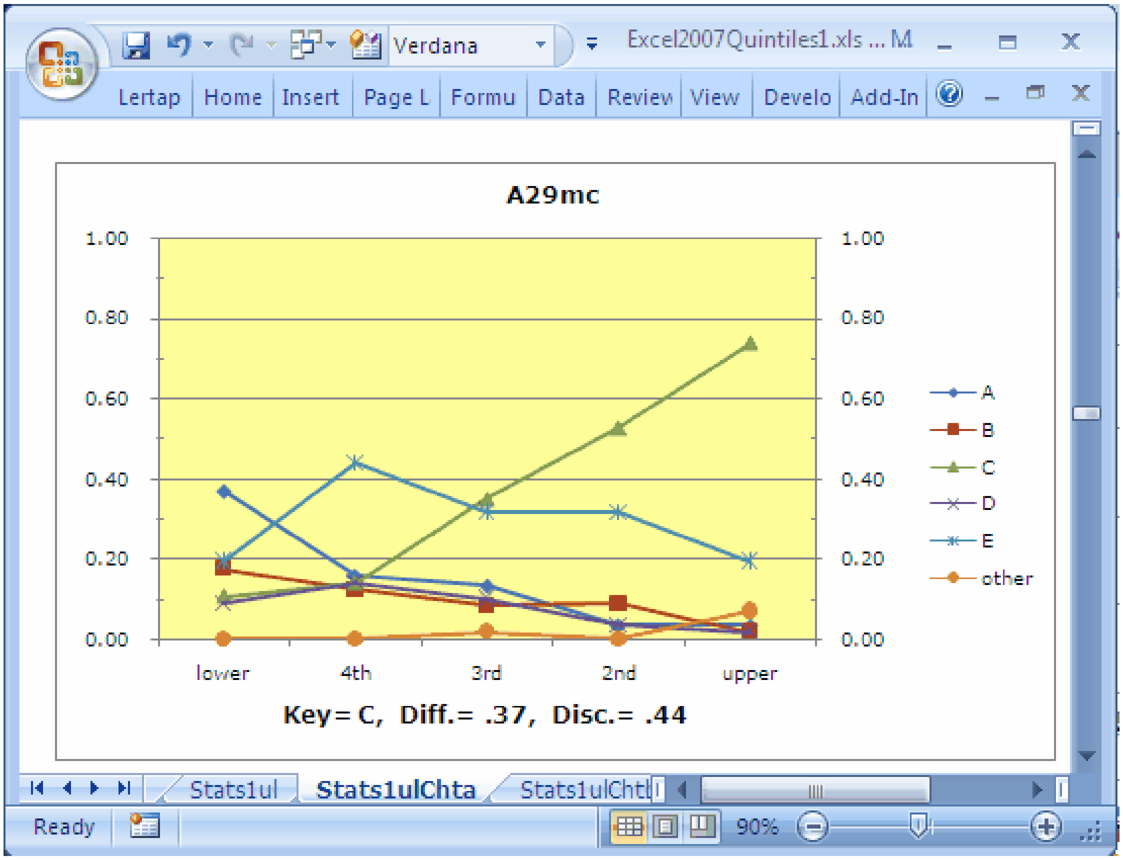


The correct answer to Q2, E, has been replaced by a + sign in the picture above (a definite plus, wouldn't you say?).

Note that problems can arise when creating these charts -- the number of charts which Excel can make is limited, as is the number of fonts which a workbook may have (the charts sometimes use a variety of fonts). These problems are more likely to appear when quintile plots are made from a Stats1ul report; the following topics provide more details.

4.5.3.2 Quintile plots

When a Stats1ul-style report is active, a click or two on this icon will produce charts such as this 'un:



Note: these plots took a big leap forward with the release of version 5.10.5 in January 2015. They're now a bit fancier, and, in particular, more complete as they can pull in information from Stats-f and ECStats-f summaries, resulting in a significantly more comprehensive summary of item performance. Read all about it here, and see numerous examples here.

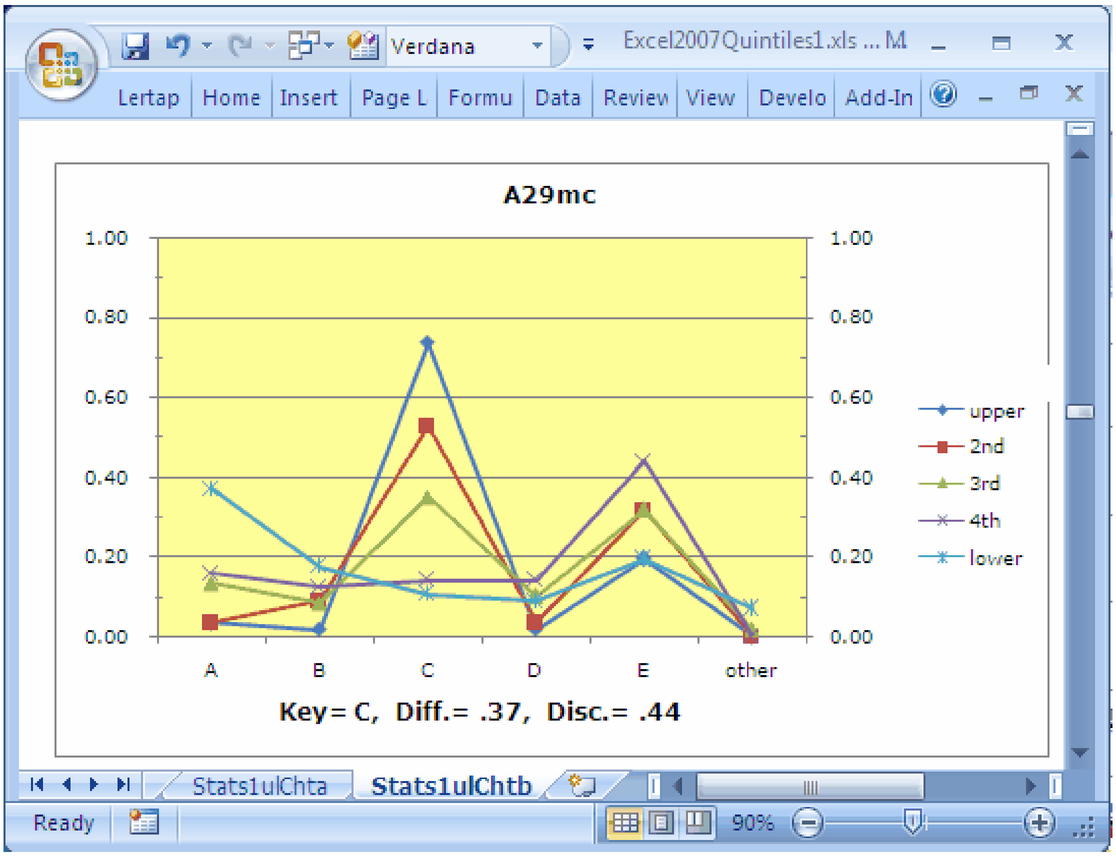
The 'quintile-a' plot pictured above shows how five item options, {A,B,C,D,E}, performed over five groups: 'lower' achievers to 'upper' achievers, with three intermediate levels of achievement. The variable used to index achievement may be internal or external; internal is the default, and is simply the total test score. The number of groups is usually five (hence the term "quintiles"), but may be less -- it's an option set in the System worksheet (see this topic).

The 'other' line generally represents respondents who omitted the item.

Wainer (1989) referred to plots similar to the one above as 'option trace lines'. If you haven't seen these before, have a careful look at their message: the green line (with small triangles on it) corresponds to the keyed-correct answer, C. The proportion of people in the lower group who selected this option is 0.11, a figure which steadily increases as we get into higher levels of achievement; by the time we get to the 'upper' group, the proportion able to pick out the correct option has increased to 0.74. The graph clearly shows an (almost) linear relationship between achievement and ability to identify the item's correct answer.

Distractors A and B foil fewer and fewer respondents as achievement level increases. Distractor E is quite popular with the lower achievement groups, falling away in the top (upper) group.

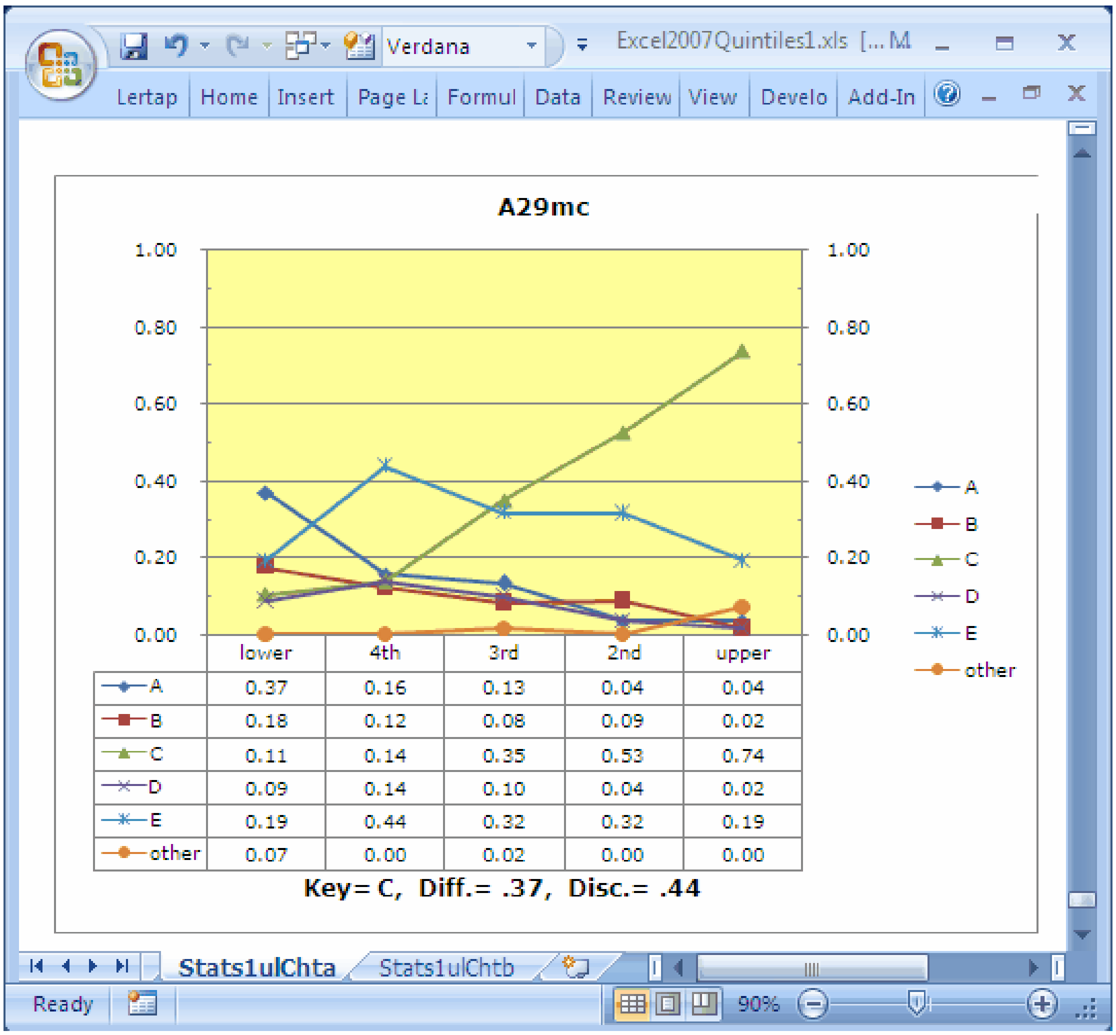
Okay? Great; now let's swap things around a bit, and look at the data from another angle:



This graph, a 'quintile-b' chart, plots achievement groups over response options. It shows, in the case of item A29mc, that the most popular response is in fact C, the correct answer. Distractor A nicely pulls off the three lowest achievement groups, especially the very bottom (lower) one. E is a fairly popular choice for the three middle achievement groups, with more than 40% of the 4th group (the next to lowest group) distracted by it.

Items which are good at discriminating the knows from the know-nots will have a particular quintile-b profile: the trace for the upper group, the "knows", will be the highest peak above the keyed-correct option, and the lowest peak for all the other options. The more it's the highest above the key, and the more it's the lowest everywhere else, the more the item discriminates.

Lertap's 'quintile plots' are of course just Excel charts. As such you can change them easily. A chart may be selected in a number of ways; once selected a great variety of chart options may be played with. Below we've asked Excel to include the actual data from which the 'quintile-a' chart was derived (Excel refers to the table below the graph as the "data table"):



These plots are handy. For some users, they may well be more informative indicators of item functioning than any other Lertap report or graph. How do you get 'em? Simple. First, make sure you've got the number of groups setting in the System worksheet right. You can have from 2 to 10 groups.

Then, make Stats1ul (or Stats2ul, ...) the active worksheet by clicking on it. Next just mouse up to the Lertap toolbar, and click on .

Lertap will click into action, opening up a new worksheet called either Stats1ulChta, or Stats1ulChtb, depending on a setting in row14 of the System worksheet. The default action: the first time you click on the icon, 'quintile-a' plots are made for each item, and placed in Stats1ulChta. The next click creates 'quintile-b' plots, recorded in the Stats1ulChtb worksheet. You can reverse the order via that setting in the System worksheet. To use different colours for the trace lines, see the Chart colors topic.

The Diff and Disc figures seen in the plots are taken from the "b" report which corresponds to the subtest involved. If you're clicking off of a Stats1ul report to get your plots, the Diff and Disc values are taken from lines in the Stats1b report. Note that the Stats1ul report has its very own diff and disc values; these are conventional upper-lower indices; in a way they're inferior to the values found in Stats1b in that they're just based on results from two groups -- the lowers and the uppers -- whereas the corresponding values in Stats1b are based on all test takers. (The Disc figure in the Stats1b report is a corrected point-biserial correlation coefficient.)

Note: items with negative discrimination ("disc.") are flagged in the plots by the use of red coloring. The item's label will be red, and the statistics below the chart will also be red. An example may be seen by clicking here.

SAQ: *If I change the number of groups, maybe even to just two, can I still get 'quintile plots', even though I no longer have true achievement quintiles?* Sure.

SAQ: *And, can I get these marvellous graphs even when my Stats1b report is based on a mastery test?* Yes, go for it (Master)!

Related tidbits:

These charts changed substantially when version 5.10.5 was released in 2015. It's now possible to incorporate information from Stats-f and ECStats-f summaries. This document has the scoop you'll want to see; note that the "Using Lertap Quintile Plots" PDF document referenced in the next sentence has many examples of the new charts.

There's a paper, a best seller, which has a lot more to do about using quintile plots: UsingLertapQuintilePlots.pdf (pdf file, about 400 KB).

How to print quintile plots is discussed in this topic.

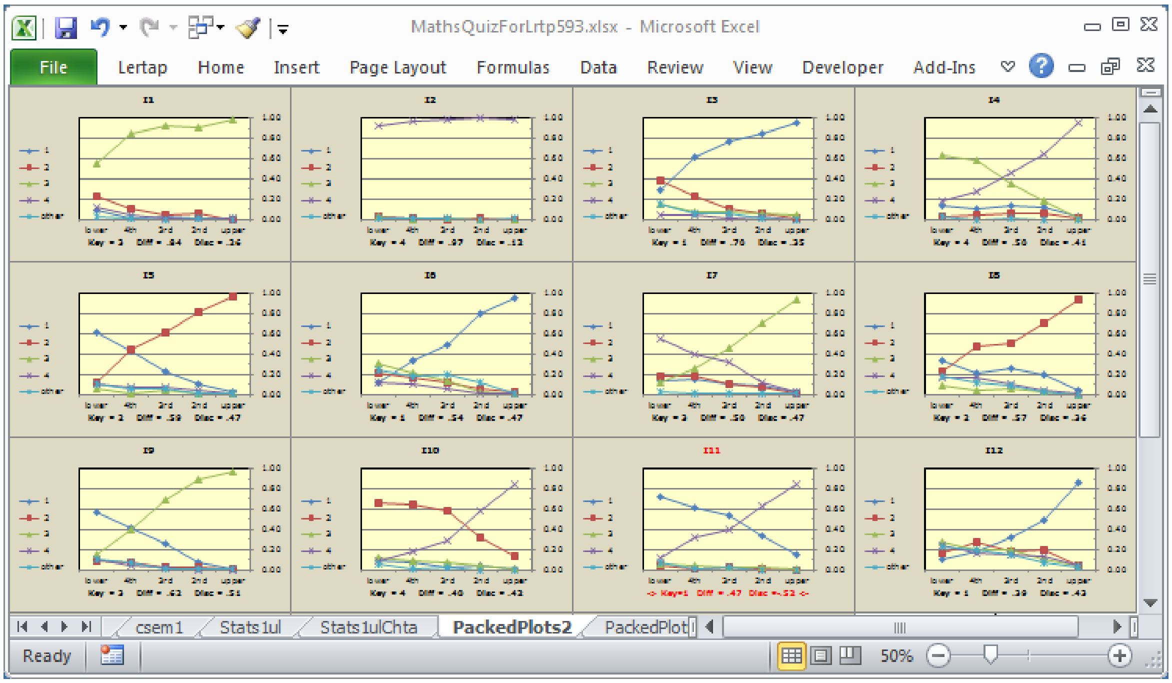
4.5.3.2.1 ChartChangers

It's possible to make quite a number of changes to Lertap's quantile plots, or "charts".

An oft-used built-in option is a switch which turns on, or turns off, the data table from which the trace lines in the charts are derived. How to flick this switch? See the quintile options topic.

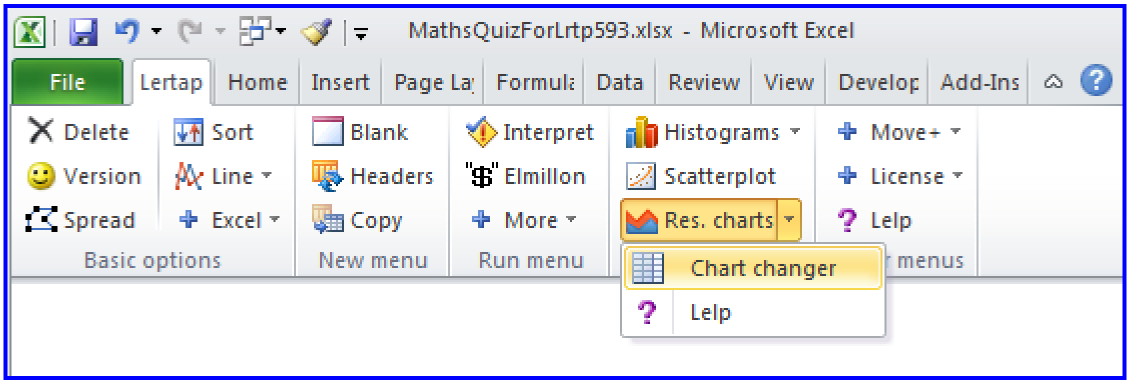
Packing the plots

The screen snapshot below indicates how "packing" the plots may lead to a useful gestalt, a condensed overview of item performance. This snapshot displays plots for the first twelve items from the MathsQuiz dataset:



Note: items with negative discrimination ("disc.") are flagged in the plots by the use of red coloring. The item's label will be red, and the statistics below the chart will also be red. I11 is an example above. The "Disc." values in the charts are brought in from Statsb reports.

If you have Excel 2010, plot packing may very readily be accomplished with an in-built macro called ChartChanger3. It's activated by clicking on the small arrowhead next to the Res. charts option:

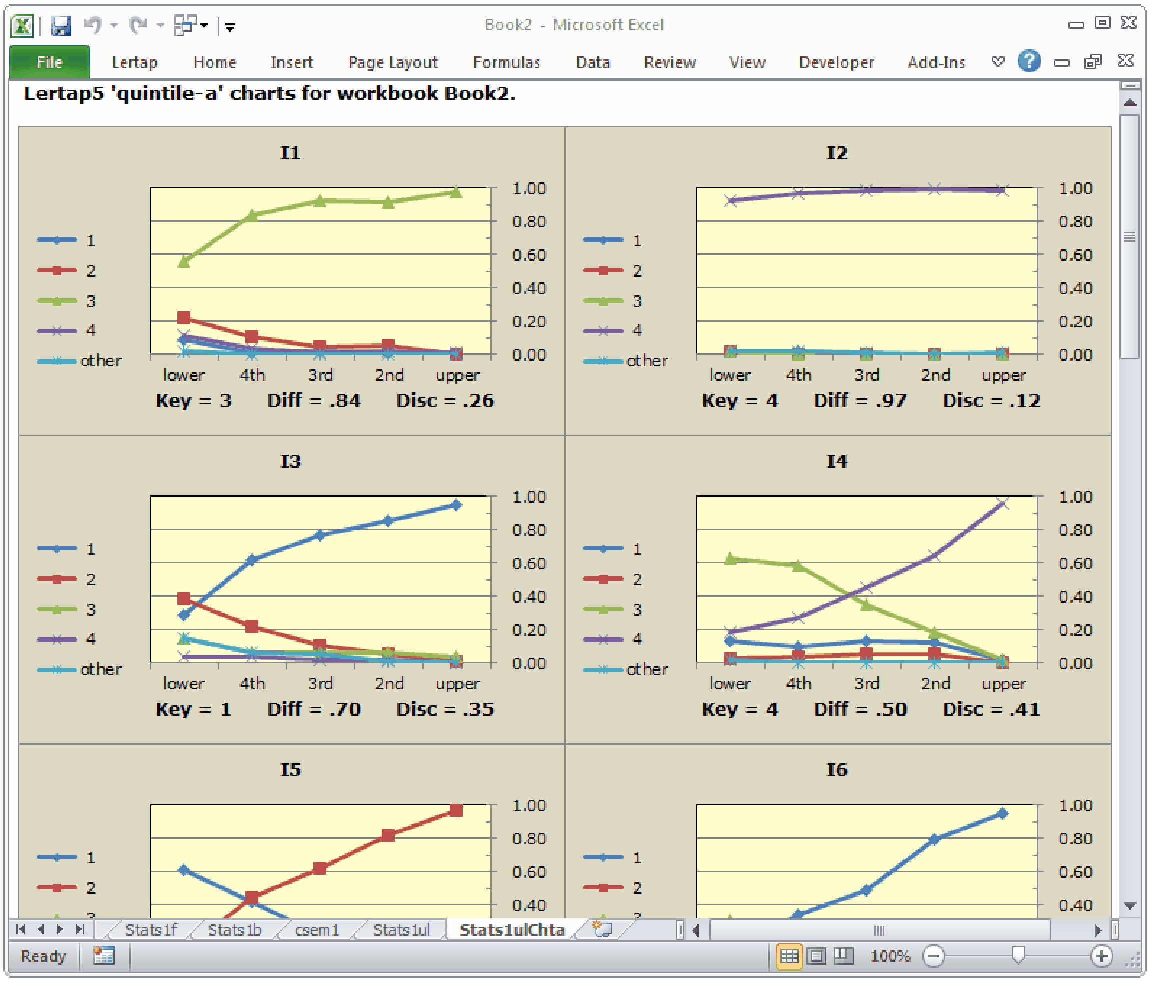


A special setting makes it possible to get normal quintile plots and ChartChanger3's packed plots with just a single click. This setting, and several others pertaining to the use of ChartChanger3, are described in the next topic (following page).

Another way to get packed plots is to use "ChartChanger2". Users with any version of Excel are able to use it, including those with Excel 2011 on a Macintosh. It's part of the set of macros found in the Lertap5MacroSetA.xlam file. This file is a standard Lertap component; when you start up Lertap, the file is automatically opened, and its macros are made ready for use by Excel and Lertap.

ChartChanger2 is not as simple to use as ChartChanger3, but it does have a bit more flexibility. It is accessed via the "Macs" menu; please refer to the Macs menu topic for more specific details on how to get it running.

Here's one example of what can be done by inviting ChartChanger2 to change the layout of a set of quintiles (these charts are based on the "MathQuiz" dataset):



In this case, ChartChanger2 was directed to make a grid having two charts per row with no blank lines between the rows.

If you're using a large monitor with your computer, you could have four or five charts per row and end up with a display which quickly captures the performance of all test items. (A useful tool for an item-review session with colleagues.)

The ChartChanger2 macro makes it a simple matter to resize quintiles. In the example above, the charts were made with the Excel 2010 version of Lertap. The first quintile was reduced in size, the macro was run, and the display above was the result. The Excel 2007 version behaves the same way.

While we're talking about changing quintiles, why not take a minute to consider the matter of deleting some of the quintiles? If you'd like to delete a quintile, just right-click on it and select "Cut". Bingo -- it's gone.

But doing this, deleting a chart, will leave a hole or a gap in the display. Not to worry -- just run ChartChanger2 again.

We could also discuss how to use Excel to select all of the charts on any worksheet in one quick go: from the Excel 2007 or Excel 2010 Home tab, find and click on the "Find & Select" option. From the list which drops down, choose "Go to". In the wee dialog box which then displays, click on the "Special" button at the lower-left. Finally, from the larger dialog box which will pop up (or down), select "Objects" and click OK. (You can quick-start this process by using the F5 key at the top of your keyboard.)

What about printing quintiles? Read a how-to on printing, and gain more insights into using ChartChanger2 here, at this topic.

4.5.3.2.1.1 ChartChanger3 settings

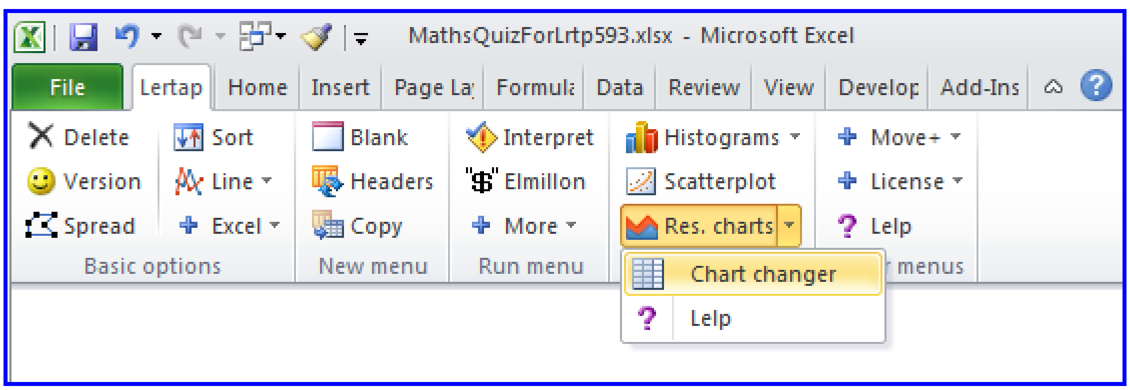
ChartChanger3 is by far the easiest way to pack your plots if you have Excel 2010 or later.

To use it you must first have a set of quintile-like plots on hand (we say "quintile-like" as quintile implies five groups; if you have fewer groups the plots are not really quintiles, but, well, just play along, if you would).

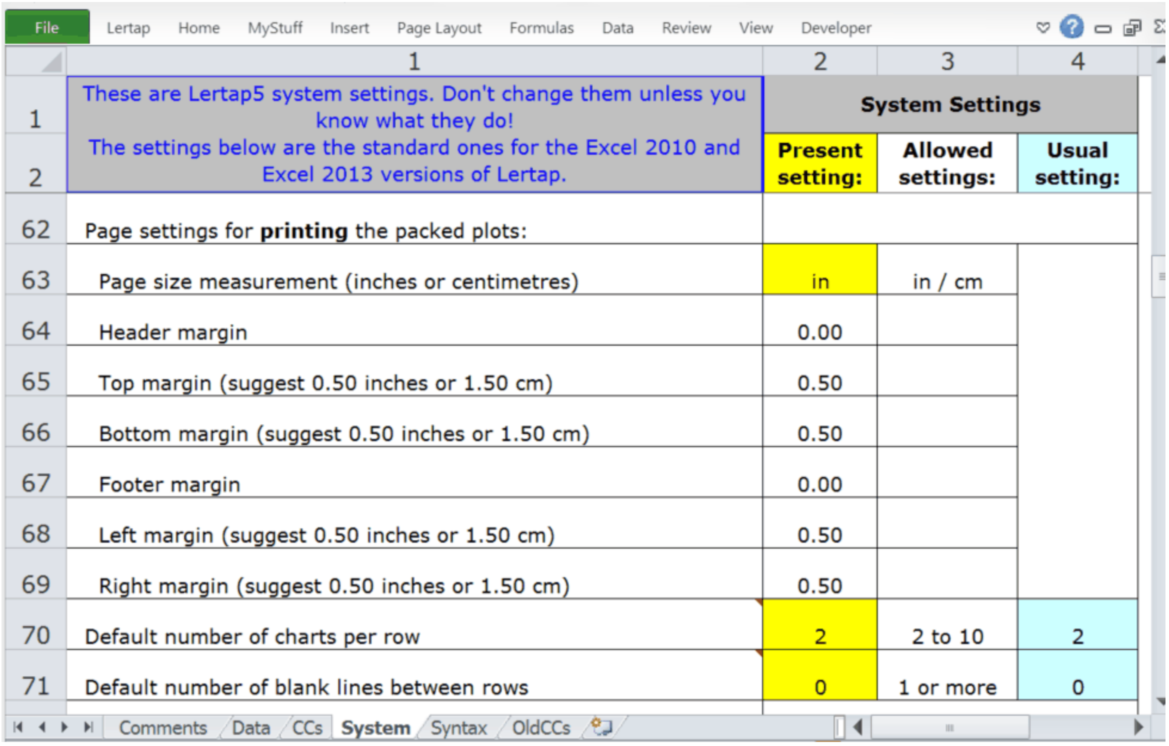
A reminder on how to get quintile plots. Start by viewing a Stats1ul sheet. Then click , the Res. charts option.

This will result in a new worksheet, or Lertap "report", with one plot for each item arranged in a top-down manner: item one at the top, followed by item two below it, and so on. Such reports have names like 'Stats1ulChta'.

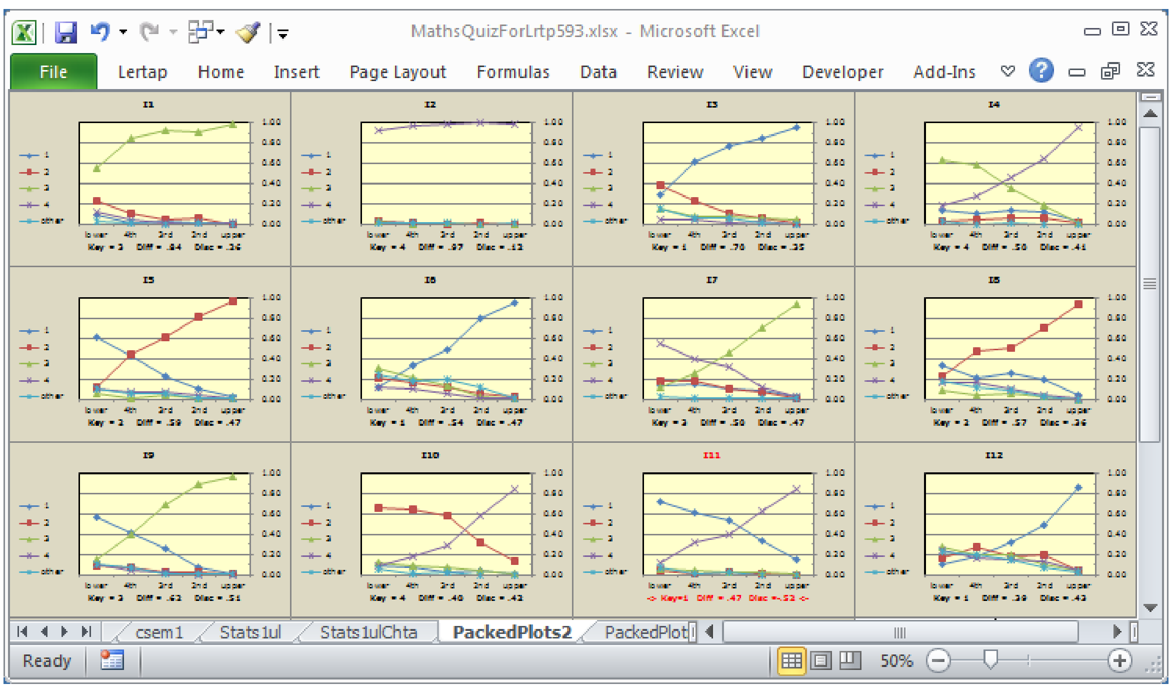
While viewing these quintiles, ChartChanger3 is then activated by clicking on the 'Chart changer' option.



What we've just described is a two-step process. You can make it a single step by going to the System worksheet and changing Row 60's 'Present setting' from 'no' to 'yes'. Once this is done, a new worksheet, "PackedPlots", will accompany the Stats1ulChta report (the rows seen below start at Row 62 -- scroll up to Row 60).

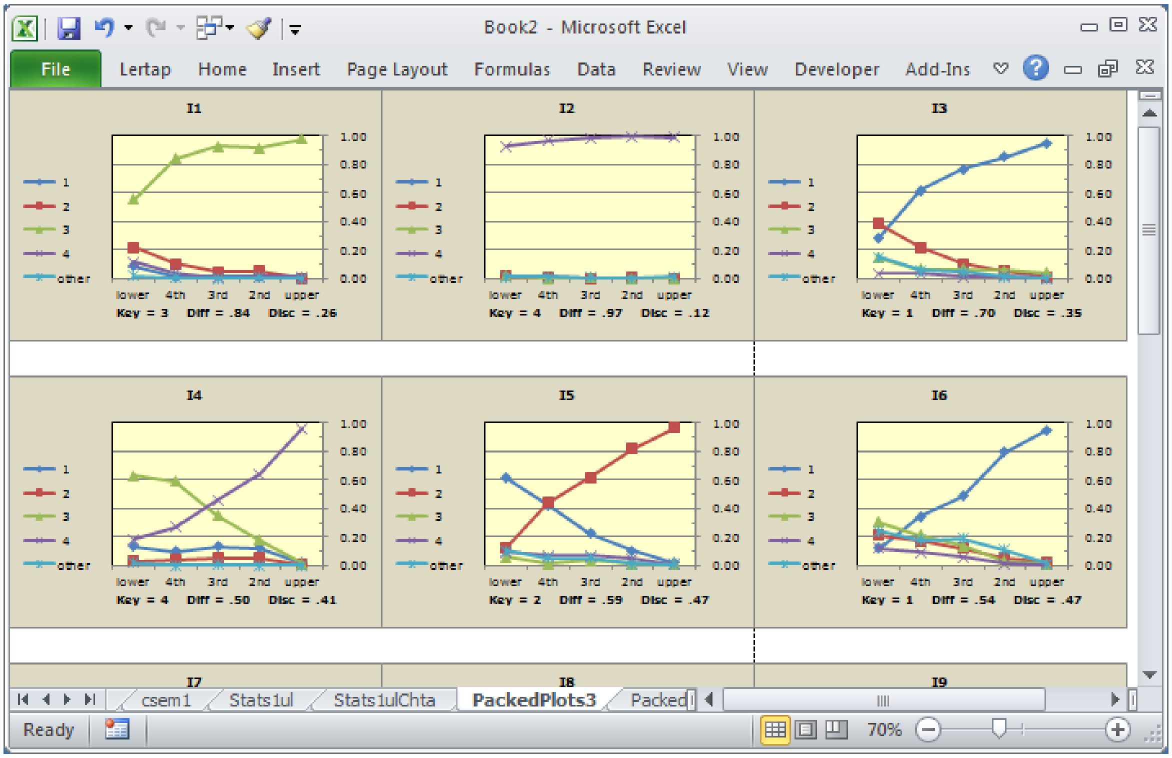


Plot packing involves moving charts so that they lie more than one to a row. In the example below there are four to a row.



The number of charts in a row is controlled by System worksheet Row 70. You can set it to whatever you prefer; 2 is a suggested default value.

It is possible to insert a blank row between the packed plots, as seen here:



The number of blank rows, or lines, is controlled by Row 71 in the System worksheet.

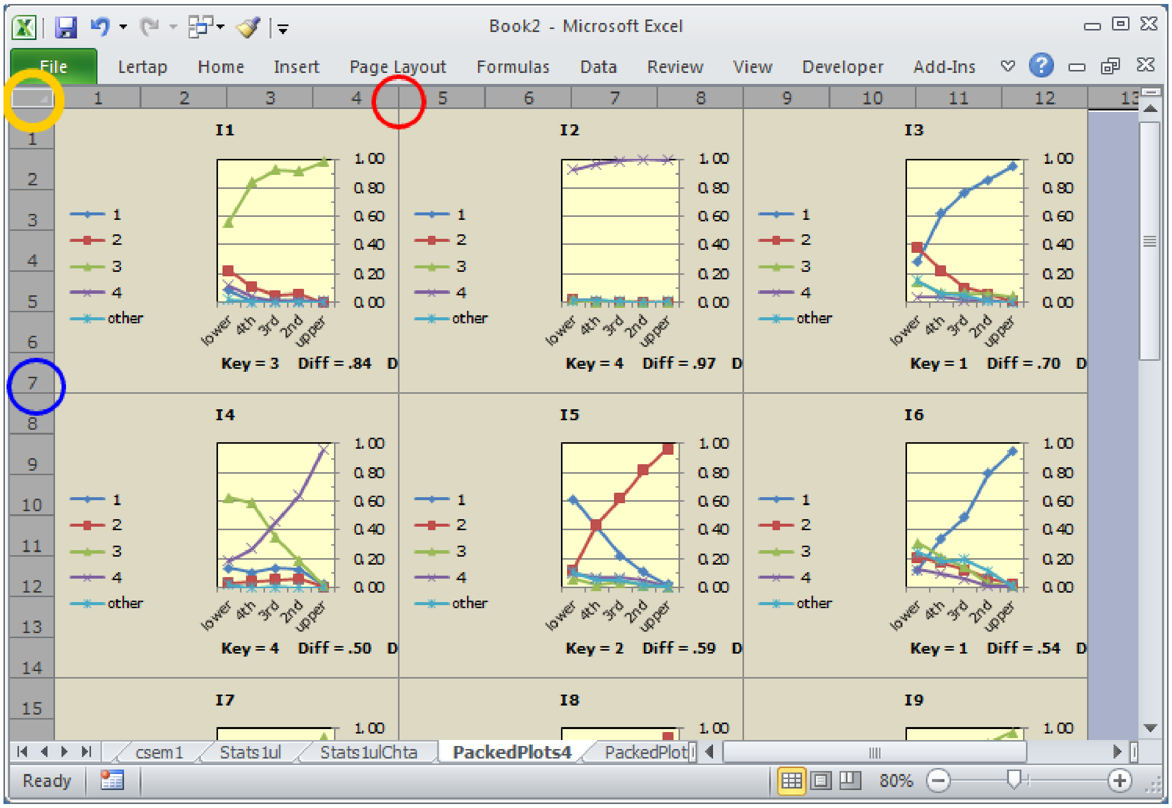
If the "interactive mode" setting in Row 61 is set to 'yes', then the values entered in rows 70 and 71 are ignored, and you're asked to enter them as soon as you've activated ChartChanger3. However, if "production mode" is set to "yes" in System row 35, then the "interactive mode" setting is ignored and the values in rows 70 and 71 are applied.

The Page settings rows in the System sheet are there to make it easier to print packed plots. You'll have to experiment with these, finding the best values by using Excel's many print options and then recording the settings which seem to work best for you in Rows 64 to 69 of the System worksheet.

Note that the margin settings in rows 64 to 69 are read as either inches (in) or centimeters (cm) depending on what's in the yellow box in row 63 (must be either in or cm). An inch = 2.54 cm. A centimeter = 0.39 in.

There's a special topic in the Lertap sample datasets website which discusses printing quintiles. Have a look at it, and then note that ChartChanger3 makes it easier to re-size charts and align them to the Excel grid. Much easier. Much much easier.

After ChartChanger3 has run, it will automatically select all of the packed plots and wait for you to re-size them. And, as you do, they automatically align with the Excel grid. (Aligning with the grid is useful as it makes page breaks easier to adjust when it comes to printing.)



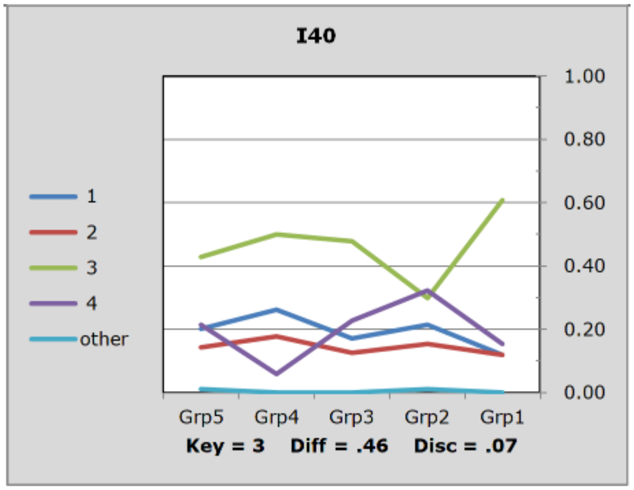
This snapshot shows what ChartChanger3 will sometimes do: create plots which are too squished. Fix them by following the comments in the next paragraph.

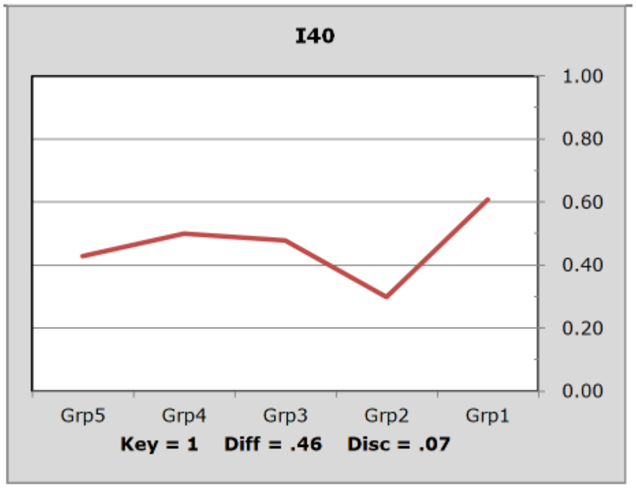
Once all the plots have been selected (they should be automatically selected; if not click on the small box circled in yellow), *slowly* drag a column divider, such as that circled in red above, to the right to make the plots wider. To make them taller, or shorter, slowly drag a row divider, such at that circled in blue. As you do these things the plots will still be aligned with Excel's underlying grid. To get them to print well you'll have to adjust the page breaks, but that's fun and simple: to see how, take in the very bottom of this topic.

4.5.3.2.1.2 ChartChanger4

This macro has a simple but quite useful task: to remove response trace lines from a "quintile plot" so that only the trace for the keyed-correct response remains.

Here are some before-and-after snapshots, the original quintile plot followed by the "cleaned" plot.:





The color of the keyed-correct response trace line has changed from green to red, but the response pattern for the correct answer is now easier to see and interpret.

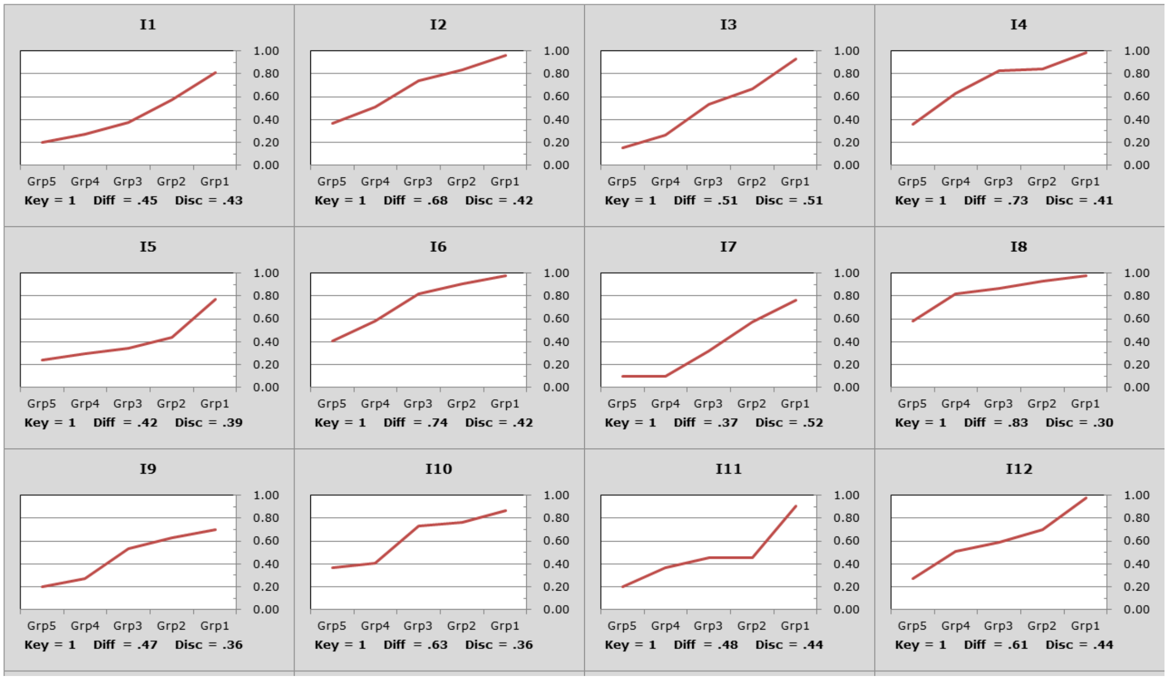
How to get a "cleaned" quintile plot?

Use the "BinaryItems" macro. Follow the steps under "How to use this macro" and "How to get the quintile plots".

If you don't see the ChartChanger4 macro listed, visit this page to find out how to activate it.

See the next topic for a snapshot of some "packed plots" made from the "cleaned" quintiles created by the ChartChanger4 macro.

The bouquet of "cleaned" quintile plots seen below was obtained by using another Lertap macro called "ChartChanger4".



A likely use for these plots may be appreciated by users who want to check to see that the response trace line for an item's keyed-correct option rises from left to right across the plot. This is almost always desired, especially when users are thinking about moving on to some sort of IRT analysis (item response theory). A hiccup in the rise may make it difficult for an IRT program to find model fit, as was the case when I40 was processed by two of the most popular IRT programs, Bilog-MG 3 and Xcalibre. (I40's plot is seen on the previous page. The I40 trace dips at Grp2, an undesirable outcome.)

Note that it is possible to change the number of groups used in these plots. There may be as many as 10 groups. This topic mentions how to change the number of groups from the default setting of five.

4.5.3.2.2 EC quintile plots

Special note: the charts displayed in this topic were produced by the Excel 2003 version of Lertap, now very dated.

It is possible to make both quintile-a and quintile-b plots with an external criterion score. When the Run menu's External criterion analysis option is selected, an upper-lower (groups) worksheet is created, assuming the subtest involved is a cognitive one, and the option to create upper-lower analyses has been set to Yes in the System worksheet.

When an external criterion is used, the upper-lower (groups) worksheet produced by Lertap will be called ECStats1ul, or, more generally, ECStatsXul, where "X" corresponds to the subtest involved. If the ECStatsXul worksheet is active, clicking on the charts icon will get the quintile plots rolling.

Now, part of the process of setting up an external-criterion analysis involves selecting a score from the Scores worksheet; in fact, it's this score which defines the external criterion.

Lertap will check the selected score to see if it might correspond to a categorical variable, such as Gender or, perhaps, Region. (Note: the Recode macro available via the Move+ Menu is useful for working with categorical variables. This macro will, for example, allow variables coded with letters to be recoded with corresponding digits.)

If the selected score is found to consist of values in the range 1 to 5, Lertap will define the number of 'upper-lower' groups as equal to the number of different values found. For example, if the selected score has only values of 1 and 2, Lertap will set the number of upper-lower groups to 2. If the selected score has values of 1, 2, 3, and 4 (for example), Lertap will set the number of upper-lower groups to 4. (This action over-rides the number of upper-lower groups setting in the System worksheet.)

Here's a practical example: 288 junior high-school students participated in a test development project which investigated the effects of coaching on test performance. About half of the students sat a practice test before taking the real one. Did this affect their achievement?

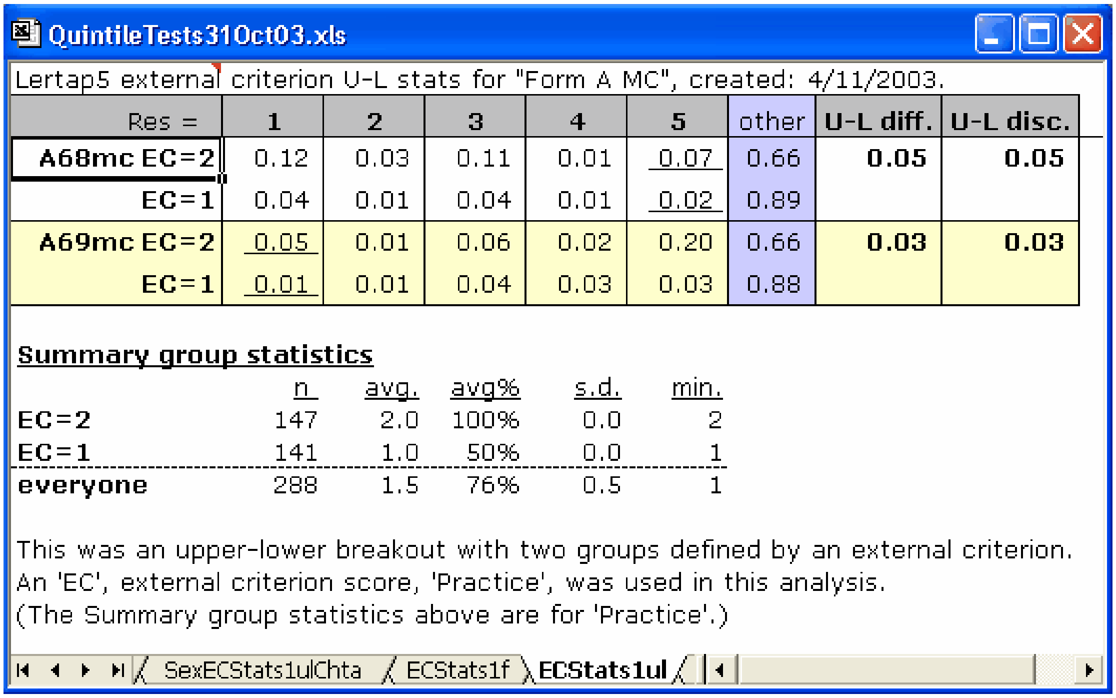
Data were entered into a Lertap workbook. One of the columns in the Data worksheet indicated whether or not the student had taken the practice test; this column was called "Practice". Practice=1 indicated the student had not taken the practice test, while a Practice value of 2 indicated that the student had sat the practice test.

Another Data column contained a code for gender, 1 for boys, 2 for girls. Other columns housed the student responses to the 70 test items.

The CCs worksheet was set up to score the 70 test items. The Run menu was then used to Interpret CCs lines, and to produce an Elmillon item analysis.

We used the Move+ menu to copy the Practice column from the Data worksheet to the Scores worksheet. Then we went back to the Run menu and started an External criterion analysis, telling Lertap to use the Scores column with Practice values as the criterion "score".

Lertap dutifully produced two new worksheets, ECStats1f, and ECStats1ul. We had a squiz of the latter -- at the very end we observed this info:


QuintileTests31Oct03.xls

Lertap5 external criterion U-L stats for "Form A MC", created: 4/11/2003.

Res =	1	2	3	4	5	other	U-L diff.	U-L disc.
A68mc EC=2	0.12	0.03	0.11	0.01	0.07	0.66	**0.05**	**0.05**
EC=1	0.04	0.01	0.04	0.01	0.02	0.89		
A69mc EC=2	0.05	0.01	0.06	0.02	0.20	0.66	**0.03**	**0.03**
EC=1	0.01	0.01	0.04	0.03	0.03	0.88		

Summary group statistics

	n	avg.	avg%	s.d.	min.
EC=2	147	2.0	100%	0.0	2
EC=1	141	1.0	50%	0.0	1
everyone	288	1.5	76%	0.5	1

This was an upper-lower breakout with two groups defined by an external criterion.
An 'EC', external criterion score, 'Practice', was used in this analysis.
(The Summary group statistics above are for 'Practice'.)

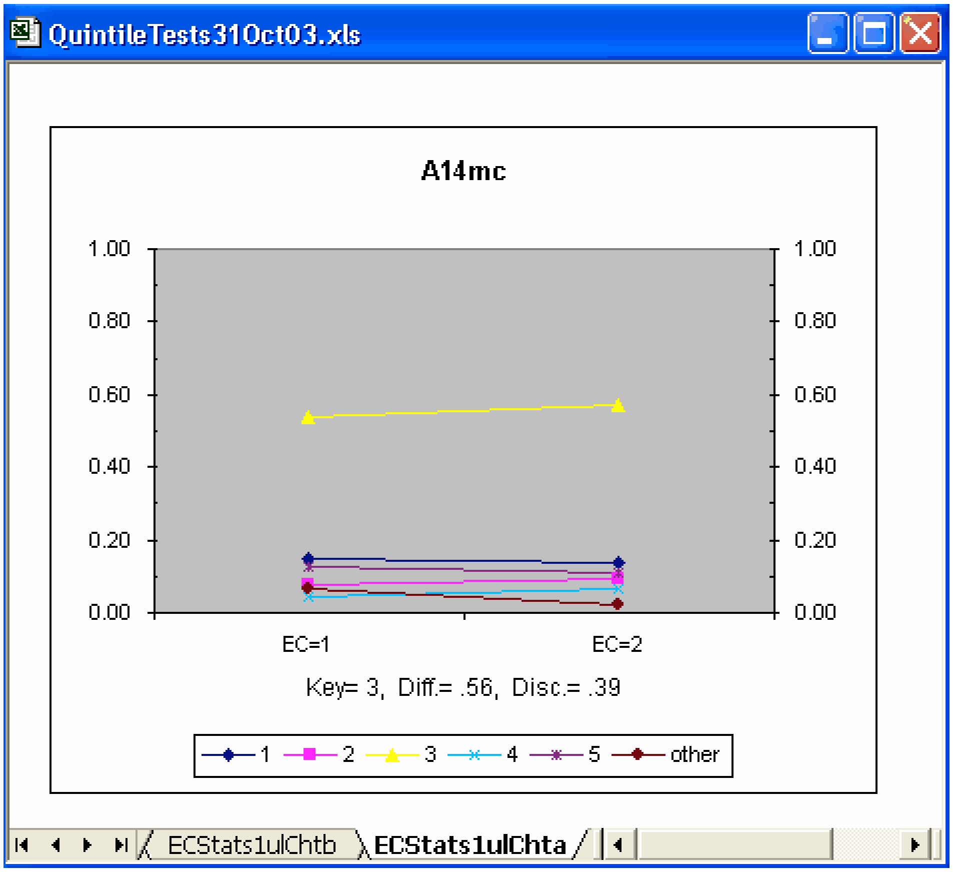
SexECStats1ulChta / ECStats1f / ECStats1ul

Lertap has denoted the two Practice values as EC=2, the Practice group, and EC=1. It will always do this, that is, always denote the categorical variable as "EC", appending its various values. Had there been four groups, we would have had EC=4, EC=3, EC=2, and EC=1.

Notice the two s.d. values? Zero point zero, and zero point zero. This is correct -- the groups have been defined by a categorical variable; all members of each group have the same "score" on this variable -- there is no variance of scores within the groups, no standard deviation.

We clicked on the charts icon once, and got a new worksheet called ECStats1ulChta. Next, we made the ECStats1ul sheet active (by clicking on its tab), and then clicked on the charts icon again to get ECStats1ulChtb.

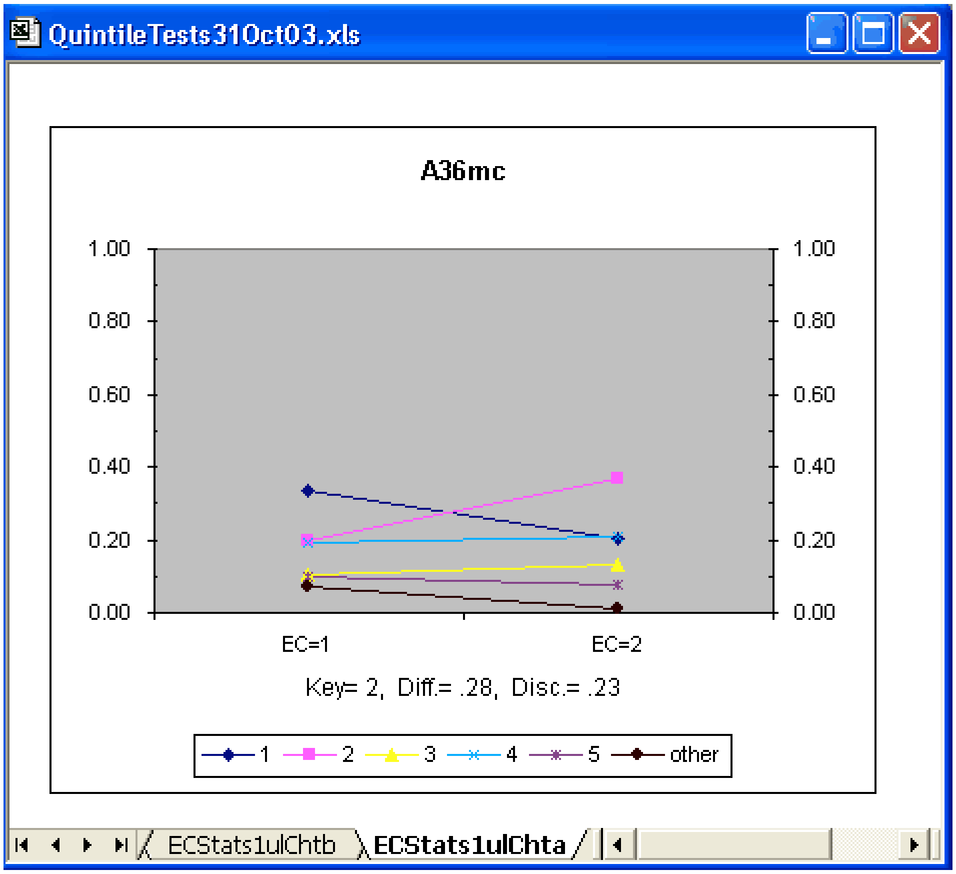
What did we find? Most of the initial items had this sort of quintile-a plot:



The plot above has lines which seem to be fairly horizontal -- we'd expect horizontal lines if there were no group differences -- truly horizontal lines indicate that the percentage of people selecting an option is the same in each group, meaning that there are no differences among the groups. (Some readers will recognise this discussion as being similar to that heard when folks sit down with their coffee / tea to yak about "item bias" and DIF, differential item functioning. Please see the Ibreaks topic for much more about DIF.)

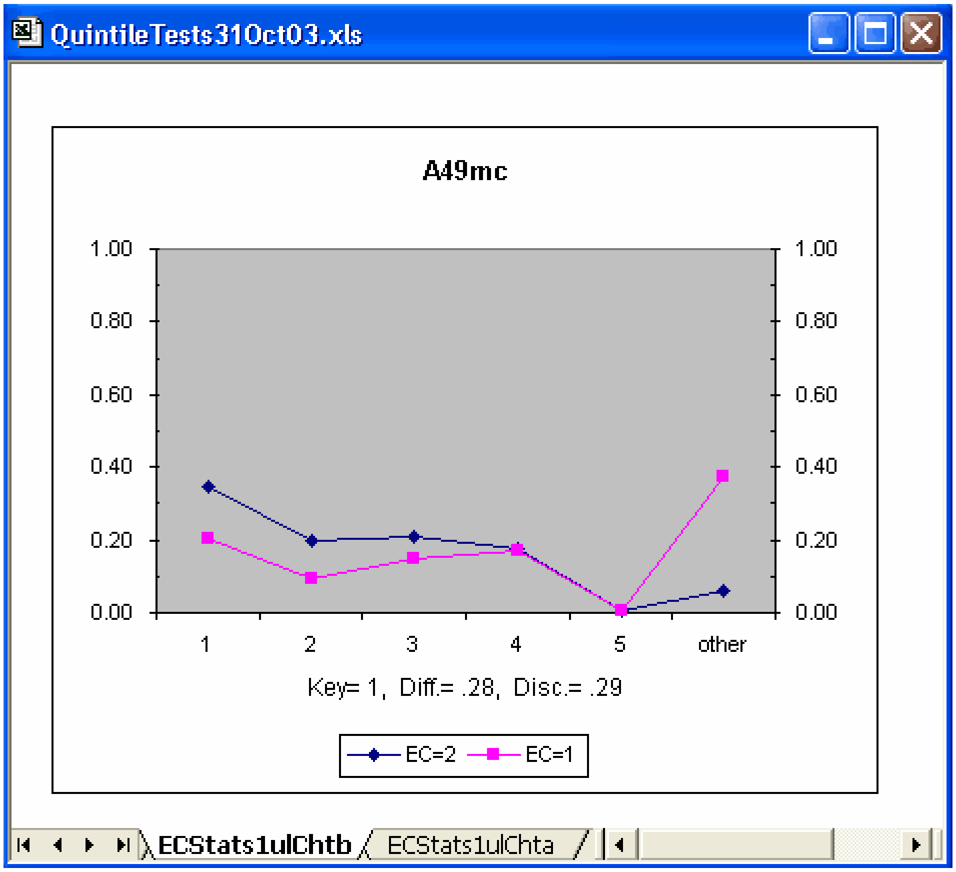
Above, the proportion in each group selecting each of the four distractors looks to be about the same (except for "other"); the proportion of correct answers (Key= 3) was about the same in each group, perhaps showing a slight practice effect (the line rises somewhat as it moves from left to right).

But now look at item A36mc:



The proportion of people who identified the keyed-correct answer (2) was noticeably higher in the practice group. The lines are no longer all close to horizontal.

There were a few other items with patterns like A36mc's. However, it was a study of the quintile-b plots which highlighted a major message: the practice group stuck it out longer -- they answered more test items; students without practice tended to get bogged down, appearing to run out of time. Look:



"Other" means a student omitted the item. Notice that almost 40% of the no-practice group omitted item A49mc, whereas the omit level in the practice group was below 10%. This pattern set in at item A49mc, and continued to the end of the test without exception. (Well, the gap did narrow somewhat after about the 65th item when the proportion of omits in the practice group began to rise rapidly.)

Of course we didn't need plots such as these to reach this finding. No; the numbers are all there in the source data, in the ECStats1ul worksheet. But you might agree that the plots are more effective in conveying the message. If there's a pattern in the data, plots such as these can help to uncover it.

4.5.3.2.3 Quintile options

There are a few options which apply to these plots. Almost all of them are controlled by the System worksheet found in the Lertap5.xlsm workbook. Below we've displayed part of the System worksheet (and please note: the next topic has even more options added with the release of version 5.10.5).

	1	2	3	4
1	These are Lertap5 system settings. Don't change them unless you know what they do!	System Settings		
2	The settings below are the standard ones for the Excel 2010, 2013, and 2016 versions of Lertap.	**Present setting:**	**Allowed settings:**	**Usual setting:**
9	Should **brief item stats** sheet be output?	yes	yes / no	yes
10	**Should upper-lower stats sheet be output for cognitive tests?**	yes	yes / no	yes
11	**Minimum percentage** score for "mastery" level:	70%	10% to 90%	70%
12	**Percentage** in Upper & Lower groups:	27	> 0	27
13	**Number** of "upper-lower" **groups** (quantiles):	5	2 to 10	5
14	**Primary** (first) **quantile plot** type**:**	A	A or B	A
15	Should quantile plots include a **data table**?	no	yes / no	no
16	(empty)			

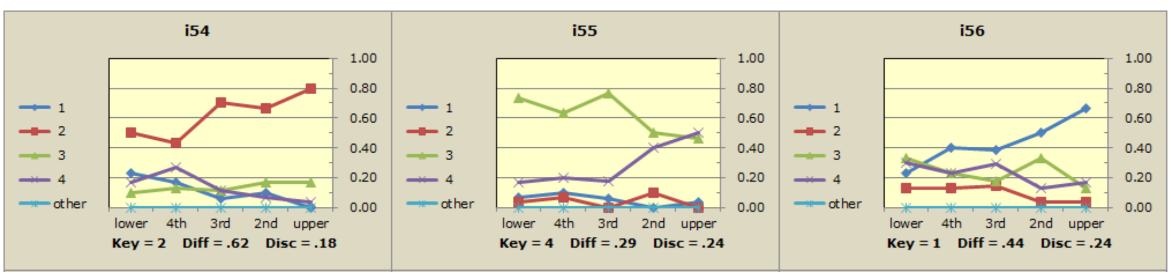
The "quintile options" are set in rows 13 through 15.

Note: more than a dozen new options were added in 2015 with the release of version 5.10.5. Please refer to this key document for much more information, and some fancy new examples.

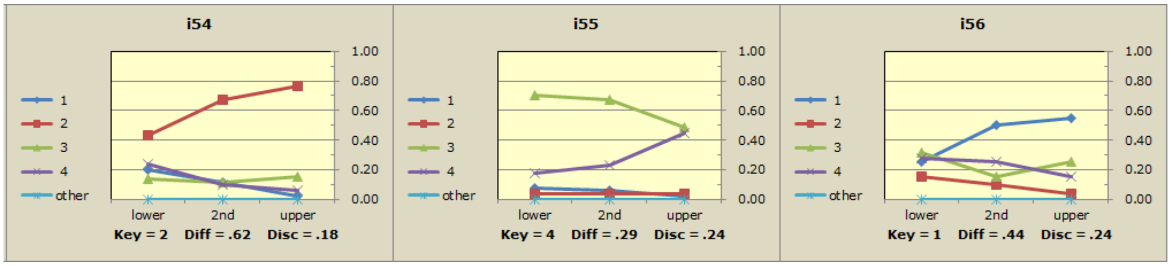
The number of groups for these plots is seen above in row 13. Its default setting is 5, a setting which gives rise to the term "quintile" plots; the more general term, applicable to any number of groupings, is "quantile".

Changing the number of groups can often result in smoother plots. Ideally we'd like to have, say, at least 50 students in each group -- less than this and these plots can look a bit ragged.

For example, consider a mid-term exam with 68 items and 154 students. With 5 groups of about 30 students each, the plots for three selected items are shown here:

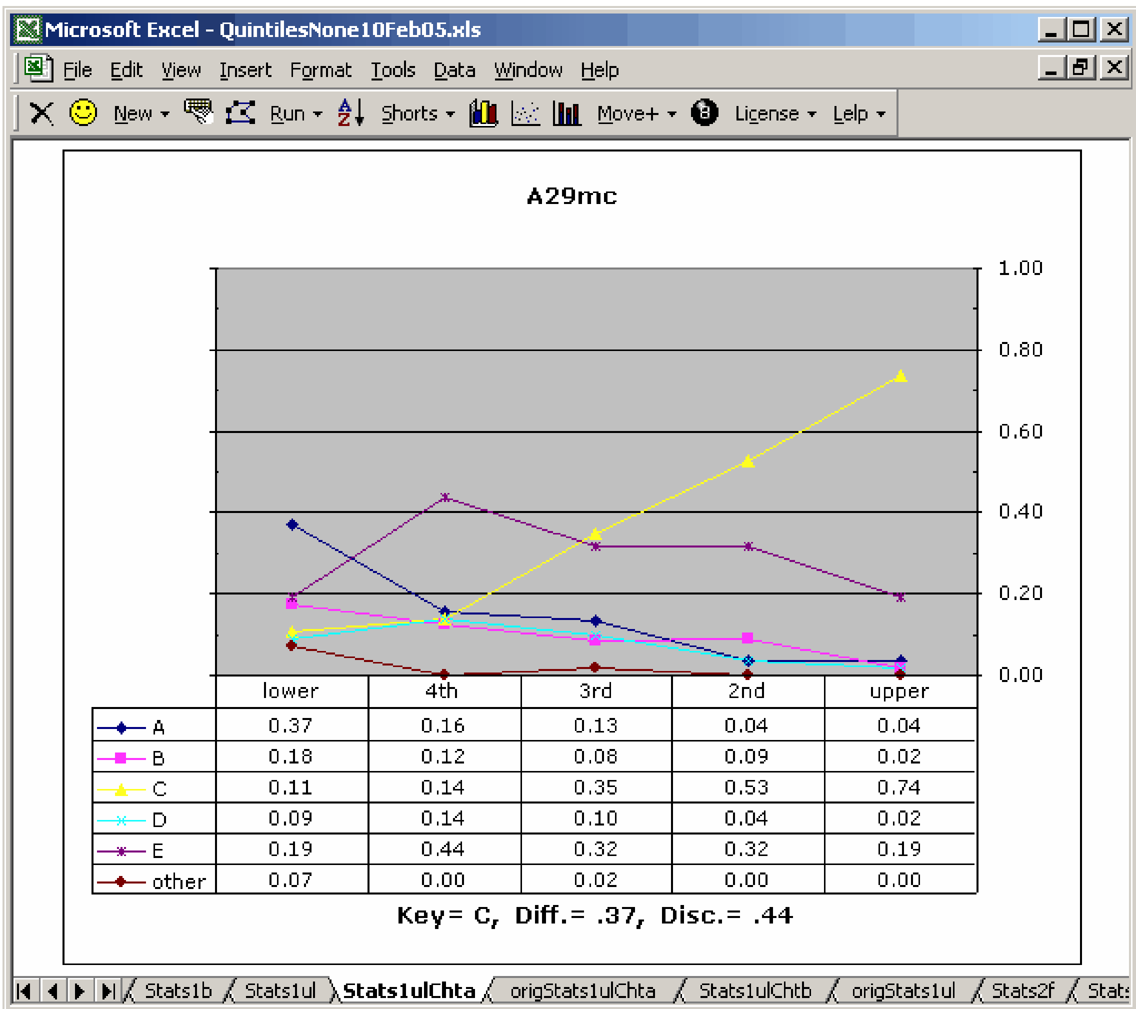


Taking the number of groups down to 3, with about 50 students in each, results in smoother plots in this case (with three groups, these plots would be called "terciles" or "tertiles").



Lertap makes two types of "quintile plots", "quintile-a" and "quintile-b", as [discussed earlier](). The setting in row 14 of the System worksheet controls which of these plot types is produced first.

The row 15 option in the System sheet, when set to yes, gets Excel to add a data table at the bottom of each plot. Here's a quintile with table:



The information in the data table is the same as that found in the Stats1ul report for the respective item.

There's a "data table toggle" which may be used to turn the data tables off. In fact, the same toggle can turn them on if they're off, even when the row15 option is set to "No".

I love quintiles, put yeesch, those colours! They don't look good at all when printed on a black and white printer. Not to worry, help is at hand: see the Chart colors topic.

4.5.3.2.3.1 Additional options

With the release of version 5.10.5 at the start of 2015, a swag of new options was added for controlling quantiles:

2	The settings below are the standard ones for the Excel 2010 and Excel 2013 versions of Lertap.	**Present setting:**	**Allowed settings:**	**Usual setting:**
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

There's so much to say about these options that we ask you to branch out to a special topic paper.

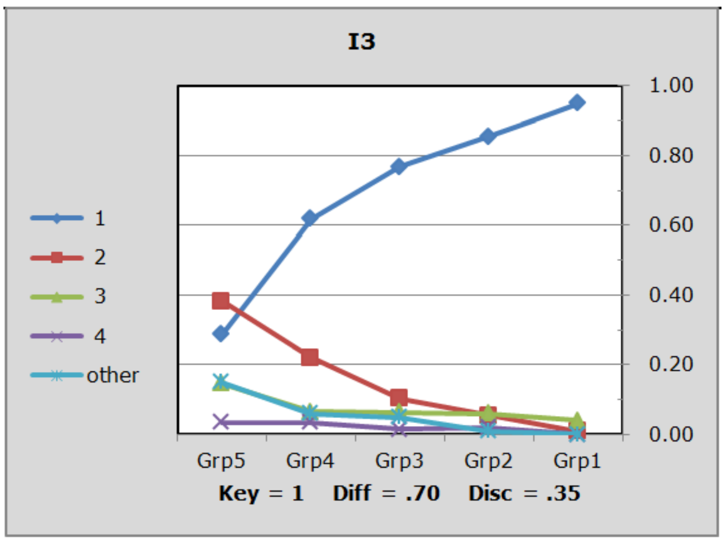
Row 95 makes reference to rows 64-69, mentioned in this topic.

In early 2016, additional options were added in rows 96, 97, and 98. Then, in the Lertap 5 versions of 1 June 2018, row 99 came to have an entry too:

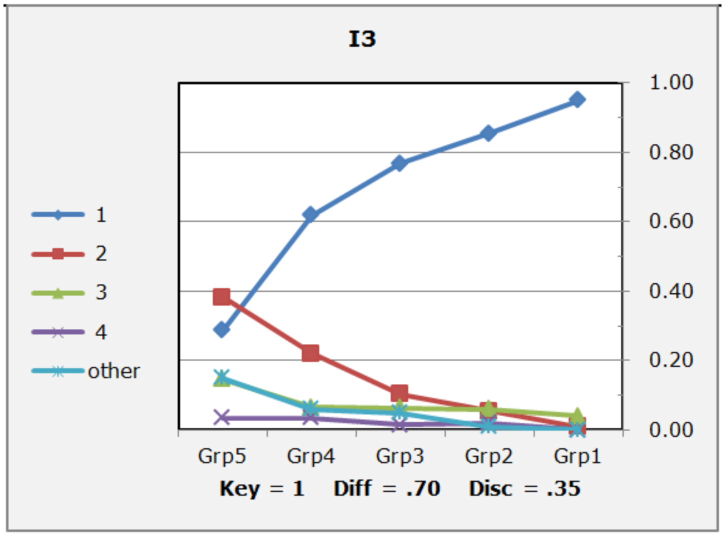
96	Remove the series markers from the plot lines?	yes	yes/no	yes
97	Present black & white option for the plots?	yes	yes/no	no
98	Use very light backgrounds in the chart frames?	no	yes/no	no
99	Thickness of each plot line (except that for 'other')	2.75	1.00 to 5.00	2.75

Rows 96 and 97 have to do with a special macro called "QuantileShader1".

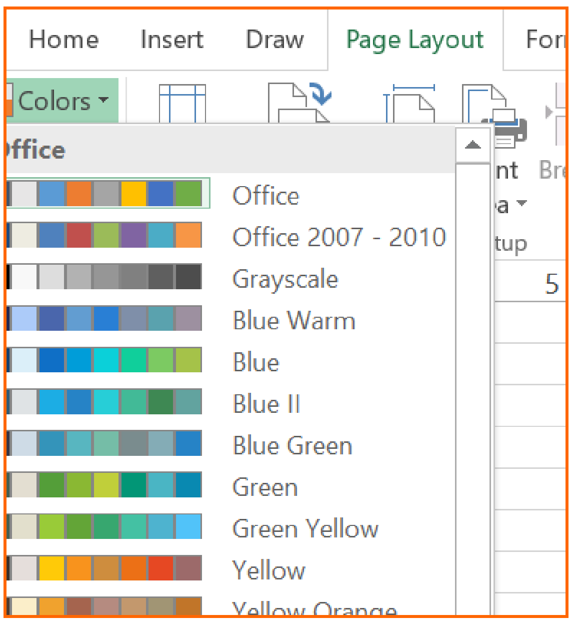
Row 98 affects the shading in quantile plots. When row 98's second column is set to "no", a quantile will look like this:



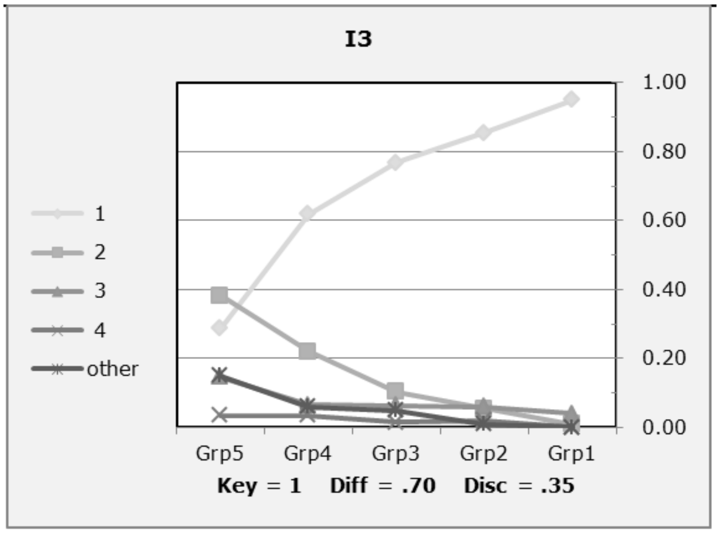
When row 98 is set to "yes", the shading of the chart area is a bit lighter:



Can all colors be washed out of these charts / plots? Yes. One way, a quickie, is to use Excel's Page Layout tab, Colors, Grayscale:



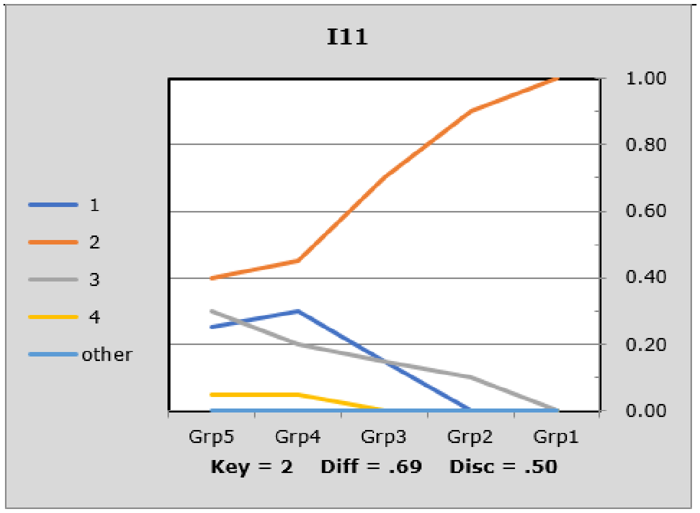
Here's the plot in Grayscale:



The "QuantileShader1" macro may also be used to alter colors.

The row 99 setting was added after Excel 2016 changed the default setting for line thickness. In Excel 2010, the lines in these plots had a default thickness of 2.25. This changed in Excel 2016, when 1.50 became the default thickness.

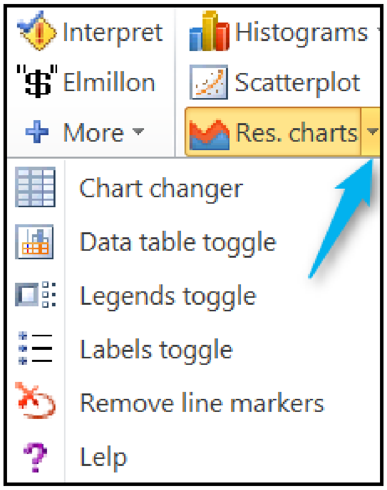
All of the plots exemplified above have a line thickness of 2.25 and were made using Excel 2010. The plot below, made with Excel 2016, used 1.50 for its line thickness, and, for this plot, line markers were not in use:



All versions of Lertap5 dated from 1 June 2018 allow quantile plot line thickness to be controlled by the setting in row 99 of the System worksheet.

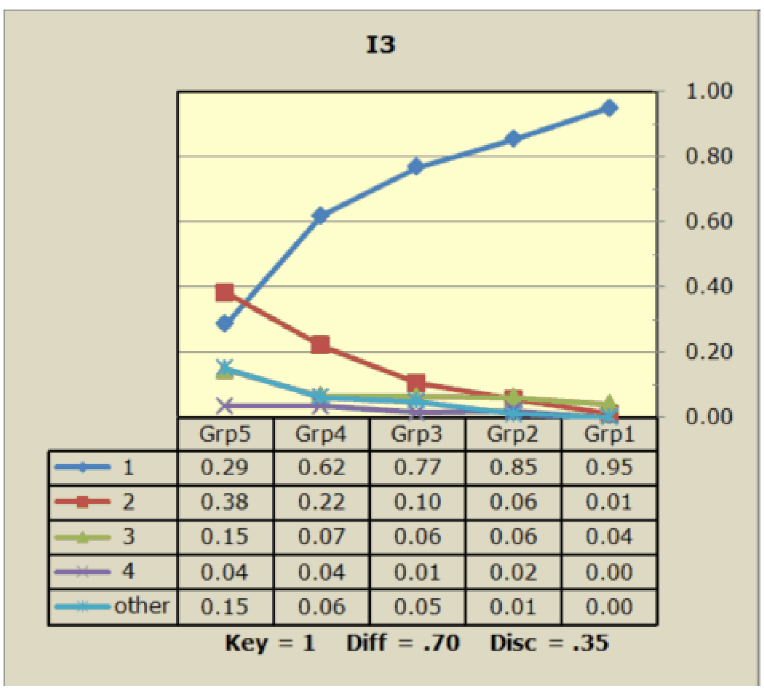
4.5.3.2.3.2 Response Charts Toggles

The screen capture below indicates how to get access to a special response charts menu. The options on this little menu can be quite handy.

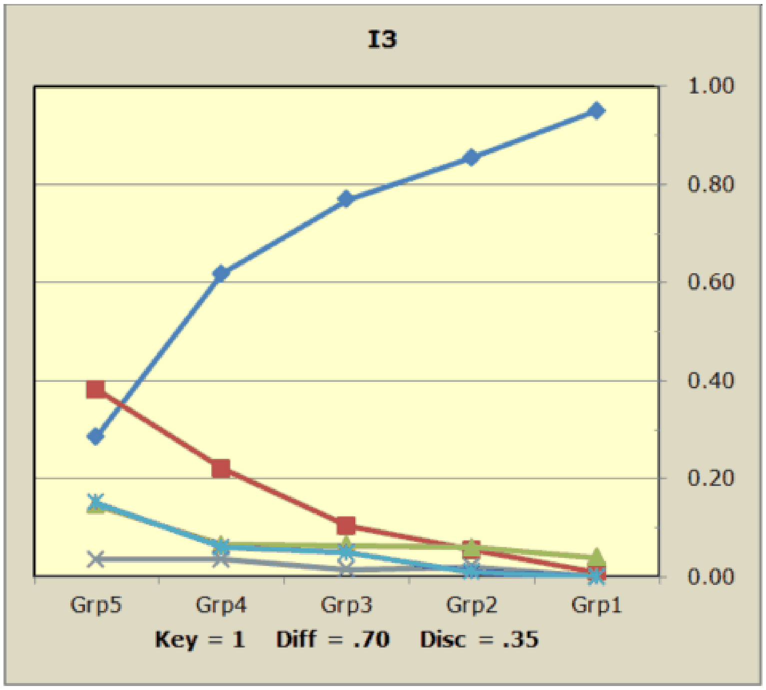


The "Chart changer" option is the gateway to a powerful feature called "packed plots".

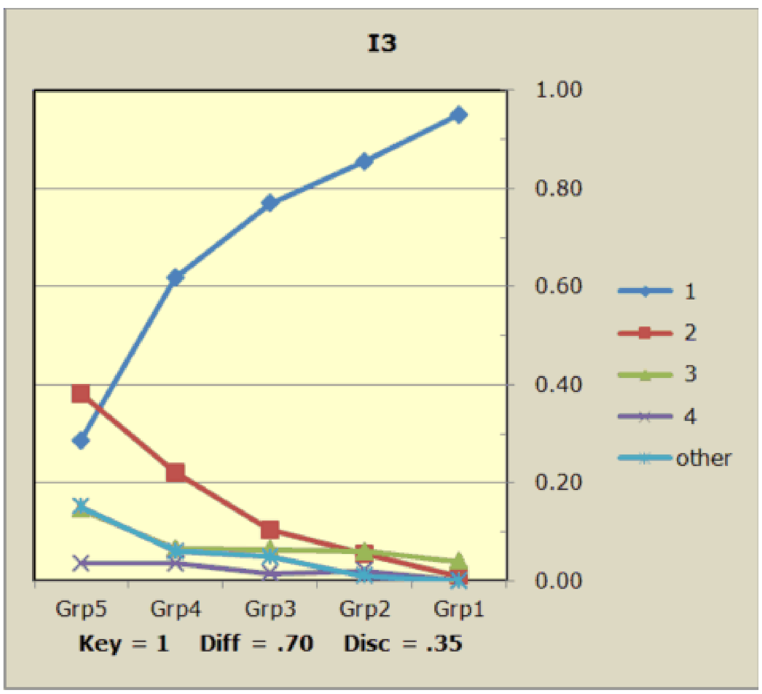
The first two "toggles" are used to show or hide data tables and legends.



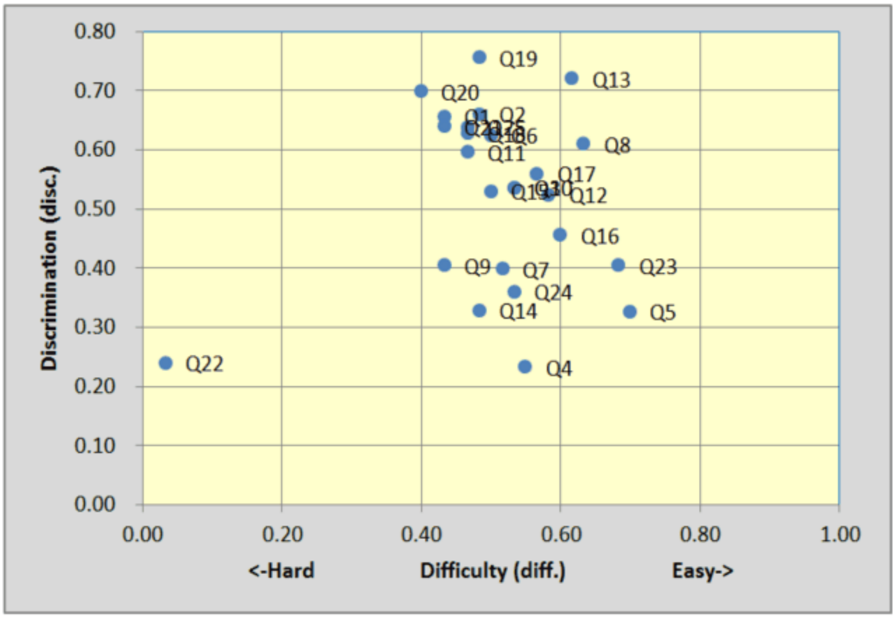
The plot above has a data table. The plot below does not.



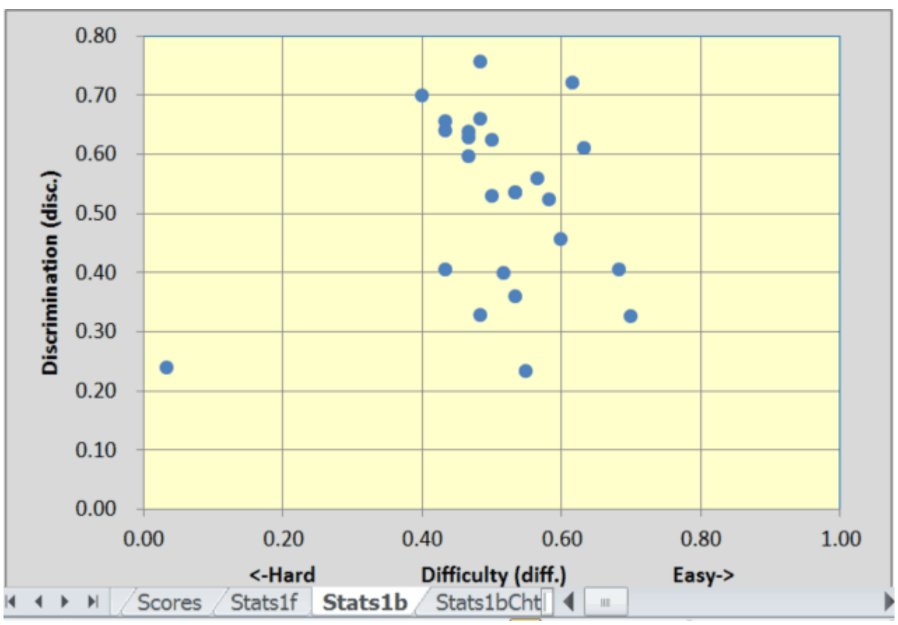
The plot above does not have a legend. The plot below does.



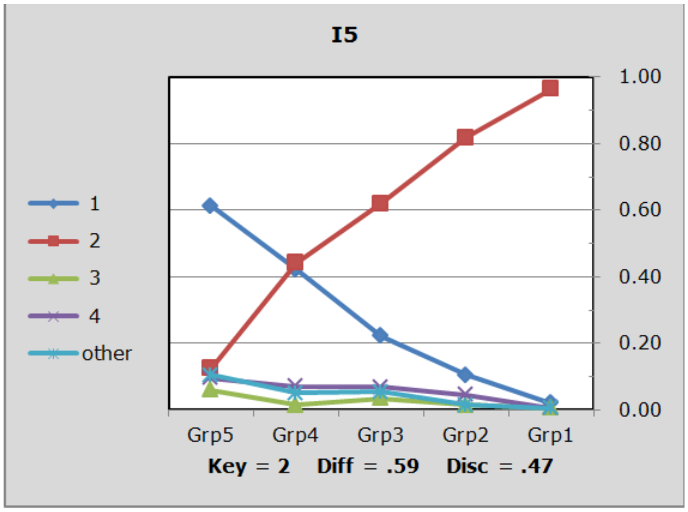
The third toggle, the "Labels toggle", is used to hide or show the item labels seen in a Stats1b scatterplot of item "diff" by item "disc".



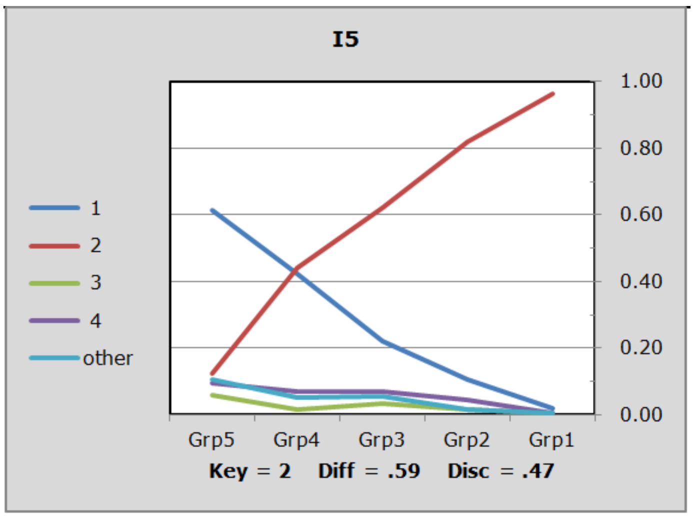
The scatter above is showing the item labels; the scatter below is not.



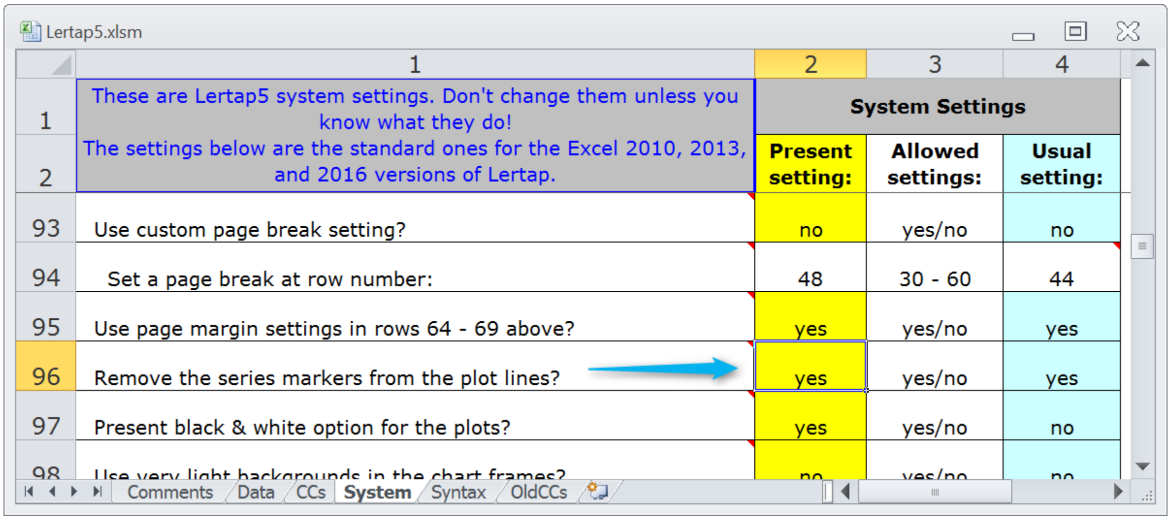
The "Remove Markers" option will delete the markers used to distinguish trace lines:



The plot above uses line markers; the plot below does not.



There's a setting in Row 96 of the System worksheet that controls whether or not markers are used when these plots are made:


Lertap5.xlsm

	1	2	3	4
1	These are Lertap5 system settings. Don't change them unless you know what they do!	System Settings		
2	The settings below are the standard ones for the Excel 2010, 2013, and 2016 versions of Lertap.	Present setting:	Allowed settings:	Usual setting:
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	yes
96	Remove the series markers from the plot lines?	yes	yes/no	yes
97	Present black & white option for the plots?	yes	yes/no	no
98	Use very light backgrounds in the chart frames?	no	yes/no	no

Comments / Data / CCs / System / Syntax / OldCCs

If this setting is "no", then the markers will appear on the plots when they're initially generated. They may be deleted by using the "Remove Markers" option. If the setting is "yes" then the markers will not appear at any time. The "Remove Markers" options is not a toggle -- once one of these plots has its markers removed, the only way to get them back is to create the plots again using the "Res. charts" option, making sure that the Row 96 setting is "no".

4.5.3.3 Chart problems

There are a few problems which can arise when asking Lertap to ask Excel to make charts.

Probably the most common of these has to do with Excel running out of chart and font resources as it goes about making its plots. For comments on this problem, please refer to the topics immediately preceding this one.

At times Excel seems to gather too much speed when making charts, and will forget to apply some of the formatting which Lertap has built in. For example, the name of the item is always supposed to be in bold face, such as Item1, but Excel can, at times forget this (perhaps on days when it's feeling meek, not bold?).

Particularly annoying is the sometimes-noted tendency of Excel to scrunch the plots, to make them squatter than they're meant to be. If you suspect your plots are a bit on the flat side, or the fat side, or the squat side, then do this: make them again, and see if that doesn't fix things.

Keep in mind that you can get into Lertap's charts, and change them to your little heart's content: make new colours, put the legend at the bottom rather than the side, add or changes titles -- a good way to pass a rainy day, or to avoid doing something more urgent but less fun.

Be sure to write to us if you'd like to talk about charts. larry@lertap.com is us.

4.5.3.4 Chart colors

It is a fairly straightforward matter to change Excel's standard colour palettes.

If you'd like to change the colours Lertap uses to make its quintile plots, histograms, and group breakout plots, you can. In fact, you can alter Excel's colour palettes so that the changes you make automatically apply to all charts in a workbook. Once you've coloured things in the way you like, you can then pick up the colours used in one workbook, and carry them over to another workbook.

If you don't have a colour printer, you may have had the occasion to frown after trying to print the coloured charts made by Lertap, another Excel application, or a program such as SPSS. Frown not, fair friend: use the procedures introduced here, along with some experimentation, to pick colours which will print well enough, letting you get by until the departmental budget has enough funds to buy a colour printer.

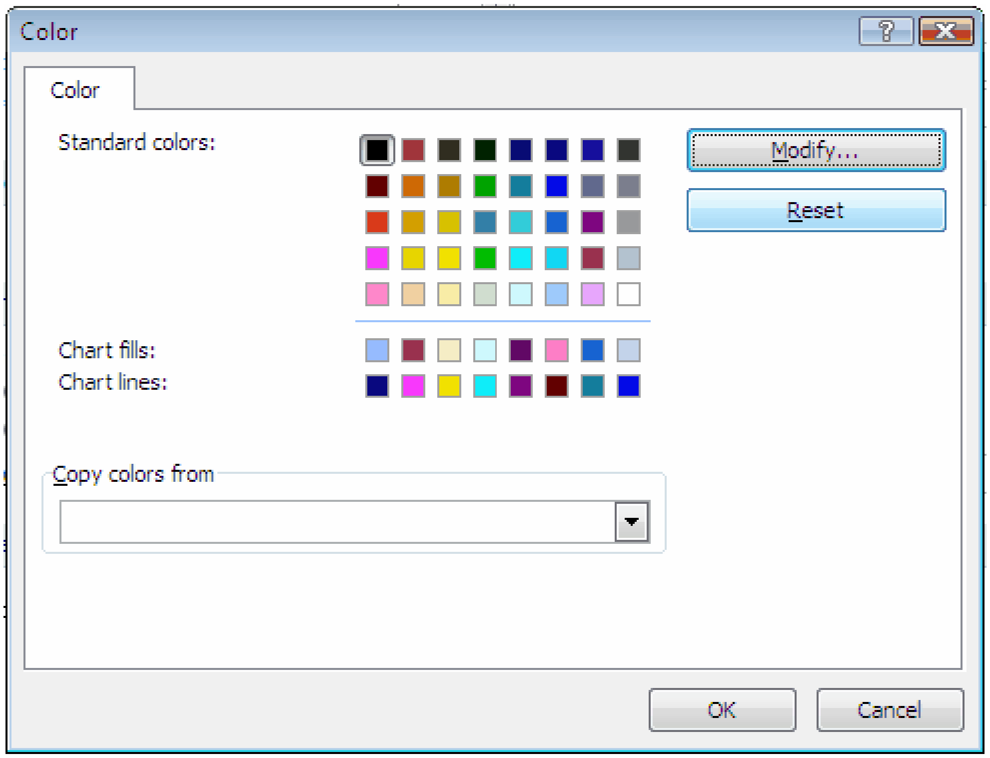
Here's what to do if you're an Excel 2007 or Excel 2010/2013 user:

Click on the Office Button (the File button in Excel 2010 and Excel 2013).

Click on Excel Options, a button found towards the bottom of the screen.

Click on Save.

Look for the section which says 'Preserve visual appearance of the workbook', and then click on Colors.



Then, for more information on how to re-map the colours, please refer to this Microsoft document.

4.6 Other menus

Way to the right-hand side of the Lertap tab for Excel 2007 there's a section for 'Other menus'.

Two sets of menu options are available: Move+ and License. Click on these options in the box above, or simply page forward to read more about them.

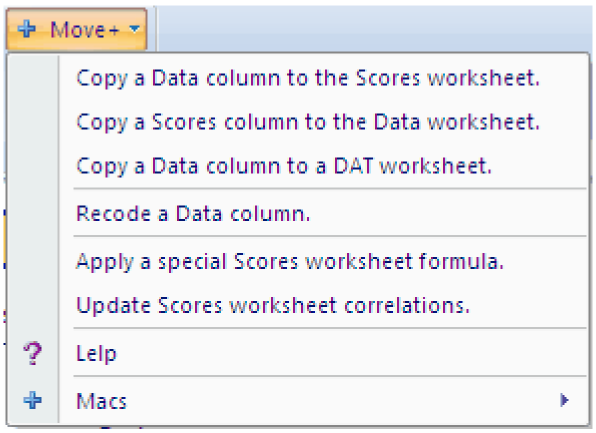
The  option opens (or should open) the very document you're reading now, but on the first page.

4.6.1 Move Menu

Among other things, the Move+ menu permits columns to be copied from/to Lertap's main information worksheets. A column in the Data worksheet may be copied to the Scores worksheet, or, when it exists, to the DAT worksheet.

The Data worksheet is created by users. The Scores worksheet is created by Lertap when the Elmillon program is run. The DAT worksheet, in turn, is created when the 'Item scores and correlations' option is selected from the Run menu.

Shifting data around from one worksheet to another was originally the Move menu's main purpose. These days, the Move menu will do quite a bit more, as you'll see if you click on the options seen in the box below, or just browse forward.



The manual briefly discusses the Move menu in Chapter 10 (in the printed manual, see p. 173). However, the manual's discussion is limited to the first two Copy options seen above -- the other options were added after the manual was printed.

4.6.1.1 Copy Data column

This option copies a designated column in the Data worksheet to the Scores worksheet.

The columns of the Scores worksheet may contain only numeric data. Before Lertap will copy a column from the Data sheet to Scores, it makes sure only numbers are found in the column to be copied.

Why will users want to copy columns from Data to Scores? There are a few reasons.

One of the most common reasons is to correlate the values found in a Data column with the values found in a Scores column. For example, the Data worksheet may have a column with SAT test scores; these are to be correlated and scatterplotted with the test scores produced by Elmillon, as found in the Scores worksheet. Copying the respective Data column to the Scores worksheet will automatically correlate the Data column's scores with the other scores made by Elmillon, and open the door to use of the Scatterplot icon on Lertap's toolbar.

Users wanting to carry out external criterion analyses sometimes have entered the criterion scores in one of the columns on the Data worksheet. These scores *must* be moved over to the Scores worksheet before they may be used as an external criterion.

When a Data column is copied to the Scores worksheet, Lertap doesn't know what to put in the MinPos and MaxPos cells after it has been copied. It writes "Unknown" in these cells, leaving it to users to put in proper values. (MinPos and MaxPos values are required by some of Lertap's routines, such as the external criterion routine; if Lertap requires these values, it will ask for them.)

4.6.1.2 Copy Scores column

This option takes all the scores found in a selected column of the Scores worksheet, and copies them to the first empty column found in the Data worksheet.

There are at least a couple of reasons why users want to do this. First, users may want to export the Data worksheet for use in another application, such as, perhaps, SPSS. However, before doing so, they'd like some of the columns in the Scores worksheet to be appended to the Data worksheet.

At other times, users may want to delete the Scores worksheet, perhaps simply to save disk space. Before doing this they'll sometimes copy one or two of the Scores columns to the Data worksheet.

4.6.1.3 Copy Data to DAT

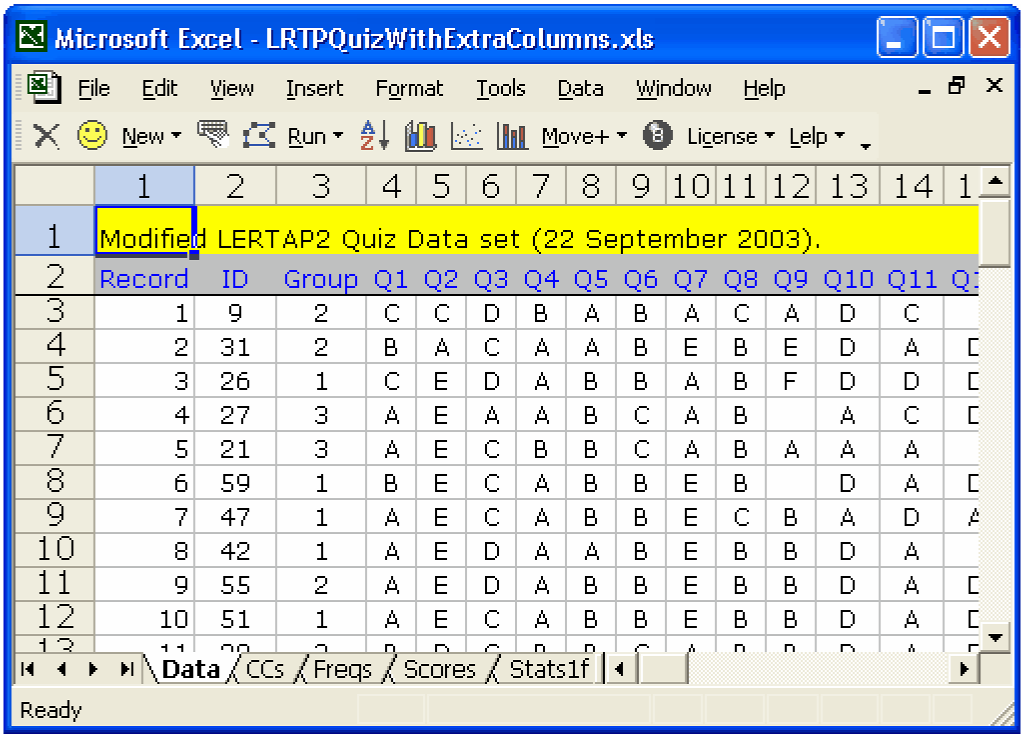
The DAT worksheet is a very special one. It's created when the option to get "item scores and correlations" is taken from the Run menu. The DAT worksheet is most likely to be used in conjunction with the Bilog and Bilog-MG computer programs.

When it's first created, the DAT worksheet contains just two real bits of information: some sort of record ID, and a string of zeros and ones representing item scores.

This is ordinarily sufficient for the Bilog program, but users of Bilog-MG often want to have more data in the DAT file. For example, they might want some sort of group identification code between the ID field and the string of item scores.

If the group identification code has been included in the Data worksheet, it may be copied over to the DAT worksheet using this option. In fact, any column in the Data worksheet may be copied to DAT. As columns are copied over, they're added after the ID field. Lertap makes an attempt to keep track of the format of the DAT records by including a Fortran format statement at the top of the DAT worksheet. This statement will usually not be adequate for Bilog-MG -- it's meant to be used as a guide.

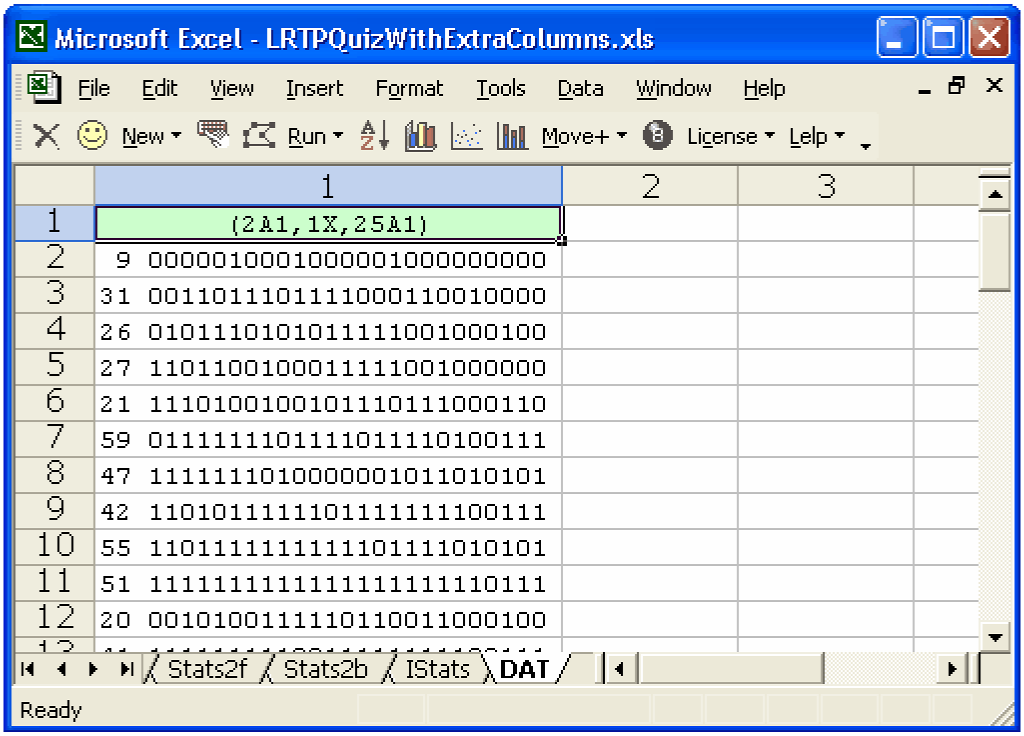
Let's look at an example, a data set having a Data worksheet as captured here when running with an earlier version of Excel:



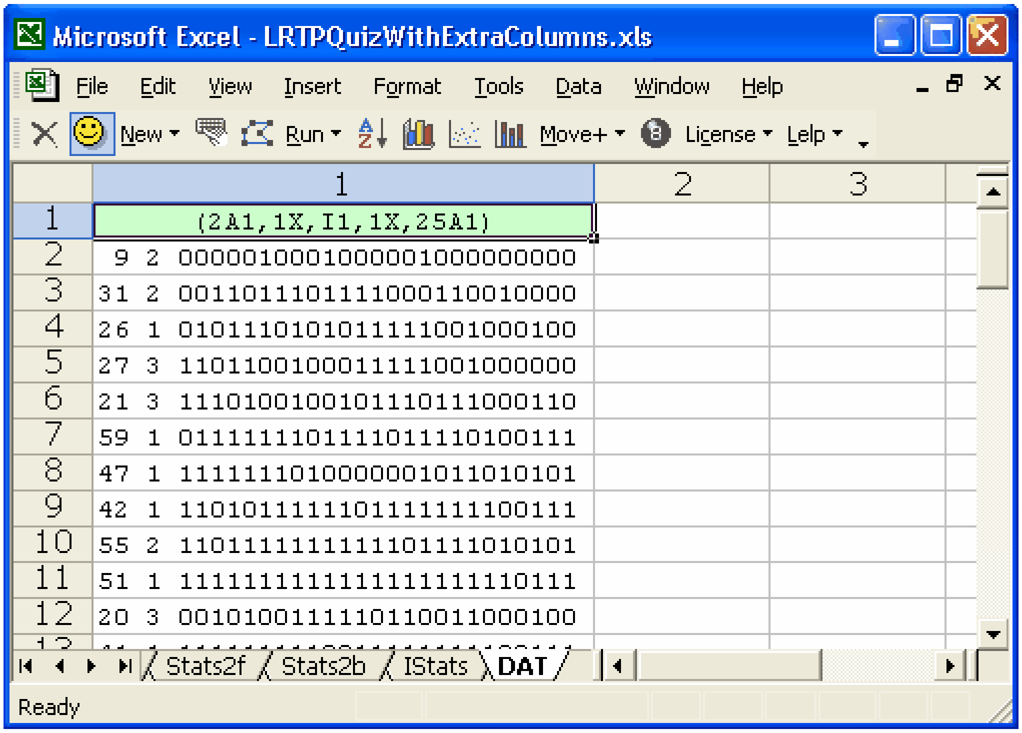
The Data worksheet has ID information in its second column, some sort of Group code in column 3, and item responses starting in column 4.

There were 25 cognitive items, each scored on a right/wrong basis, with one point for the right answer.

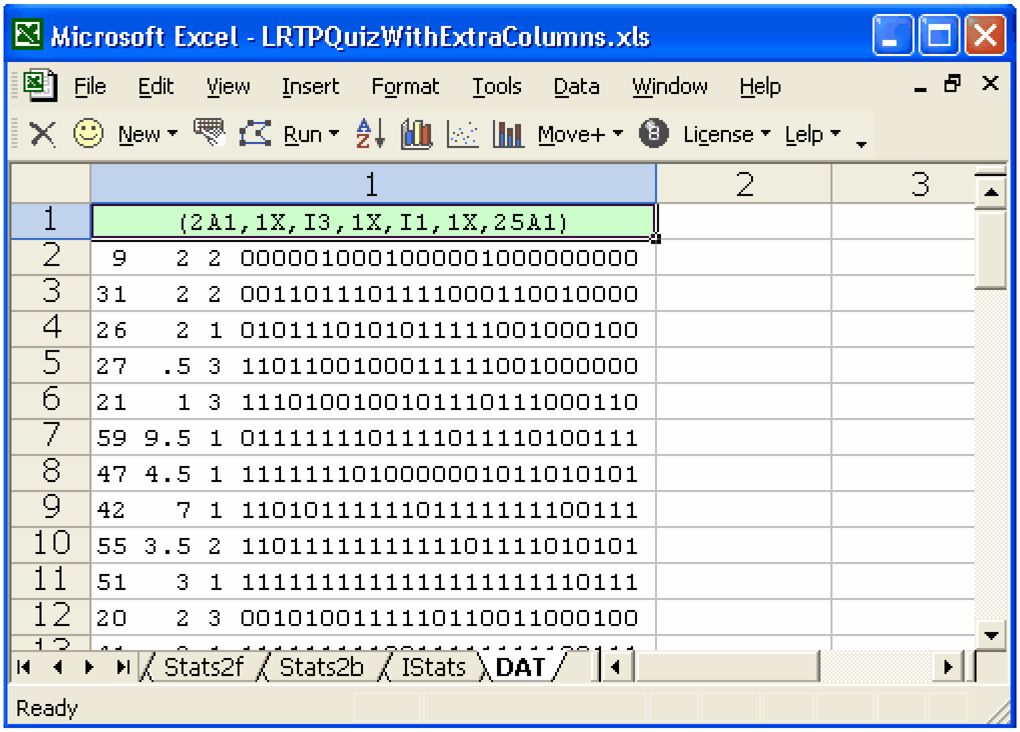
After going through the usual process of using the Run menu to "Interpret CCs lines", then "Elmillon item analysis", then get "item scores and correlations", the DAT worksheet looked like this:



The Move+ menu was then used to "Copy a Data column to the DAT worksheet" as we wanted to bring over the Group code from column 3 of the Data sheet. After doing this, the DAT file changed, as shown below:



Notice how the Fortran format statement has changed? In this case the statement is in fact correct, that is, as far as Bilog-MG is concerned. However, as more columns are copied over from the Data worksheet, the format statement continues to use "I" as a field identifier, which may or may not be correct as far as Bilog-MG is concerned. For example, we brought over a numeric column from the Data worksheet, prompting the DAT file to look like this:



Now the Fortran format statement is no longer correct. What Lertap has called an "I3" field should be "F3" -- in Fortran, "I" is used to denote a field containing an integer; "F" is used to denote a real number with a floating decimal. But this ain't a real problem. When you save the DAT worksheet as a text file, you'll end up deleting the Fortran format statement, and Lertap's minor *faux pas* will go unnoticed.

How to save the DAT sheet as a text file? Just click here, and read on.

4.6.1.4 Recode Data column

Suppose you had a column in the Data worksheet called "Gender", with entries of F for female, and M for male. Then, suppose that, for some reason, you'd like to instead have a code of 1 for female, and a code of 2 for male.

The recode option would be for you.

Suppose you had a Data worksheet, with a column for "Country", with entries such as NZ, AU, CA, US, DO, and VZ. Then, suppose for some reason you'd like to create a new column, to be called "Language", with NZ AU CA and US to all be coded EN, with DO and VZ to be coded SP. (Apologies to CA residents who speak SP.)

The recode option would do the job for you.

Or, suppose you'd like to get group breakouts, with only CA and US selected. You could use Recode, entering a new code of "Exclude" for all records without CA or US, after which you'd go for those breakouts.

Finally, although you know how to use the *tst card to select only certain Data records, and the NumericFilter2 special macro, you'd like to just delete Data records with, say, SP in the Language column. The Recode option could do it.

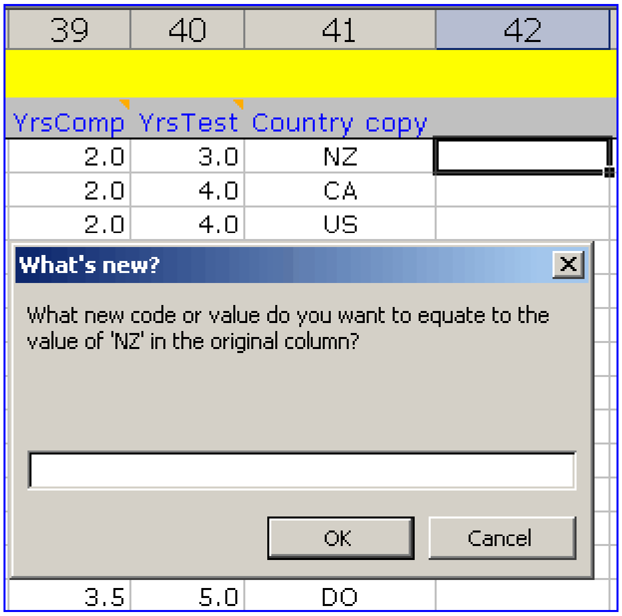
How it works

You go to the Data worksheet, and click on the Recode a Data column option under the Move+ menu.

You indicate the column which has the source data, that is, the column with the codes you want to work with.

The Recode macro copies the column to the far right side of the worksheet. Note the use of the word "copies" -- the macro does not delete or alter the original column in any way.

Then, the macro looks at the first entry in the copied column, that is, in row 3 of the newly-copied column. Let's say it finds a value of NZ. This little snap indicates what next happens:



The macro adds another new column to the worksheet, immediately to the right of the copied column.

It then asks you what NZ should become. You enter EN in the little 'What's new?' box, and click OK. All rows whose entry in the copied column is NZ will then have EN in the adjacent column, that is, in the recoded column.

After this the macro returns to look at the next entry in the copied column. It'll find CA (according to the little snapshot above). You'll be asked what CA is to equate to in the new column. You enter EN.

And so forth. We might end up with rows looking like this:

41	42
Country copy	Country recoded
NZ	EN
CA	EN
US	EN
NZ	EN
VZ	SP
VZ	SP
VZ	SP
VZ	SP
VZ	SP
VZ	SP
AU	EN
DO	SP

Got the idea? It's pretty simple (which will reflect poorly on you if you didn't get the idea).

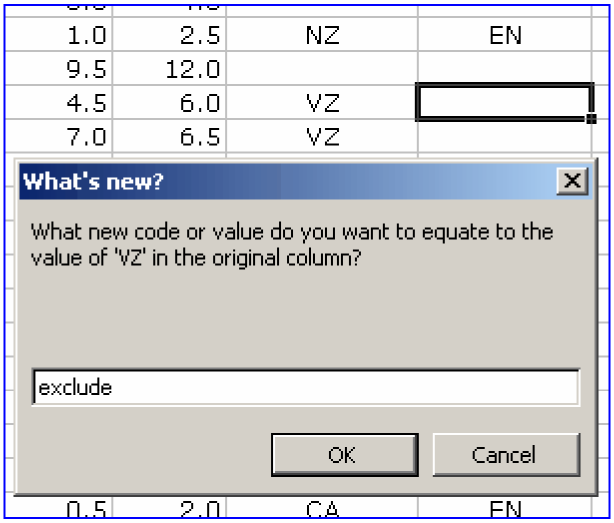
At the end of this process, the Recode macro leaves you with two new columns: the copy of the original column, and a column with the recoded values. It's easy to see if the macro has done what you intended -- just scan down the columns. If the result is not what you had in mind, just delete these two columns, and start again.

On the other hand, if the result is in fact what you wanted, you might then want to delete the column which has the copy of the original column. You don't have to do this, but if you do you'll save a bit of space. There are, after all, only a certain number of columns which a worksheet may have (256 was the limit in earlier versions of Excel, and this limit could be an issue at times -- however, with Excel 2007, you can have over ten thousand columns!).

Astute readers might have their hands up at this point: *You said you wanted to have a new column called "Language", but instead you have "Country recoded" at the top of the new column. You haven't finished, have you?*

Correct (ho-hum). We're left with the back-breaking task of typing 'Language' into the cell which presently has 'Country recoded'.

If your ultimate objective is to get group breakouts, and you'd like to exclude all records with DO or VZ in the original column, you'd respond thusly:



You don't have to type the whole word; just 'ex' will do. When the breakouts routine runs, it will ignore all rows which have been excluded in this manner.

Similarly, if there are records you want to delete, enter the word 'delete' in the little box, or just 'del', without the apostrophes. Lertap will set about deleting rows from the Data worksheet once the Recode macro has worked completely down the original column.

Please note that Lertap will say No-No! if you're asking for records to be deleted from Data when your workbook also has a Scores worksheet. There's a very critical correspondence between the Data and Scores worksheets, and Lertap tries its best to see that this correspondence is not disturbed.

Herewith all the usual warnings about deleting records from Data: you cannot recover them. Best to make a copy of the workbook before deleting records, something you can do by using the New menu.

If you click on OK without entering anything in the little box, the Recode macro will use whatever value you last entered. This makes it a bit easier to apply the same new code multiple times.

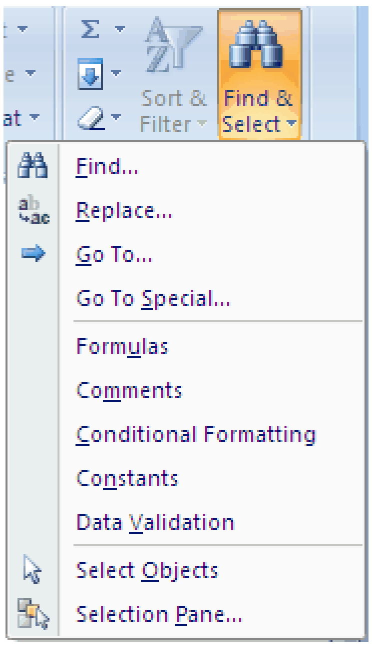
If the original value is blank, or empty, then, to maintain the blank, press your keyboard's space bar once, and then click on OK. Otherwise, if you don't want blanks in the new column, simply enter something in the little box, and, in the blank of an eye ....

Finally: as you may know, Excel has its own recode facilities, and they're quite respectable. If you page forward to the next topic you'll see.

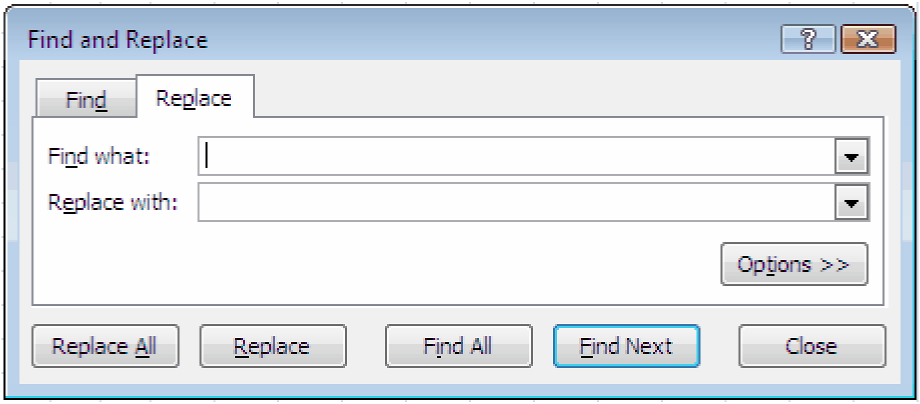
4.6.1.4.1 Excel's recoder

The Recode macro described in the preceding topic is not really a recoder. It doesn't alter the contents of the original Data column; instead it copies the indicated Data column, and then lets you create a new column with values, or codes, based on those found in the original column. This is much more along the lines of creating a "new variable", or of "transforming" an original variable to a new one (to use terms which may be familiar to SPSS users).

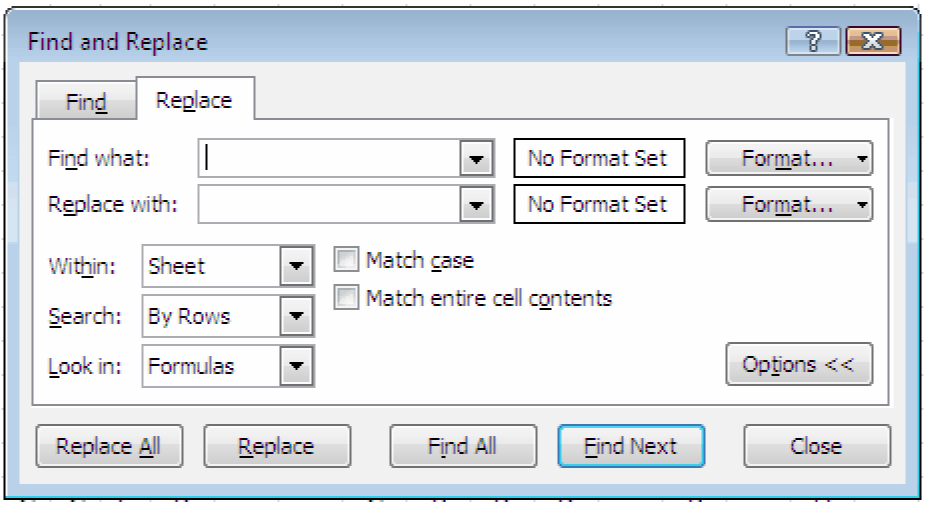
Excel 2007 has Find & Select options under the Editing section of the Home tab (look way to the right-hand side of the Home tab).



Click on Replace, and the following dialog box will pop up.

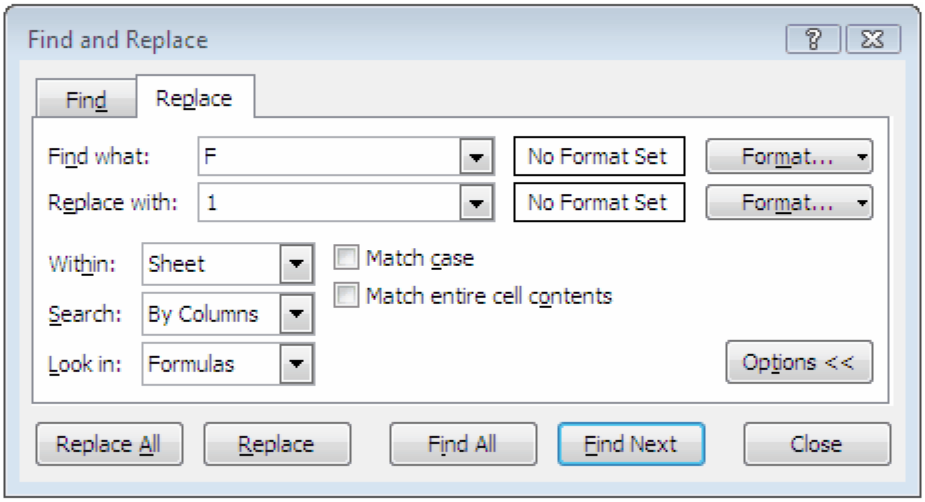


The Options>> button is what data recoders will want to use:



To give an example of using Excel to recode a column, let's say that we wanted to change every occurrence of 'F' in a given column to '1'.

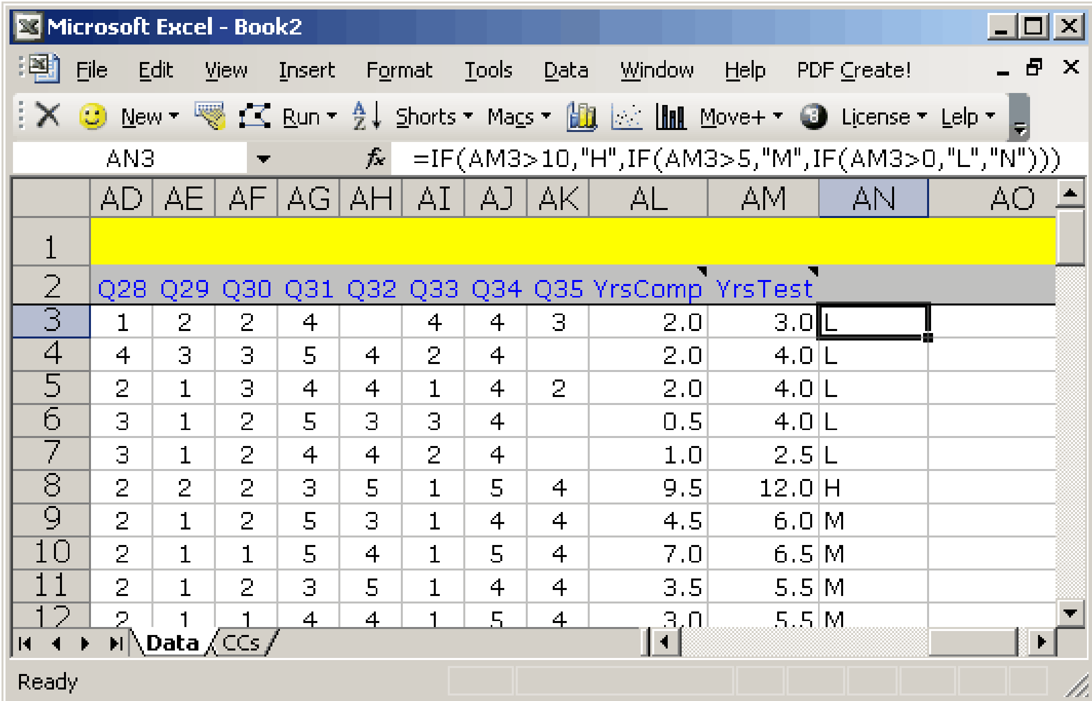
To do so, we'd select the column, and then fix up the dialog box so that it looks something like this:



If we dared to click on Replace All, Excel would dutifully find and replace all Fs with 1s in the selected column. We could have selected more than one column -- this is a quick and effective way to truly recode values over a number of columns.

It is very possible to get Excel to do more. We might have a column with numeric values, such as "number of years of test experience", and wish to create a new column with a coded experience letter, such as "H" for high, "M" for medium, "L" for low, and "N" for none.

Take a deep breath and look below at the =IF statement seen in the Formula Bar:



The =IF formula does the recoding for us. It says that, if YrsTest (column AM) has a value greater than 10, then column AN is to have an "H", otherwise, if column AM's value is greater than 5, then column AN is to show "M", otherwise, if column AM's value is over 0 (zero), show "L", else show "N".

(Note that we used Lertap's Excel shortcuts option to "Change the referencing style" so that the column headings are letters, not numbers -- this can often make writing Excel formulas easier.)

This looks very, very IFfy, you say? Well, among all the nice things we might say about Excel, one is that there are lots of resources to turn to when help is needed. You can try Excel's Help -- look up 'create conditional formulas' for some very helpful hints and examples. Or, try the internet. Or the local bookstore (perhaps there's now *Recoding Excel Columns for Dummies*!). Or even try us at: lertap5@gmail.com.

4.6.1.5 Apply a formula

Lertappers sometimes want to create a new score by transforming or combining one or more of the scores found in the Scores worksheet.

For example, let's suppose that a user wanted to apply a linear transformation to one of Scores' scores, of a type commonly found in texts and reference books:

$$y = mx + b$$

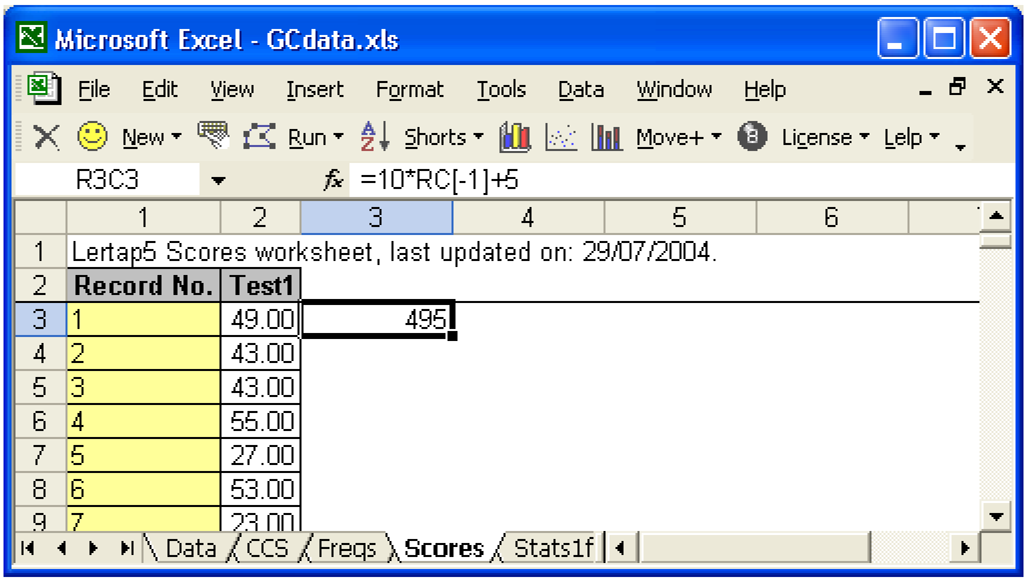
Here, "x" is a score which already exists, while "m" and "b" are constants. The new score is "y".

Let's suppose that m=10, and b=5, making the equation

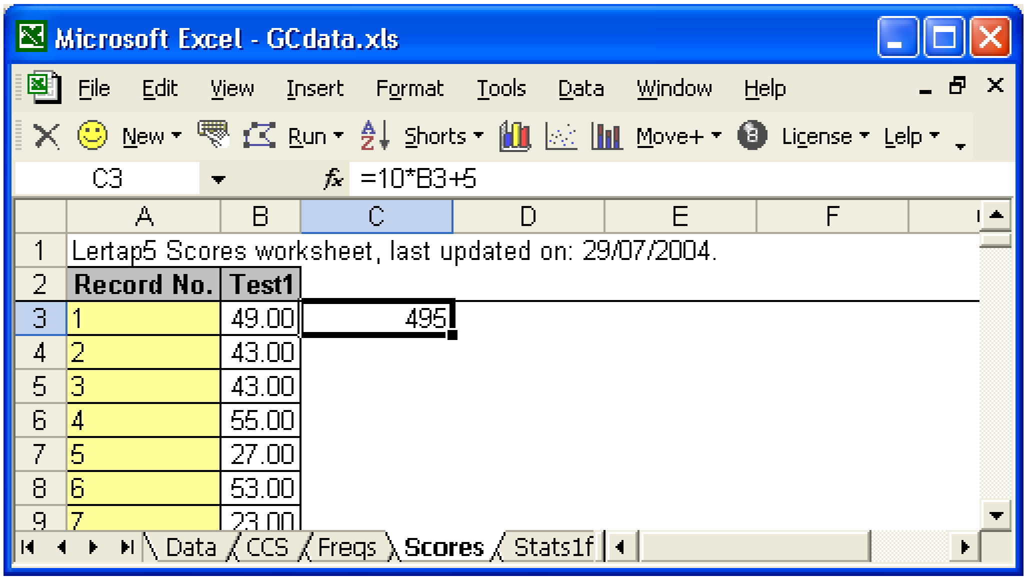
y = 10x + 5

Have a squiz now of the workbook below (taken when running an earlier version of Excel). The original score, "x" in the equation above, is called Test1, found in Scores column 2. Get out your best glasses, and look carefully at the Excel Formula Bar, the one which begins with R3C3, and contains a formula which a user entered: =10*RC[-1]+5.

The RC[-1] is Excel's way of referring to a value found in the same row (R), one column to the left (C[-1]).



Not everyone likes to work with the R1C1 Excel referencing style seen above in the Formula Bar. Many users click on Lertap's Excel shortcuts to change the referencing style so that columns are labeled with letters, as shown below:



Notice how the formula has changed to =10*B3+5? B3 refers to the cell where "x", the original score, is found.

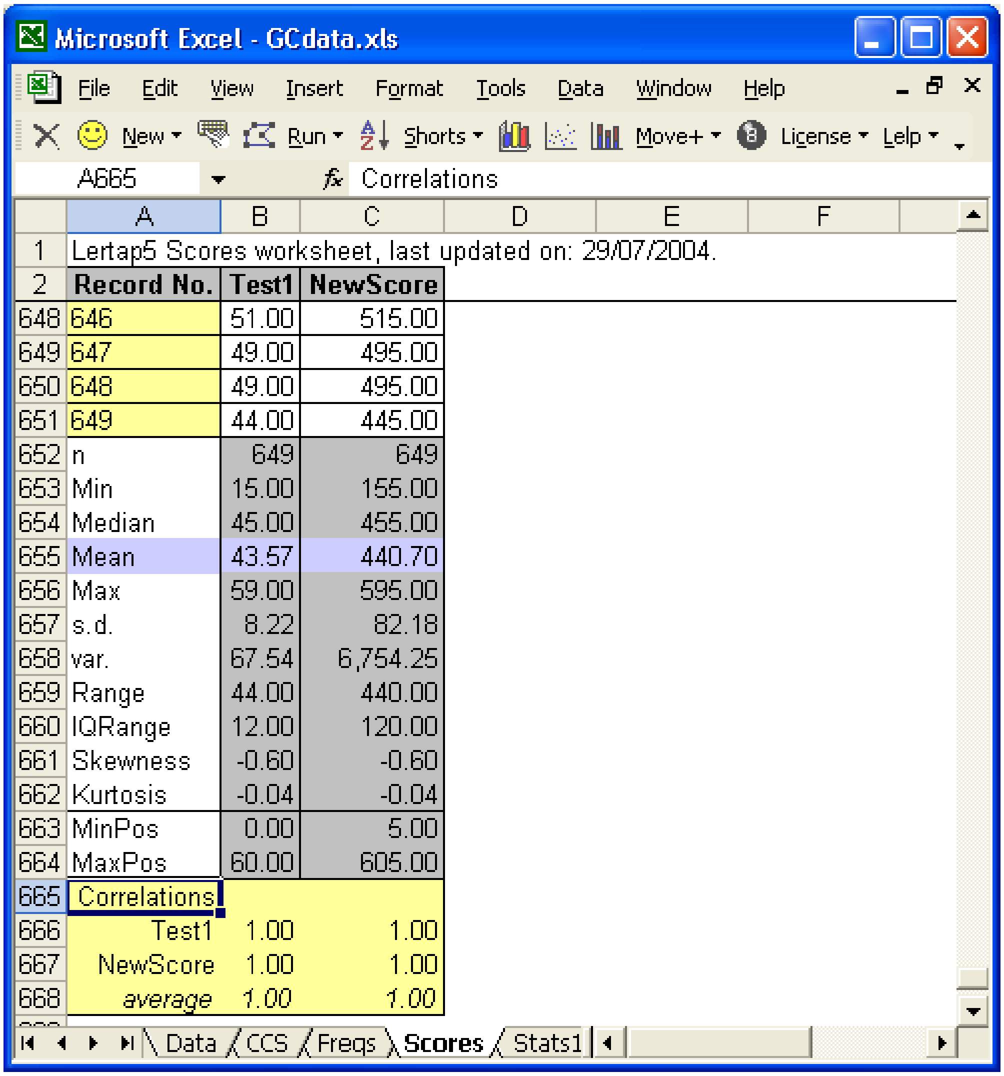
Both of these formulas (formulae) say the same thing -- they just use different referencing styles (you can pop back and forth between the referencing styles as much as you wish).

Okay? We've got a user who wants to make a new score by multiplying the original score, Test1, by 10, and adding 5 to the result.

S/he begins by selecting the cell immediately to the right of the first Test1 score, and then enters the formula by actually typing it in, starting with the equals (=) sign.

After typing the formula, the user presses the <Enter> key, and Excels displays the value of the new score, which in this example is 495.

If this is what's wanted, the user then selects the cell with the new formula by clicking on it, cruises up to the Move+ menu, and clicks on "Apply a special Scores worksheet formula". Your beloved little Lertap then applies the formula to all other original scores, determines the descriptive statistics related to the new score, and updates the correlation matrix found at the bottom of the Scores worksheet, lo:



The label given to the new score, NewScore, may of course be changed.

Don't like the results? Select the new score's column, delete it, and start again.

Want to know more about working with formulas in Excel? Look for assistance in Excel Help (there's lots -- you might start by searching Excel Help for "create a formula").

4.6.1.6 Update correlations

This option refreshes part of the Scores worksheet -- the part where the correlations appear, which is always at the bottom of the worksheet.

You might want to use this option after you have deleted one of the columns in the Scores worksheet. When a Scores column is deleted, the correlations become messy, with Excel tending to display REF# messages, or something equally ugly. Updating the correlations via this option will clean things up.

Remember that deleting rows in the Scores worksheet is a big no-no. If a row is deleted, Lertap loses track of what's what in terms of the data, and who's who.

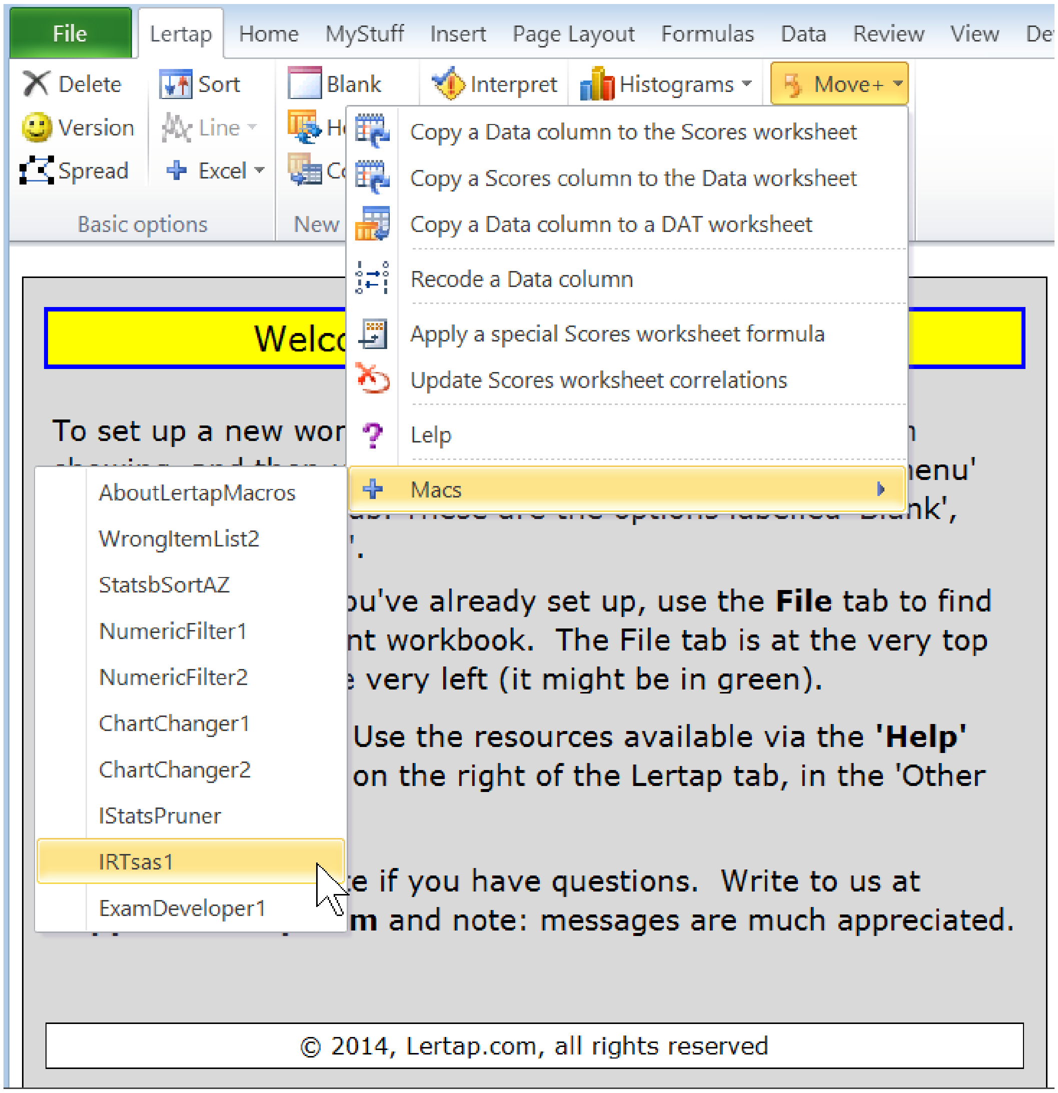
You may delete the whole Scores worksheet, and you may delete one or more columns of scores without affecting Lertap's internal data structure. But don't delete any of the rows, okay? Promise?

4.7 Macs Menu

There are quite a number of things to say about the Macs menu, foremost among them: this is a very volatile topic. Many of the screen shots seen in this topic, and those immediately following, may be out of date when you get around to comparing them with what you might see on the version of Lertap you've got running at the moment.

The reason for this? The macros accessed via the Macs menu are wide open, and subject to change at almost anytime.

But let's start at the start. The way to get to the Macs menu in the Excel 2007/2010/2013/2016/365 versions of Lertap is to click on Move+, and then on the + Macs line. The Macs menu will be similar to this (in this example, the cursor arrow is shown pointing to the IRTsas1 option):



Keep in mind that your Macs menu might not look like this, okay? We may have made some changes and not yet updated this topic's displays.

The names of the options seen under the Macs menu are the names of macros. Above, the names start with "AboutLertapMacros", and end with "ExamDeveloper1".

Macros? Whatsit? If you don't know what a macro is, open Excel Help, and toss it a search term such as "about macros", "using macros", or "creating macros". Macros are based on computer code; they're special-purpose computer programs, often quite brief, designed to accomplish something not presently built into Lertap itself. They're somewhat akin to Excel "Add-ins" -- they add special little features and capabilities.

For your information, the macros accessed via the Macs menu are stored in a special Excel file called Lertap5MacroSetA.xlam.

There will be more macros available than those displayed in the Macs menu. This topic explains how to change the macro names displayed in the list.

How to get the Lertap5MacroSetA.xlam file? You need not worry -- when Lertap starts up in Excel 2007/2010/2013/2016/365 this macro file is automatically summoned behind the scenes. (See this topic for more information.)

How to tell what a macro does? Click on it and see what it says; it won't hurt you, it won't actually do anything until you read its wee introduction, and click on some sort of 'run me!' option (also, see one of the titbits below--some macros are discussed in this document).

It is entirely possible to change the macros, and even to add your own. Interested? Look at the paper referenced under the titbit below. The ability to add your own home-brewed macros to Lertap opens up an expansive horizon of possibilities.

If you have a serious interest in tailoring Lertap macros to match local needs, do one or all of these things: read the titbit below; read this topic; write to us at larry@lertap.com. If you have an interest in a special macro, but don't have time to make it, write to us to see if we might take on the task. If your idea is one which could be of benefit to other users, there's a fair chance we'd write the macro for free, especially if it's winter, too cold to do much outside, there's no cricket or baseball to watch, and petrol is still too costly to drive north to the tropics (to people in the southern hemisphere, the tropics lie in a northerly direction).

Related titbits:

A paper which has more on the development of macros may be admired at: AngelLearningLertapMacros1.doc (a Microsoft Word file, about 170 KB). This paper has quite a bit to say about how to get Lertap to use data files created by Angel Learning. A more recent document discusses macros used for working with data from Pearson VUE.

These macros are discussed in this document: *BinaryItems, ChartChanger1, ChartChanger2, ChartChanger4, difR, EIRTanalysis1, Exam Developer, ImportCSV, ImportCSV2, IRTsas1, IStatsPruner, NumericFilter2, Omega1, PearsonVUE, QuantileShader1, RaschAnalysis1, StatsbSortAZ, ScrunchBoss, TotalTest1, WrongItemList2,* and *ZoomAllSheets*.

This topic discusses how to change the list of macros seen in the Macs menu.

4.7.1 EIRTanalysis1

Support for an Excel "Add-In" called EIRT was added in July 2015. This add-in was initially created in 2006 and has been updated regularly since then by its three authors. EIRT has been available on "Source Forge" for some time, where equivalent IRT routines for use with the R language are found in a package called RIRT .

Here's a description from EIRT's Readme.doc file:

> This add-in for Excel provide a 'wizard' to estimates the items and latent variables from the responses of subjects to a questionnaire. The supported dichotomous IRT (Item Response Theory) models are the 1PLM (one parameter logistic model), the 2PLM (two parameter logistic model) and the 3PLM (three parameter logistic model). For polytomous items, Bock's nominal model, and Samejima's graded model are supported. Two nonparametric methods (kernel regression an PMMLE) are also supported. The estimations methods available are the MMLE (marginal maximum likelihood estimator) and the BME (Bayes modal estimator, for dichotomous models only) for the parametric estimation of items, the PMMLE (Penalized MMLE) and the kernel regression (Nadaraya-Watson) for the nonparametric estimation of items, and the EAP (expectation a posteriori) and WMLE (Warm's maximum likelihood estimator) for the latent variable predictions.

From the description, you might sense that English is not the native language of the authors (in fact it's French). You'll get the same impression when running EIRT, but you shouldn't be put off -- EIRT blends in nicely with Lertap 5, and has power.

Read much more about it in this document.

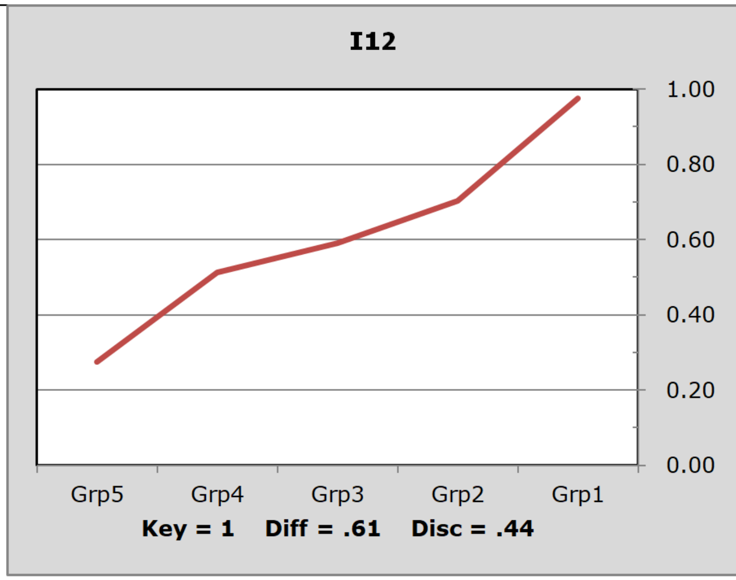
4.7.2 Binary items

The term "binary items" refers to test items which have been scored (usually) on a right-wrong basis. Very often a zero is used to indicate the item "score" for a wrong answer, with an item score of 1 (one) used for right answers.

This type of item scoring is also referred to as "dichotomous scoring", and items scored in this manner are often referred to as "dichotomous items".

The analysis of dichotomous test items is a common task undertaken in "IRT", item response theory. The purpose of the BinaryItems macro is to assist IRT users by making it possible to obtain simple quantile plots which reflect how the proportion of items-correct varies by "proficiency" level, where proficiency refers to a measure of ability, or of knowledge.

The following quintile plot traces the proportion of correct answers in each of five groups, ranging from "Grp5" to "Grp1", with Grp5 composed of the weakest students, and Grp1 the strongest.



I12 (c13) *alpha with & without: 0.935 0.934*

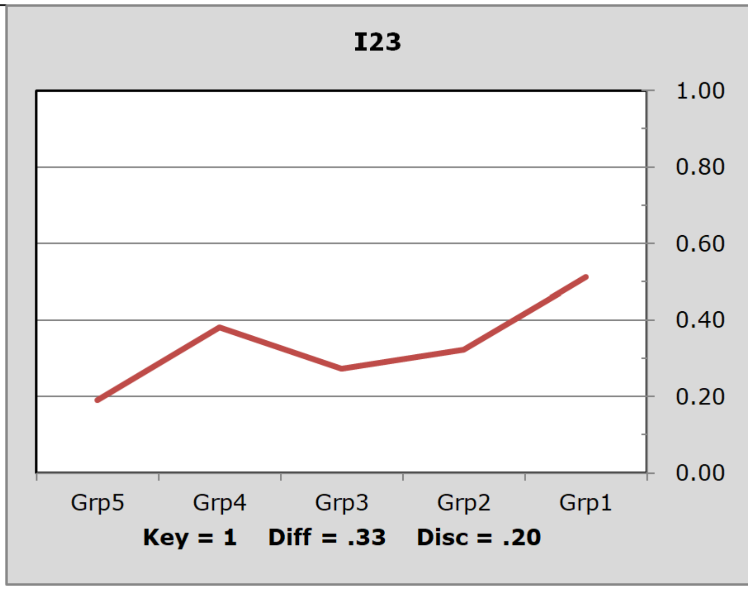
option	wt.	n	p	pb(r)	b(r)	avg.	z
0	0.00	165	0.39	-0.46	-0.58	35.08	-0.58
1	1.00	259	0.61	0.44	0.56	51.75	0.37

These results for I12 are from the "HalfTime" sample dataset. The five groups were defined, in this case, by using the total test score obtained by 424 postgraduate students of statistics. Total test scores were sorted from low to high, with Lertap getting Excel to pick off the lowest 20% for Grp5, the second-lowest 20% for Grp4, ..., and the highest 20% for Grp1.

The plot indicates a positive relationship between the proportion of correct answers and "proficiency". The stronger the student group, the higher the proportion correct. Here the lowest group had a proportion-correct of just under 0.30 while the strongest students had a proportion-correct of almost 1.00.

In this case, "proficiency" was taken as the total test score. More generally, proficiency may be any sort of score. Lertap's support for "external criterion" scores is used when proficiency has been measured by something other than total test score.

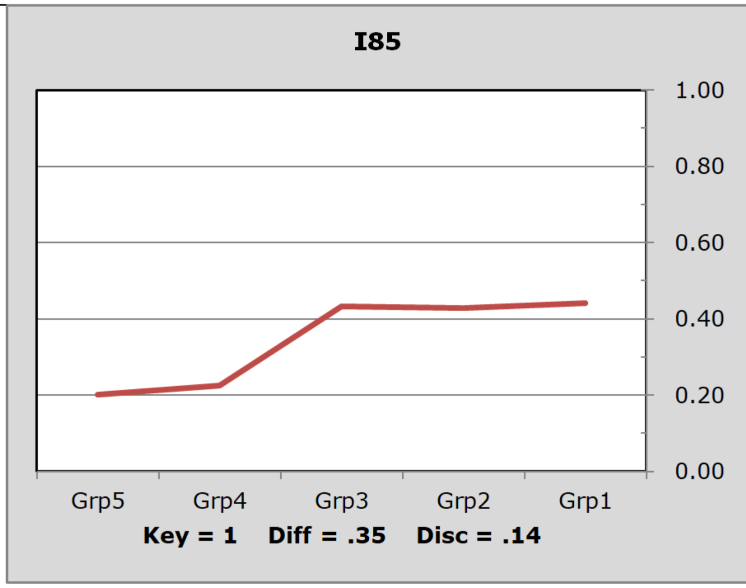
I12's plot displays the response pattern desired for discriminating test items, that is, items capable of identifying weak and strong students. Unfortunately, it is often the case that not all items in a test will display the desired pattern. The two items below have response patterns that are not all that uncommon.



I23 (c24) *alpha with & without: 0.935 0.935*

option	wt.	n	p	pb(r)	b(r)	avg.	z
0	0.00	282	0.67	-0.23	-0.30	42.38	-0.16
1	1.00	142	0.33	0.20	0.26	50.99	0.32

I23 was not a good discriminator. Some of the weaker students (Grp4) had a higher proportion-correct than those in Grp3 and Grp2.



I85 (c86) *alpha with & without: 0.935 0.935*

option	wt.	n	p	pb(r)	b(r)	avg.	z
0	0.00	277	0.65	-0.16	-0.21	43.16	-0.12
1	1.00	147	0.35	0.14	0.18	49.22	0.22

I85 was another item with poor discrimination. In this case Grp3, Grp2, and Grp1 all had about the same proportion correct and, in each case, it was below 50%.

Some readers may wonder if we might expect a relationship between plots such as those seen here and the success of an IRT model to fit the data (or the compatibility of the data with the IRT model)? Yes, definitely. We might expect I12 to "pass IRT muster", but probably not I23, nor I85.

How to use this macro

This macro assumes that you, the user, are looking at a conventional Lertap dataset involving the use of cognitive test items, such as, for example, HalfTime (a simpler example, with a more straightforward CCs worksheet, would be MathsQuiz).

It's assumed that you've applied the Interpret and Elmillon options, and also the "Item scores and correlation" option in order to get an IStats worksheet.

Once these assumptions are met, simply select the macro and let it do its job. It will create a new workbook with a Data worksheet comprised of item scores, and a CCs worksheet ready to use with Interpret and Elmillon.

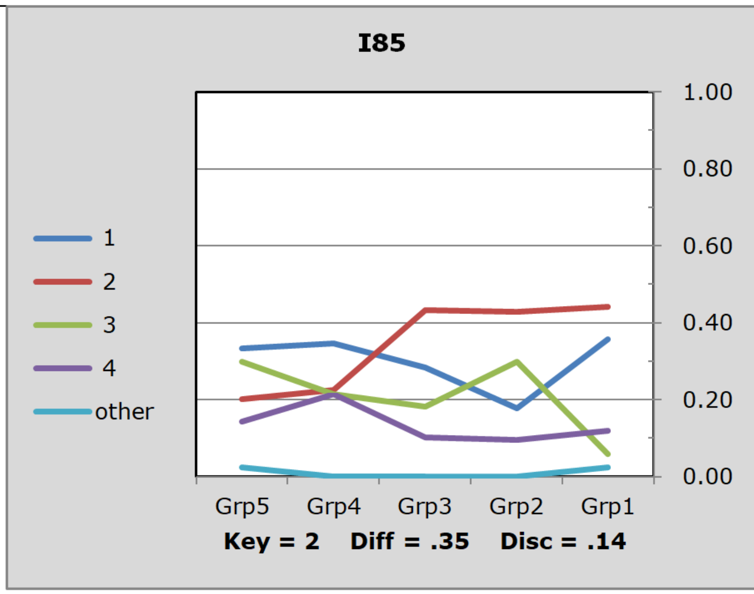
How to get the response plots

Open the Stats1ul worksheet produced by Elmillon, and then click on "Res. charts" option. This will make the Stats1ulChta worksheet with its series of quantile plots, one for each item. The initial plots will display three response trace lines: one for wrong answers, one for correct answers, and one for "other". Clean these up by using the ChartChanger4 macro to get plots like those above.

When this macro will not be needed

Using this macro, in conjunction with the ChartChanger4 macro, is nothing more than a bonus activity designed to make it easier to see how items are discriminating, and judge how well they might fare when submitted to an IRT analysis.

Compare, for example, the above plot for I85 with this one, a conventional quintile plot for I85 which includes trace lines for the distractors:



I85 (c86) *alpha with & without: 0.935 0.935*

option	wt.	n	p	pb(r)	b(r)	avg.	z	
1	0.00	127	0.30	0.03	0.04	46.06	0.04	<-aa
2	1.00	147	0.35	0.14	0.18	49.22	0.22	
3	0.00	89	0.21	-0.18	-0.25	39.21	-0.34	
4	0.00	57	0.13	-0.07	-0.10	42.28	-0.17	
other	0.00	4	0.01	0.04	0.14	51.75	0.37	<-aa

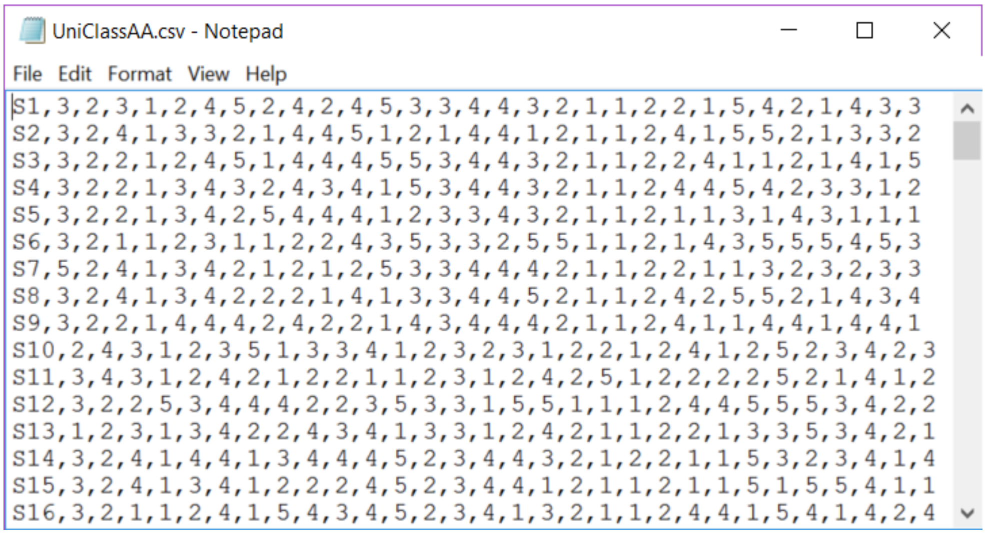
This plot also shows that I85 is a poor discriminator, but the plot itself might be regarded as a bit "noisy" in comparison. Related comments are found in the ChartChanger4 topic, and its subtopic.

4.7.3 ImportCSV

A csv file is a file which contains data values separated by commas. These files are often created by "OMR" scanners, such as those from Scantron.

This topic presents examples of two typical csv files, and then shows how the ImportCSV macro is used to convert them to an Excel workbook ready for use with Lertap 5. (Refer to this topic for a definition of "Lertap-ready".)

Example 1: introduction



UniClassAA.csv - Notepad

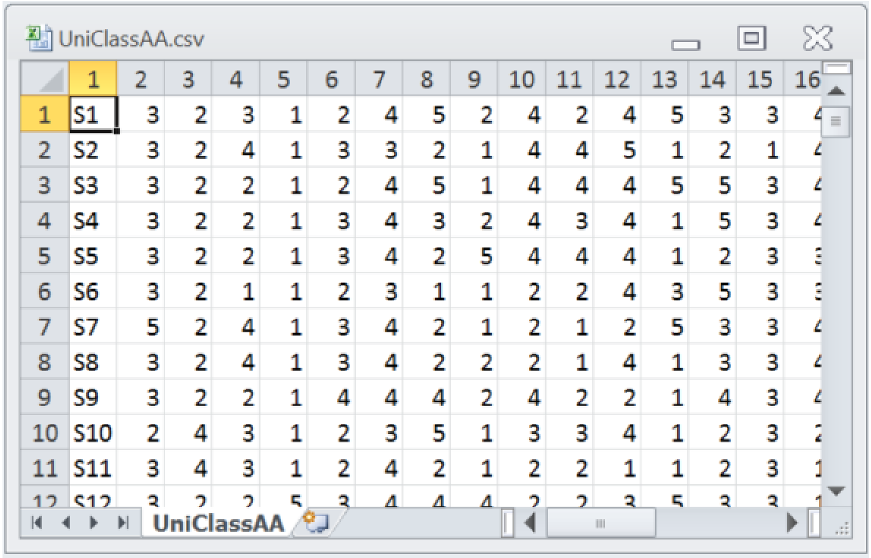
File Edit Format View Help

```
S1,3,2,3,1,2,4,5,2,4,2,4,5,3,3,4,4,3,2,1,1,2,2,1,5,4,2,1,4,3,3
S2,3,2,4,1,3,3,2,1,4,4,5,1,2,1,4,4,1,2,1,1,2,4,1,5,5,2,1,3,3,2
S3,3,2,2,1,2,4,5,1,4,4,4,5,5,3,4,4,3,2,1,1,2,2,4,1,1,2,1,4,1,5
S4,3,2,2,1,3,4,3,2,4,3,4,1,5,3,4,4,3,2,1,1,2,4,4,5,4,2,3,3,1,2
S5,3,2,2,1,3,4,2,5,4,4,4,1,2,3,3,4,3,2,1,1,2,1,1,3,1,4,3,1,1,1
S6,3,2,1,1,2,3,1,1,2,2,4,3,5,3,3,2,5,5,1,1,2,1,4,3,5,5,5,4,5,3
S7,5,2,4,1,3,4,2,1,2,1,2,5,3,3,4,4,4,2,1,1,2,2,1,1,3,2,3,2,3,3
S8,3,2,4,1,3,4,2,2,2,1,4,1,3,3,4,4,5,2,1,1,2,4,2,5,5,2,1,4,3,4
S9,3,2,2,1,4,4,4,2,4,2,2,1,4,3,4,4,4,2,1,1,2,4,1,1,4,4,1,4,4,1
S10,2,4,3,1,2,3,5,1,3,3,4,1,2,3,2,3,1,2,2,1,2,4,1,2,5,2,3,4,2,3
S11,3,4,3,1,2,4,2,1,2,2,1,1,2,3,1,2,4,2,5,1,2,2,2,2,5,2,1,4,1,2
S12,3,2,2,5,3,4,4,4,2,2,3,5,3,3,1,5,5,1,1,1,2,4,4,5,5,5,3,4,2,2
S13,1,2,3,1,3,4,2,2,4,3,4,1,3,3,1,2,4,2,1,1,2,2,1,3,3,5,3,4,2,1
S14,3,2,4,1,4,4,1,3,4,4,4,5,2,3,4,4,3,2,1,2,2,1,1,5,3,2,3,4,1,4
S15,3,2,4,1,3,4,1,2,2,2,4,5,2,3,4,4,1,2,1,1,2,1,1,5,1,5,5,4,1,1
S16,3,2,1,1,2,4,1,5,4,3,4,5,2,3,4,1,3,2,1,1,2,4,4,1,5,4,1,4,2,4
```

This example is from the Uni Class A sample dataset; 127 students answered 30 multiple-choice test items. On each item, response codes {1,2,3,4,5} were used.

Every line in the csv file begins with a simple student ID code, such as S1, followed by the student's answer to the first item, such as 3, followed by her/his answer to the second item, such as 2, and so on. There were a total of 127 lines in this csv file, one for each student. Each line had an entry, or a "value", for each of the 31 "fields" of data. Each field's value is separated by commas.

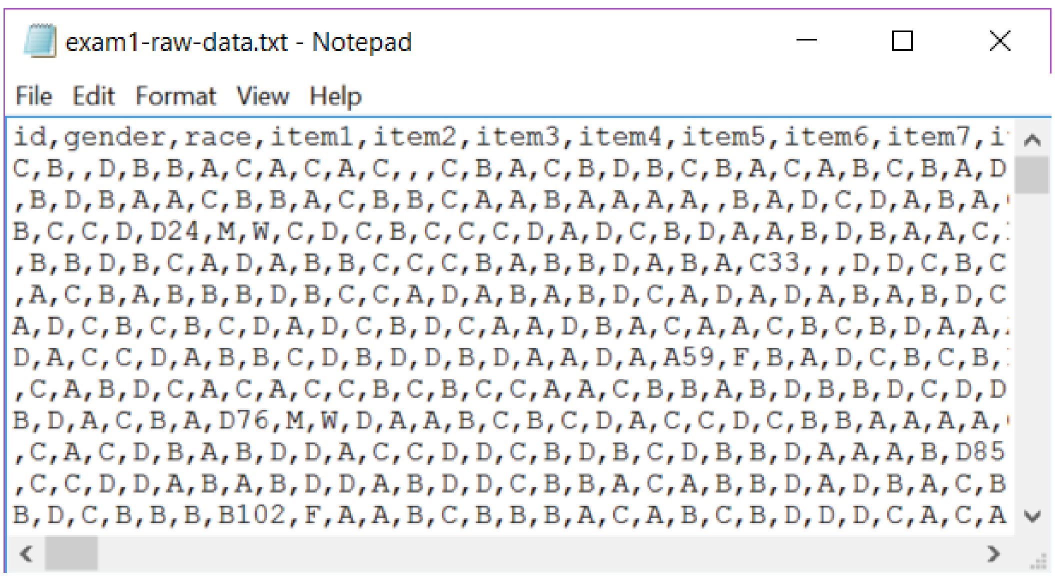
Note that the Windows "Notepad" app was used to open this file -- this is seen in the heading. Excel is used just as often as Notepad to open csv files. Here's how UniClassAA.csv looks when opened in Excel 2010:



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	S1	3	2	3	1	2	4	5	2	4	2	4	5	3	3	4
2	S2	3	2	4	1	3	3	2	1	4	4	5	1	2	1	4
3	S3	3	2	2	1	2	4	5	1	4	4	4	5	5	3	4
4	S4	3	2	2	1	3	4	3	2	4	3	4	1	5	3	4
5	S5	3	2	2	1	3	4	2	5	4	4	4	1	2	3	3
6	S6	3	2	1	1	2	3	1	1	2	2	4	3	5	3	3
7	S7	5	2	4	1	3	4	2	1	2	1	2	5	3	3	4
8	S8	3	2	4	1	3	4	2	2	2	1	4	1	3	3	4
9	S9	3	2	2	1	4	4	4	2	4	2	2	1	4	3	4
10	S10	2	4	3	1	2	3	5	1	3	3	4	1	2	3	2
11	S11	3	4	3	1	2	4	2	1	2	2	1	1	2	3	1

The "UniClassAA" Excel worksheet had 127 rows and 31 columns.

Example 2: introduction



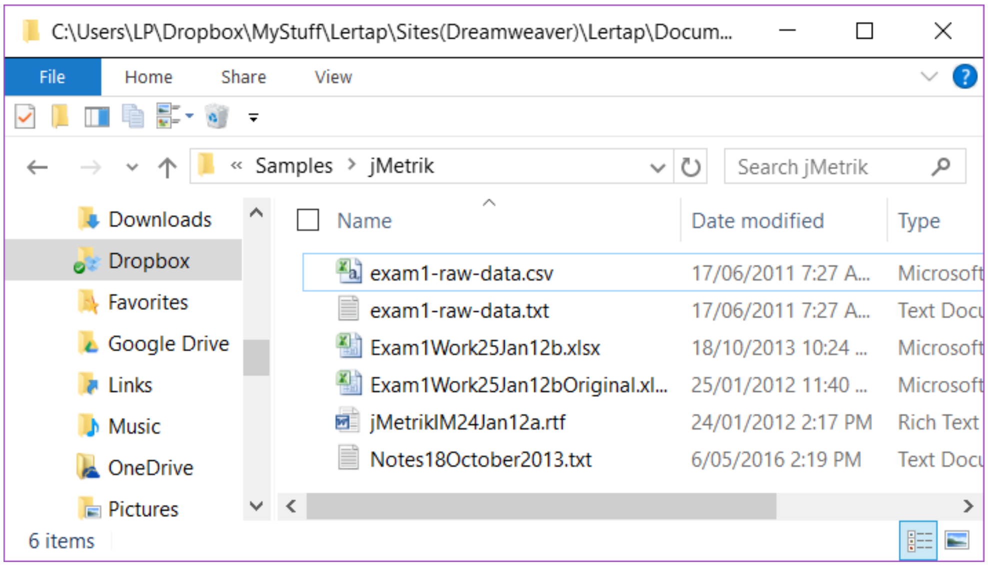

```
id,gender,race,item1,item2,item3,item4,item5,item6,item7,i
C,B,,D,B,B,A,C,A,C,A,C,,,C,B,A,C,B,D,B,C,B,A,C,A,B,C,B,A,D
,B,D,B,A,A,C,B,B,A,C,B,B,C,A,A,B,A,A,A,A,,B,A,D,C,D,A,B,A,
B,C,C,D,D24,M,W,C,D,C,B,C,C,C,D,A,D,C,B,D,A,A,B,D,B,A,A,C,
,B,B,D,B,C,A,D,A,B,B,C,C,C,B,A,B,B,D,A,B,A,C33,,,D,D,C,B,C
,A,C,B,A,B,B,B,D,B,C,C,A,D,A,B,A,B,D,C,A,D,A,D,A,B,A,B,D,C
A,D,C,B,C,B,C,D,A,D,C,B,D,C,A,A,D,B,A,C,A,A,C,B,C,B,D,A,A,
D,A,C,C,D,A,B,B,C,D,B,D,D,B,D,A,A,D,A,A59,F,B,A,D,C,B,C,B,
,C,A,B,D,C,A,C,A,C,C,B,C,B,C,C,A,A,C,B,B,A,B,D,B,B,D,C,D,D
B,D,A,C,B,A,D76,M,W,D,A,A,B,C,B,C,D,A,C,C,D,C,B,B,A,A,A,A,
,C,A,C,D,B,A,B,D,D,A,C,C,D,D,C,B,D,B,C,D,B,B,D,A,A,A,B,D85
,C,C,D,D,A,B,A,B,D,D,A,B,D,D,C,B,B,A,C,A,B,B,D,A,D,B,A,C,B
B,D,C,B,B,B,B102,F,A,A,B,C,B,B,B,A,C,A,B,C,B,D,D,D,C,A,C,A
```

This example is from a 56-item multiple-choice test taken by 6,000 elementary school students in the United States. The file was generated by an OMR scanner; it read the bubble sheets students used to indicate their answers, and created a "txt" file. There's more about this particular dataset in another topic.

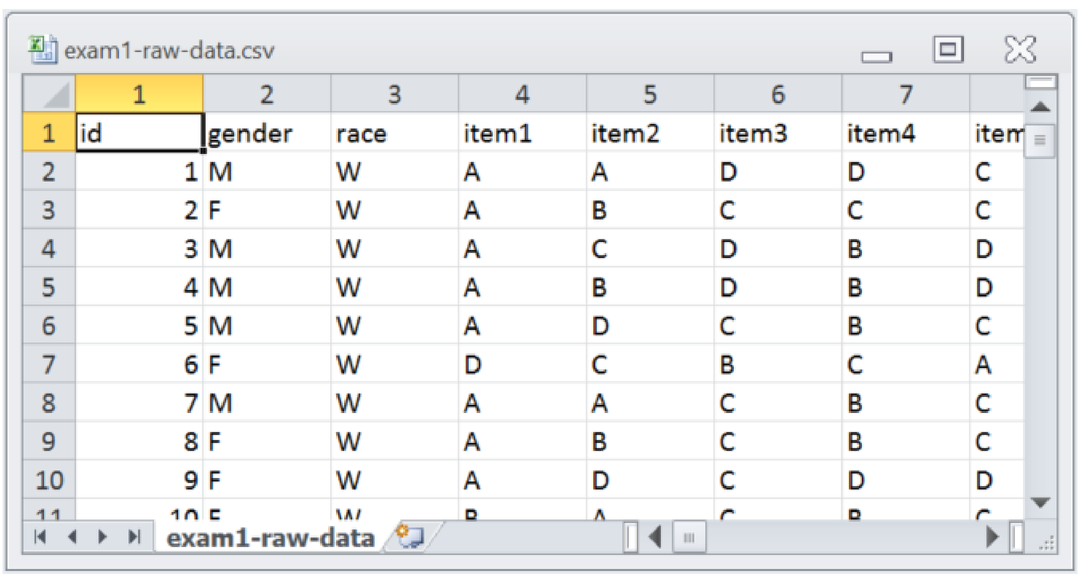
In this case, the instructor had the scanner insert a data line in the first row with the names (or titles) of the data fields found in each line of the txt file. It may be seen that the first field was "id", then "gender", then "race", followed by 56 fields for the item responses; the first of these fields was called "item1".

As mentioned, this file has a txt extension. Excel can open txt files by applying its "Text Import Wizard"; an example is provided here.

However, this file uses commas to separate each field -- to get Excel to open it without having to use the Text Import Wizard, all that's needed is to copy the file and change the copy's extension to "csv" instead of txt:



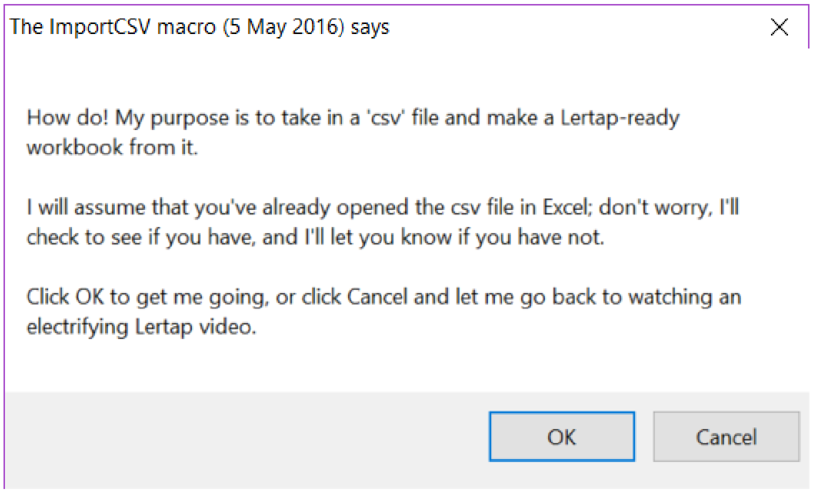
Double-clicking on the csv file above gets Excel to open it:



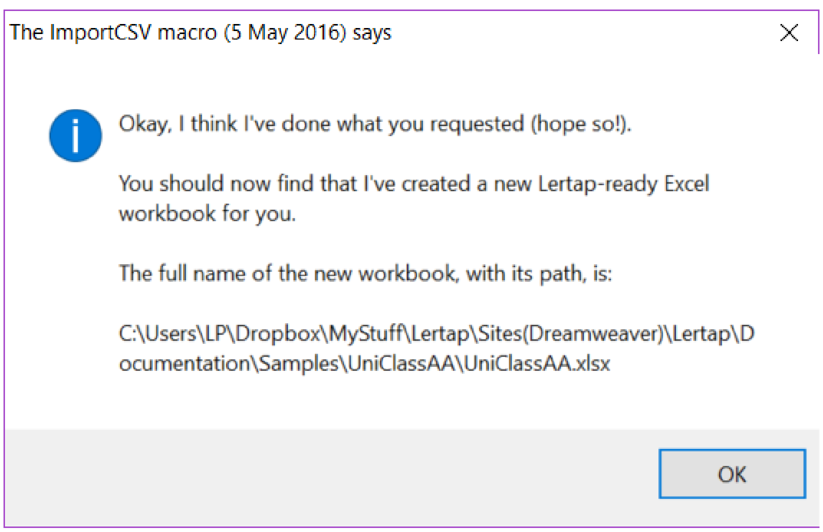
Okay. We've got two typical csv examples to work with. The presentation below shows how the ImportCSV macro is used to create a new Excel workbook for each example.

Using the ImportCSV macro with Example 1

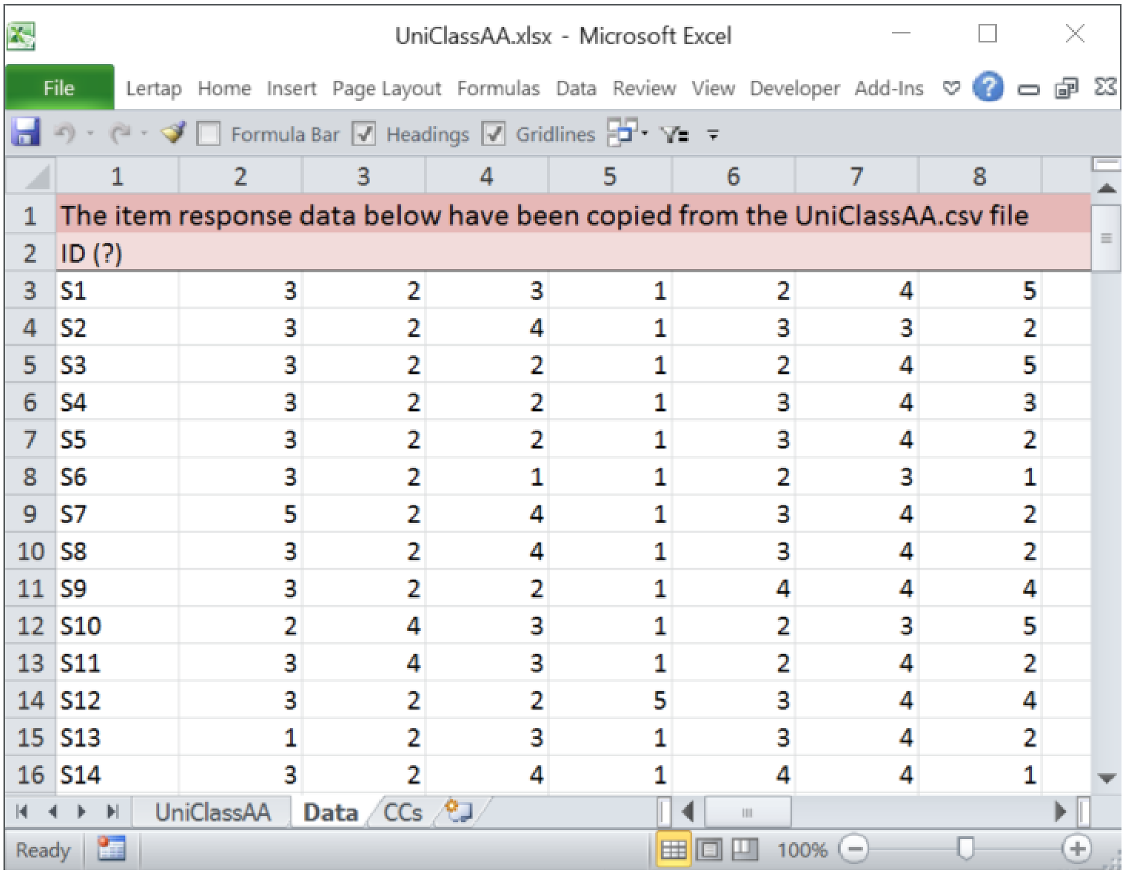
The ImportCSV macro is found under the Macs Menu. It displays the following message when it starts.



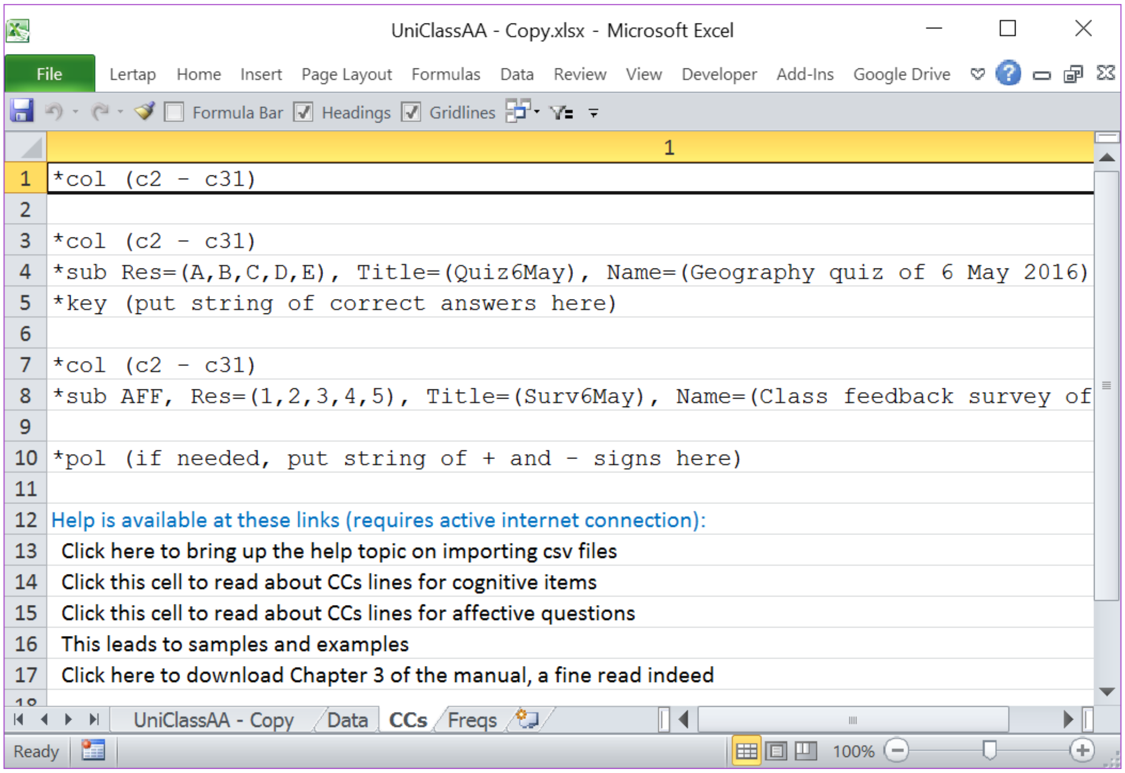
After the OK button is clicked, the macro works quite rapidly, ending with this message:



The macro will make a new Excel workbook and add two worksheets to it. The screen snapshots below show worksheet names in the tabs at the bottom of screen -- "UniClassAA" has the original data as brought in from the csv file; "Data" is a copy of UniClassAA but with two new rows inserted at the top.



	1	2	3	4	5	6	7	8
1	The item response data below have been copied from the UniClassAA.csv file							
2	ID (?)							
3	S1	3	2	3	1	2	4	5
4	S2	3	2	4	1	3	3	2
5	S3	3	2	2	1	2	4	5
6	S4	3	2	2	1	3	4	3
7	S5	3	2	2	1	3	4	2
8	S6	3	2	1	1	2	3	1
9	S7	5	2	4	1	3	4	2
10	S8	3	2	4	1	3	4	2
11	S9	3	2	2	1	4	4	4
12	S10	2	4	3	1	2	3	5
13	S11	3	4	3	1	2	4	2
14	S12	3	2	2	5	3	4	4
15	S13	1	2	3	1	3	4	2
16	S14	3	2	4	1	4	4	1



The Data and CCs worksheets created by the ImportCSV macro are just "drafts" -- they need a bit of work before they can really be used.

Of these two worksheets, CCs will require the most attention.

Row 1 in the CCs sheet above is saying that item responses begin in column 2 (c2) and end in column 31 (c31). A single *col line like this one, when followed by an empty row as above, is useful when all that's wanted from Lertap 5 is a "Freqs" report.

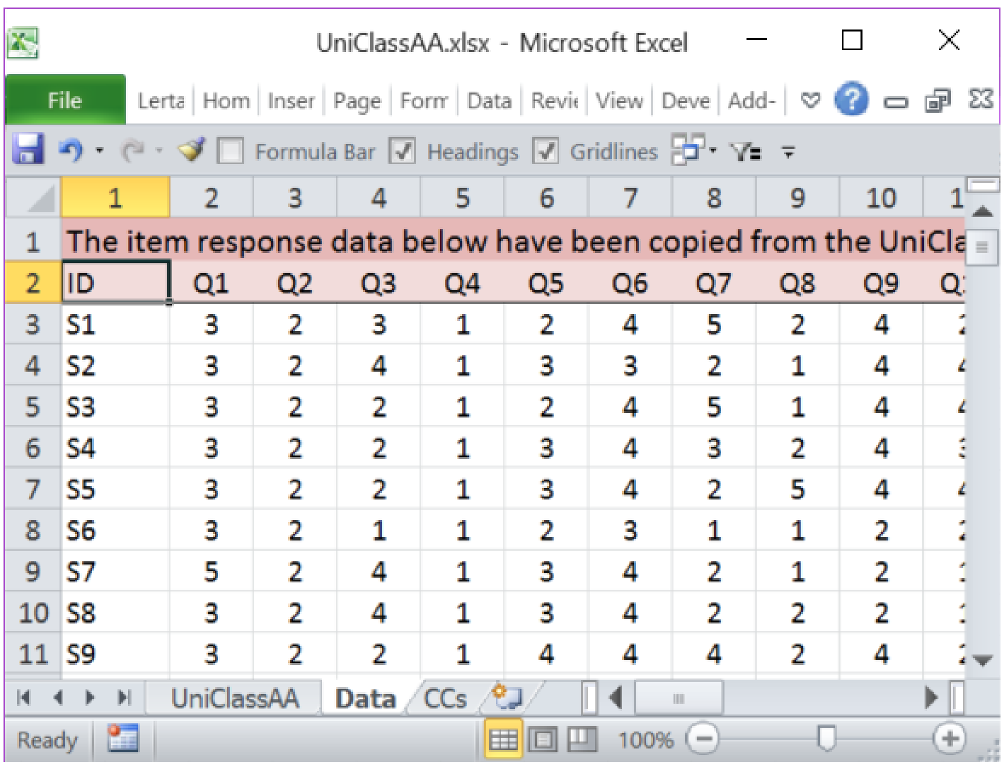
Given the CCs lines exactly as above, a Freqs report would be created if the Interpret option were selected, assuming that the *col line is accurate, that is, item responses are found in columns 2 through 31. If the Elmillon option were then taken, it would complain, saying that it cannot find a Sub sheet to work with. Elmillon will only produce the various Stats reports when the *col line is immediately followed by another line which begins with an asterisk.

If the first <u>two</u> rows in the CCs sheet above were deleted, the *col, *sub, and *key lines seen above in rows 3, 4, and 5 would then move up to rows 1, 2, and 3 and could be used to process the item responses in columns 2 through 31 as multiple-choice test items. But first the correct answers to each item would be required on the *key line, and the entries in the *sub line would undoubtedly require modification. Once these things are done, the Interpret option will create a Freqs report, and Elmillon will produce reports for a cognitive test.

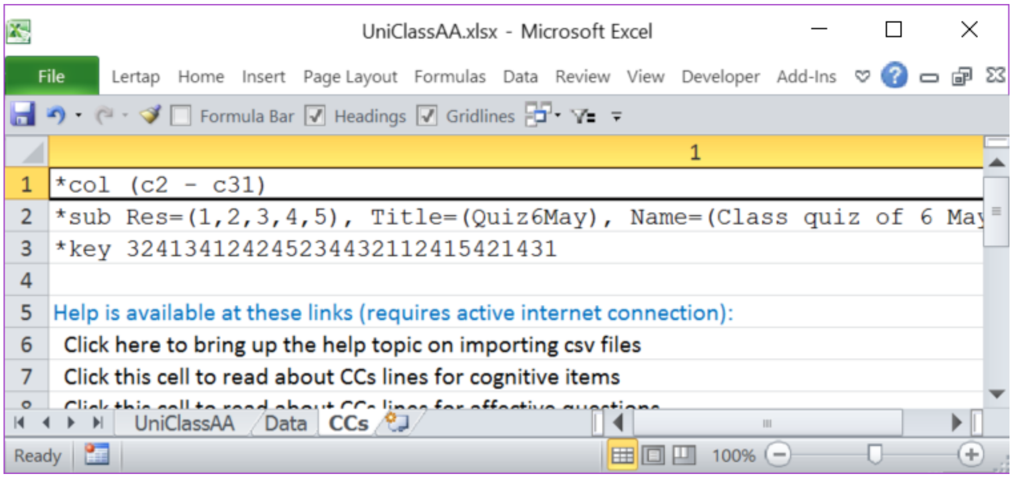
If the first six rows in the CCs sheet above were deleted, the *col, *sub, and *pol lines seen above in rows 7, 8, and 10 would move up to rows 1, 2, and 4 and could then be used to process the item responses in columns 2 through 31 as an affective scale (a survey, perhaps) . But first the Res= Title= and Name= entries in the *sub line would undoubtedly require modification. If some of the items were negatively worded, then the empty line between the *sub and *pol lines should be deleted, and +/- symbols added to the *pol line. Once these things are done, the Interpret option will create a Freqs report, and Elmillon will produce reports for an affective test.

(Note: Lertap 5 stops reading CCs lines whenever it encounters an empty line. The *pol line will be ignored if the line above it is empty, as is the case here. Deleting the empty line will get Lertap 5 to read the *pol line.)

The snapshots below display the Data and CCs sheets after they were fixed up for the Uni Class A sample .



The Data sheet in this example was easy to prepare as only the second row required modifications. An item label was inserted at the top of each column. (It's easy to do this by using Excel's option to "complete a series; fill data into adjacent cells by using the fill handle". Above, all that was required was "Q1"; the other item labels, from Q2 to Q30, were filled in simply by using the fill handle, dragging to the right from Q1.)

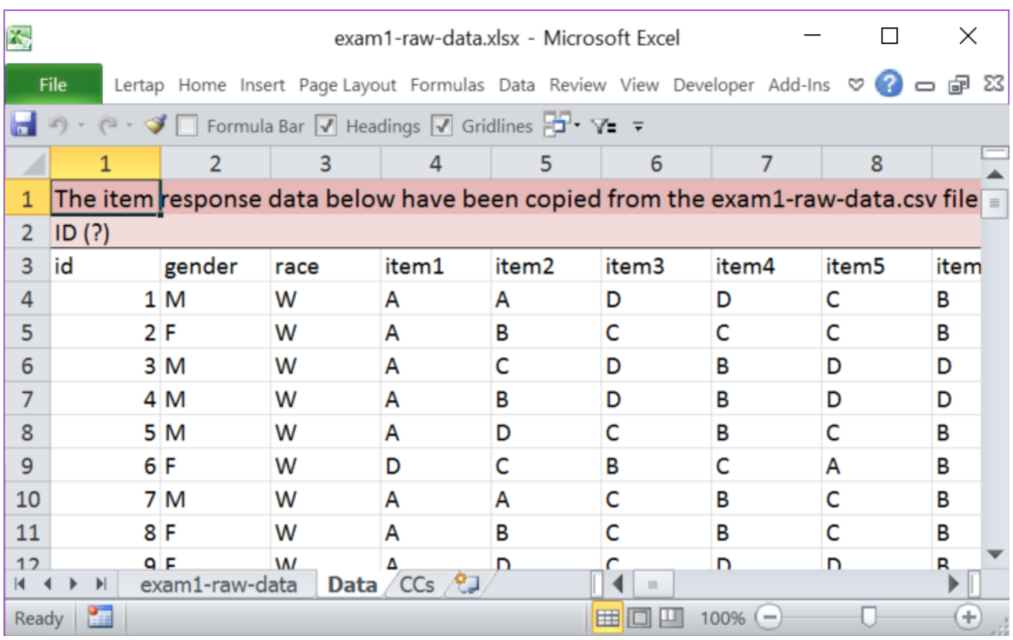


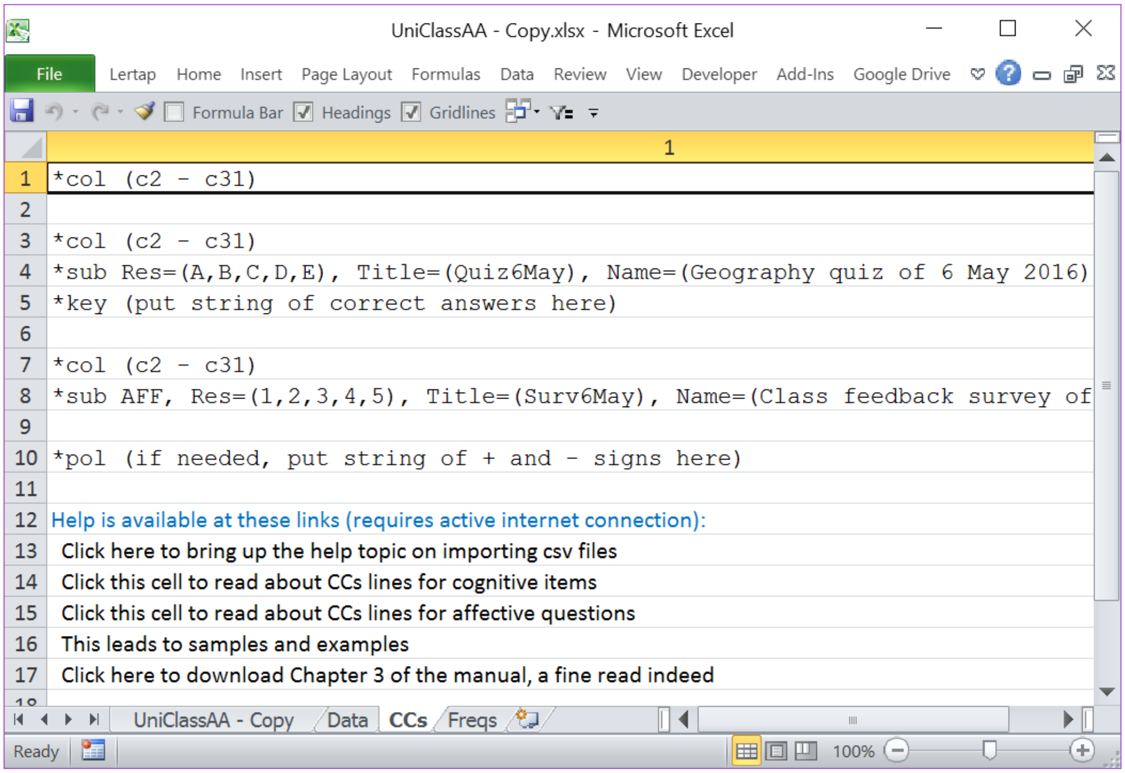
In this case, the *col line created by the macro was correct -- item responses started in column 2 and ended in column 31.

The *sub line was modified as shown above; the entries in the *key line were obtained from this webpage.

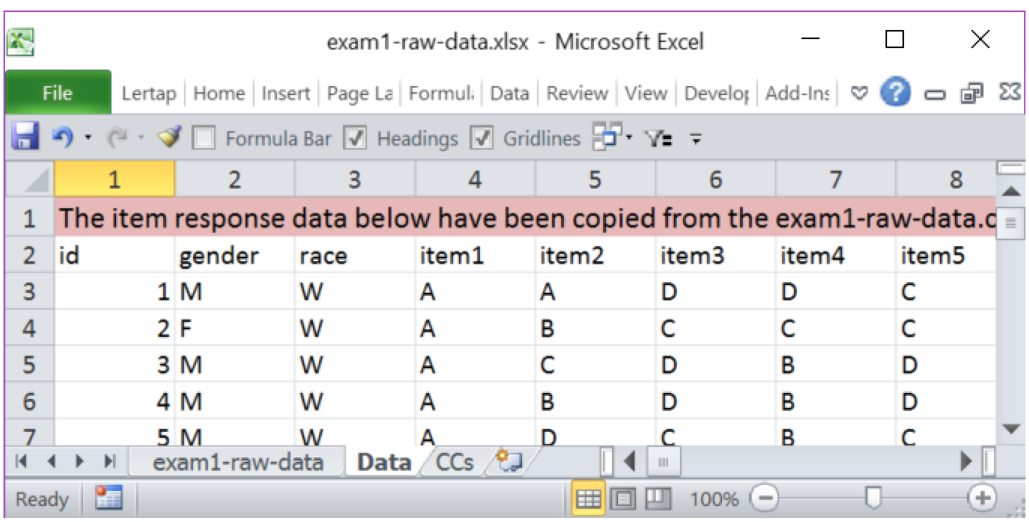
Using the ImportCSV macro with Example 2

The macro created these Data and CCs worksheets for Example 2:

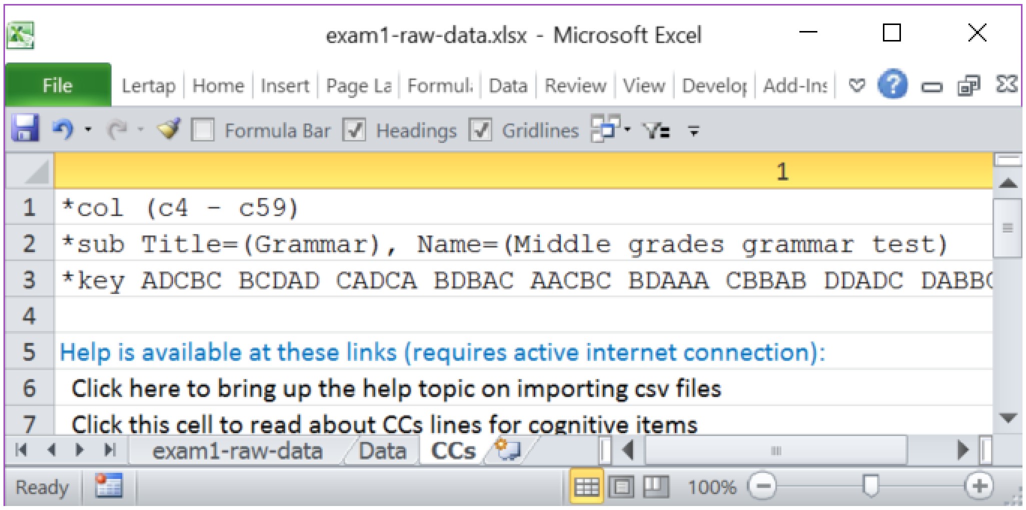




The Data sheet for this example required just one simple modification: the second row, with ID(?), was deleted.



In the CCs sheet, rows 1 and 2 were deleted. The item responses in this example begin in column 4 and end in column 59, so the *col line needed a change. Test items used {A,B,C,D} as response codes; these are the default response codes for cognitive tests, so the Res= statement is not really required. The final CCs sheet looked like this:



Related tidbits:

The *sub line in the second example could have included Res=(A,B,C,D); Lertap would have worked okay.

The entries in the *key line do not have to be grouped as seen in the second example. The string of correct answers does not have to have spaces after every five answers (note that the CCs sheet for Example 1 has a *key line with no spaces).

Often scanners will output a csv file with the correct answers to the items as the first line in the file. The "ScrunchBoss" macro may be used to read this line and create the string required by the *key line in the CCs worksheet. Once this is done, that extra line at the top of the csv file should be deleted so that Lertap 5 does not think the line corresponds to the first student's item responses.

The Help lines may be deleted. Since they have a blank line above them, Lertap never reads them.

After the macro runs, the original csv file will remain exactly as it was before.

The ImportCSV macro is "open source". It is found in the "ImportFiles" module in the Lertap5MacroSetA.xlam file. Excel's macro editor may be used to modify it.

More comments on importing data from files made by scanners are found in this topic.

4.7.4 ImportCSV2

There are times when students (or "candidates") will participate in a test which is administered over two sessions.

This is likely to result in two sets of item responses, one from the first session, and one from the second session.

The ImportCSV2 macro is of use when the responses for the first session are in one Excel row, followed by another row for the same student with her or his responses from the second session.

This macro copies all of the responses from the second session and appends them to the responses from the first session, resulting in a single Excel row with all responses (both sessions); after this the second row is deleted and the next student's responses are processed in the same way.

After doing this for all students, the macro passes control to a close cousin, the ImportCSV macro. The result will be a new Excel workbook with an xlsx extension; *the original csv file will be unchanged*.

Note:

See the "related tidbits" at the bottom of the previous page.

This macro makes a few assumptions -- they're mentioned when the macro starts up. Of these, a crucial one regards the location of the item responses on the first row: *there should be nothing to the right of the last item response in the first row.* All columns to the right of the last item response on the first row should be empty. This is because the macro will want to append the second row's item responses to the item responses in the first row.

4.7.5 IRTsas1

This macro is designed to set things up for an IRT analysis using SAS.

The macro begins by looking in the currently open Excel workbook for an IScores worksheet. If it can't find one, it then looks for an IStats worksheet and uses it to create an IScores sheet.

Once an IScores worksheet is on hand, the macro creates a new workbook and copies the IScores worksheet to it. The new workbook will be called SAS-IScores.xlsx.

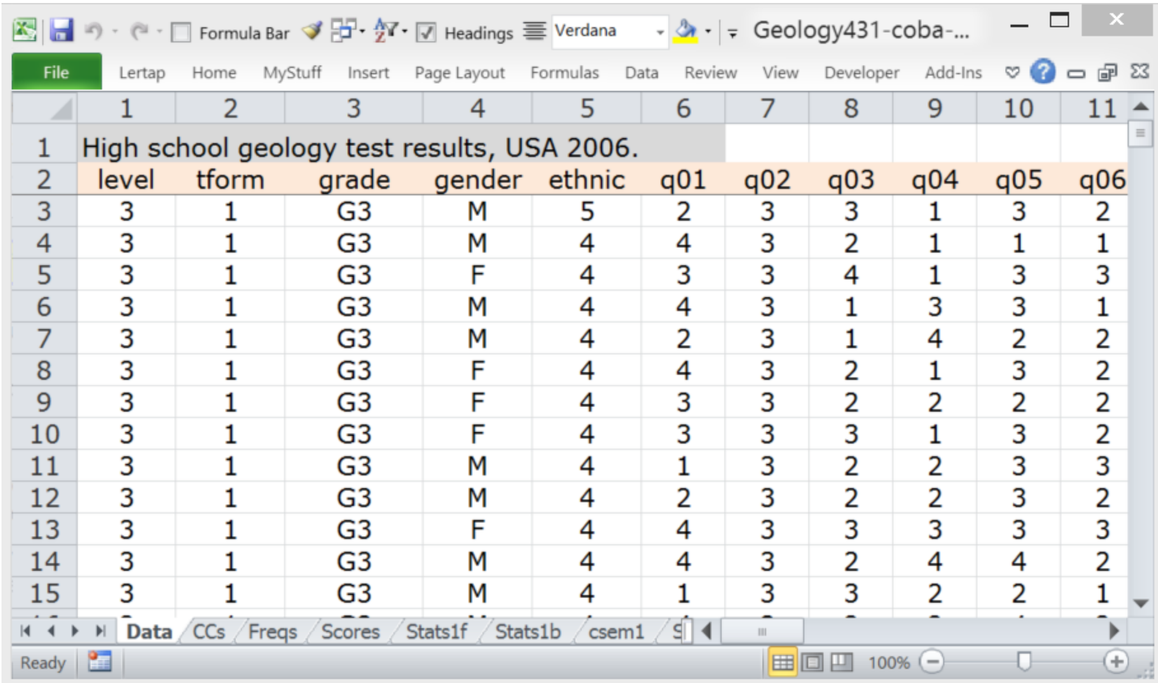
The macro will also create a bit of SAS code. This is placed in a file called SAS-IScoresProg.sas. An example of the SAS code created by the macro is shown here:

SAS-IScoresProg.sas - Notepad

```
* This SAS code was created by a Lertap 5 macro on 13/07/2015;

proc import datafile = '/folders/myfolders/SAS-IScores.xlsx'
  OUT = LertapIScores DBMS = XLSX Replace;
run;

proc print data = LertapIScores (obs=10);
 Title 'Item scores from Lertap 5 (first 10 records)';
run;

* IRT with SAS;

ods graphics on;
 Title 'SAS IRT 2PL model with Lertap IScores';

proc irt data=LertapIScores plots=(scree(unpack) icc) itemfit resfunc=twop;
  var q01-q49;
run;
```

The code seen in this little example assumes that the SAS user has set up to work with SAS Studio, as will be the case when working with the free University Edition of SAS released in the year 2014 -- users of this version may use the suggested default working folder, "myfolders", as seen above.

Note the "var q01-q49" code line above. It tells SAS that the variables to use in the IRT analysis start with the variable named "q01" and work over to the variable "q49". These names have been picked up (in this case) from the headers in a Lertap 5 Data worksheet:



The macro will change the variable names automatically, using whatever is found as headers in the Data worksheet.

Related tidbits:

This macro was inspired by two events: the release of the free SAS University Edition, and a 2014 book by Cody and Smith, "Test Scoring and Analysis Using SAS".

(Many of the item and test analysis capabilities in Lertap 5 have counterparts in SAS.)

To branch directly to the SAS IRT Proc description, click here.

4.7.6 IStatsPruner

This macro is made to take an IStats report apart.

IStats reports can be very long. There will be one row of item scores for each student, six rows of item summary data, one row of correlations for each item, from two to five rows of summary correlation data, and then one or two bands of data plotting levels of summary statistics. For a test of 60 items, taken by 400 students, the IStats report will have over 500 rows of information. There will be times when it's handier to separate the two IStats sections, that is, to have all the item scores in one worksheet, and all the correlation data in another. IStatsPruner provides the means to do this.

The macro starts up by making sure an IStats report has been created. (If you've forgotten how this is done, it's the "Item scores and correlations" option in the Run menu that creates an IStats report.)

Once it has confirmed the presence of an IStats report, the IStatsPruner macro takes IStats' item scores section and copies it to a new worksheet called IScores; it also copies IStats' item correlations section to another new worksheet called ICorrs. This makes it easier to work with these two distinct sections of information.

What happens if an IStats report doesn't exist? The usual: Excel and Lertap will let the air out of your bicycle tires, run down the battery in your mobile phone, and warm up all of the cold beers you've stashed away in what you thought was the icebox.

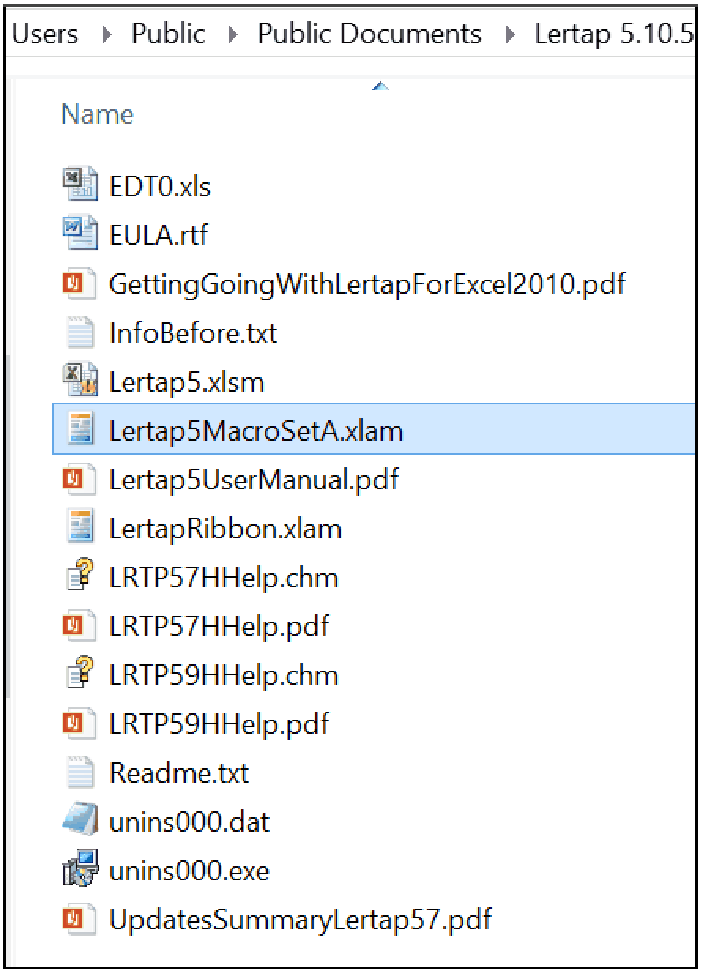
(If you don't believe this, start the macro and see what it says. There's no harm in starting it, even if you don't have an IStats report or an icebox full of beers.)

4.7.7 Lertap5MacroSetA

The Lertap5MacroSetA.xlam file is, as mentioned in the previous topic, a special Excel workbook and an integral component of the Lertap 5 system. It contains special-purpose small computer code modules called "macros".

When Lertap 5 is installed, this file of special macros, and a number of other files, are bundled together in the same folder. What folder is that? It's the one specified when Lertap was installed -- this is usually a folder nested with a user's Documents folder on a Windows computer, although in many cases it might instead be a folder nested with a Windows computer's Public Documents folder. (This topic discusses why an installer might elect to place Lertap 5 in the Public Documents folder.)

Here's a snapshot of the Lertap 5 folder contents from a computer with Lertap 5 installed under Public Documents:



The Lertap5MacroSetA.xlam file is highlighted above.

How to check the date on the Lertap5MacroSetA.xlam workbook presently installed? Use the "AboutLertapMacros" option as shown in the previous topic.

How to find out if Lertap5MacroSetA.xlam has been updated? A summary of system changes and updates, covering both Lertap 5 itself and the special macros workbook, may be seen with a click here.

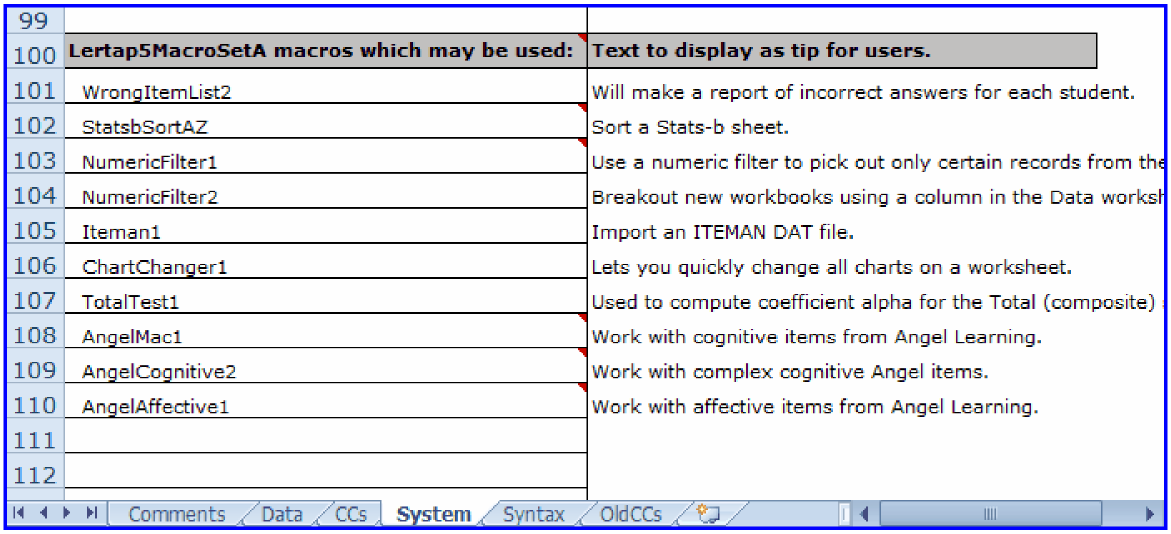
How to download Lertap5MacroSetA.xlam if you need to? Easy-peasy. Get it by mousing here. It's quite a small file and will download quickly.

How to install Lertap5MacroSetA.xlam? Just copy it into the Lertap 5 folder, letting it overwrite the former Lertap5MacroSetA.xlam file. NOTE: *it's best to close* Lertap 5 *before making any changes to the files in the Lertap 5 folder.*

4.7.8 Linking to macros

The previous topic mentioned that the macros activated from Lertap's Macs menu are stored in a file called Lertap5MacroSetA.xlam.

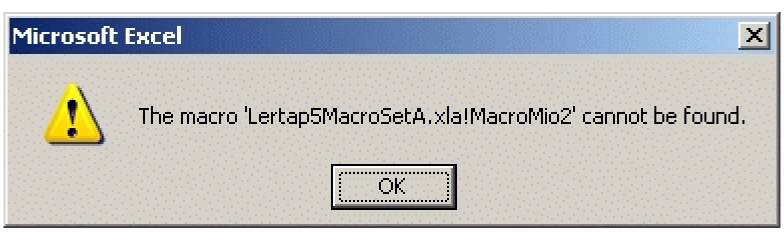
The links from Lertap to the macros in Lertap5MacroSetA.xlam are made by recording the names of the macros in Lertap's System worksheet. The following screen shot shows how part of the System worksheet looked in September, 2010:



99		
100	Lertap5MacroSetA macros which may be used:	Text to display as tip for users.
101	WrongItemList2	Will make a report of incorrect answers for each student.
102	StatsbSortAZ	Sort a Stats-b sheet.
103	NumericFilter1	Use a numeric filter to pick out only certain records from the
104	NumericFilter2	Breakout new workbooks using a column in the Data worksh
105	Iteman1	Import an ITEMAN DAT file.
106	ChartChanger1	Lets you quickly change all charts on a worksheet.
107	TotalTest1	Used to compute coefficient alpha for the Total (composite)
108	AngelMac1	Work with cognitive items from Angel Learning.
109	AngelCognitive2	Work with complex cognitive Angel items.
110	AngelAffective1	Work with affective items from Angel Learning.
111		
112		

There is a one-to-one relationship between the macro names seen in the System worksheet and the names seen in Lertap's Macs menu. What's seen in the Macs menu are the macro names found in the System worksheet.

In turn, these names correspond to the names of the macros found in the Lertap5MacroSetA.xlam file. So, if a user has placed "MacroMio6" in row 110 of the System worksheet, when the Macs menu is opened, MacroMio6 will be the tenth entry. When a user clicks on this entry, Lertap looks in the Lertap5MacroSetA.xlam file for a macro called MacroMio6. If this macro exists, it starts up. If it doesn't exist, an error message will appear, such as the one shown here:

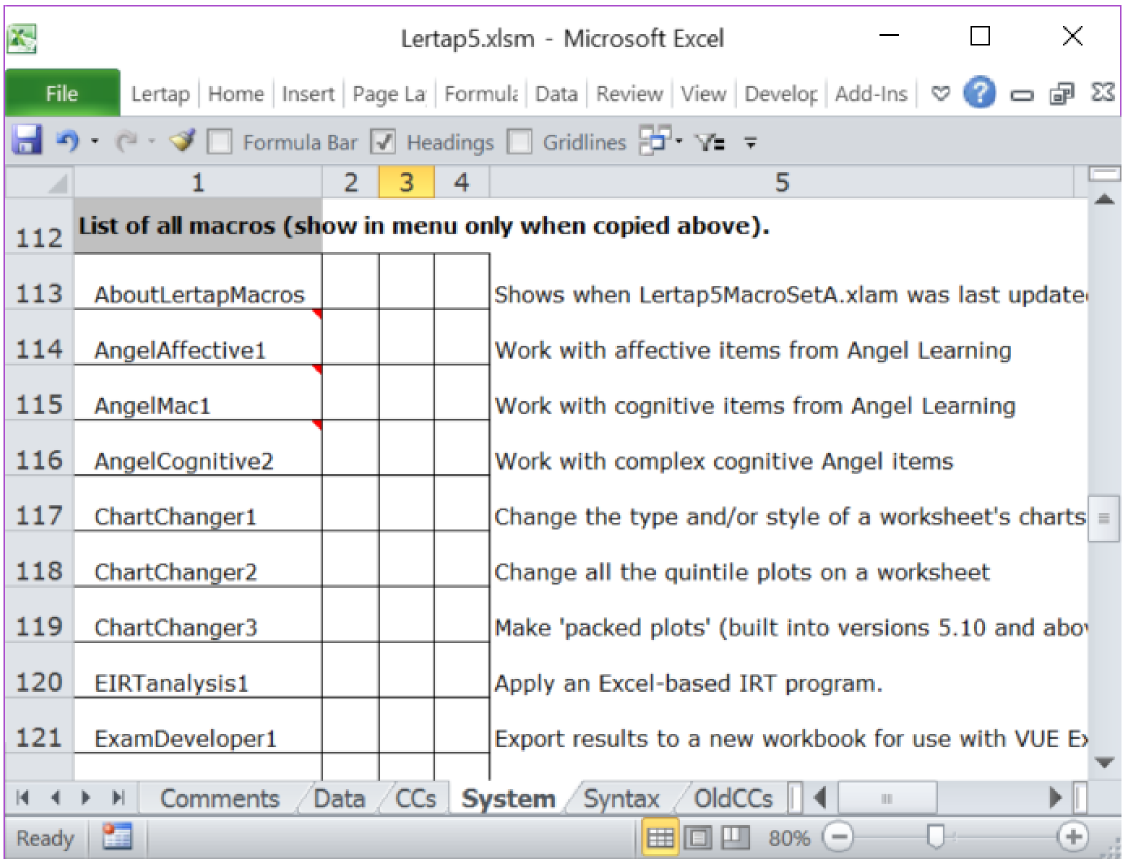


As mentioned, users may either modify the macros already found in Lertap5MacroSetA.xlam file, or create their own, saving them in the same file. You give each macro a unique name, and then place this name in one of the rows in Lertap's

System worksheet. The next time you start Lertap, the Macs menu will have been updated. Page forward for an example.

4.7.9 Active macros list

As of May 2016, the number of special macros in the Lertap5MacroSetA.xlam file had grown to be over 20. The complete list begins in row 113 of the System worksheet. A partial list is displayed below.

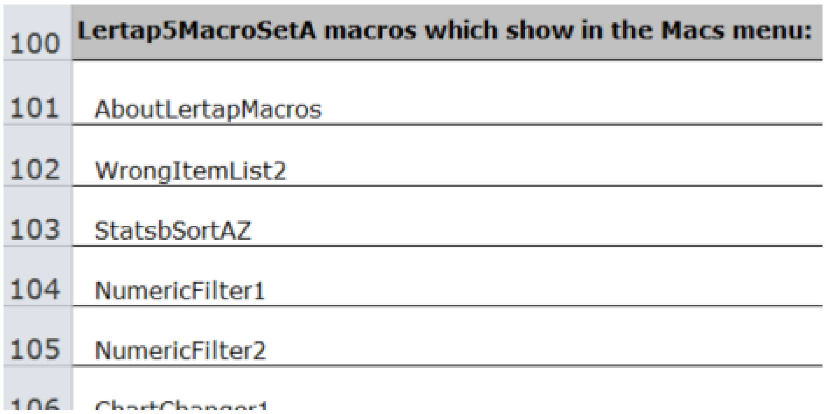


The names of the macros are found in the first column of each row. Columns 2, 3, and 4 are empty. The 5th column has a short description of what each macro does; this "description" will appear in the Macs Menu when the mouse cursor is left to hover over a macro name.

The entries in this list *should never be deleted*. The names of the macros *must not be changed*, not ever ever. The descriptions, however, may be changed to suit local needs.

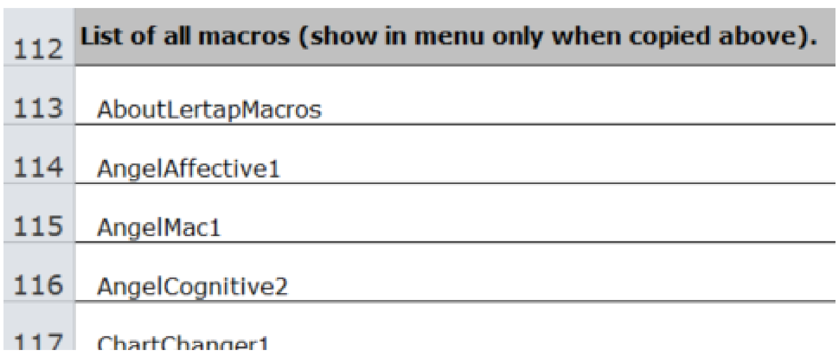
Rows 101 through 110 of the System worksheet, not shown above, contain the names of the "active" macros, the ones which display when the Macs Menu is showing. The entries in these rows have been copied from the "List of all macros" which starts in row 113.

The screen snapshot below is a sample of rows 100 to 105 of the System worksheet. These are active macros; their names will show in the Macs Menu (the names of all macros found in rows 101 through 110 show in the Macs Menu).



100	**Lertap5MacroSetA macros which show in the Macs menu:**
101	AboutLertapMacros
102	WrongItemList2
103	StatsbSortAZ
104	NumericFilter1
105	NumericFilter2

The snapshot below shows rows 112 to 116 (as of May, 2016).



112	**List of all macros (show in menu only when copied above).**
113	AboutLertapMacros
114	AngelAffective1
115	AngelMac1
116	AngelCognitive2

Let's suppose a user would like to make the AngelMac1 macro active. Let's also suppose that the user does not need to have the WrongItemList2 macro active; she would rather have AngelMac1 replace WrongItemList2 in the list of active macros.

What she would do is copy the first five columns in row 115 and paste the copy in row 102.

This does not delete the WrongItem2 macro as it will still be listed below row 112 (in May 2016 it was found in row 135).

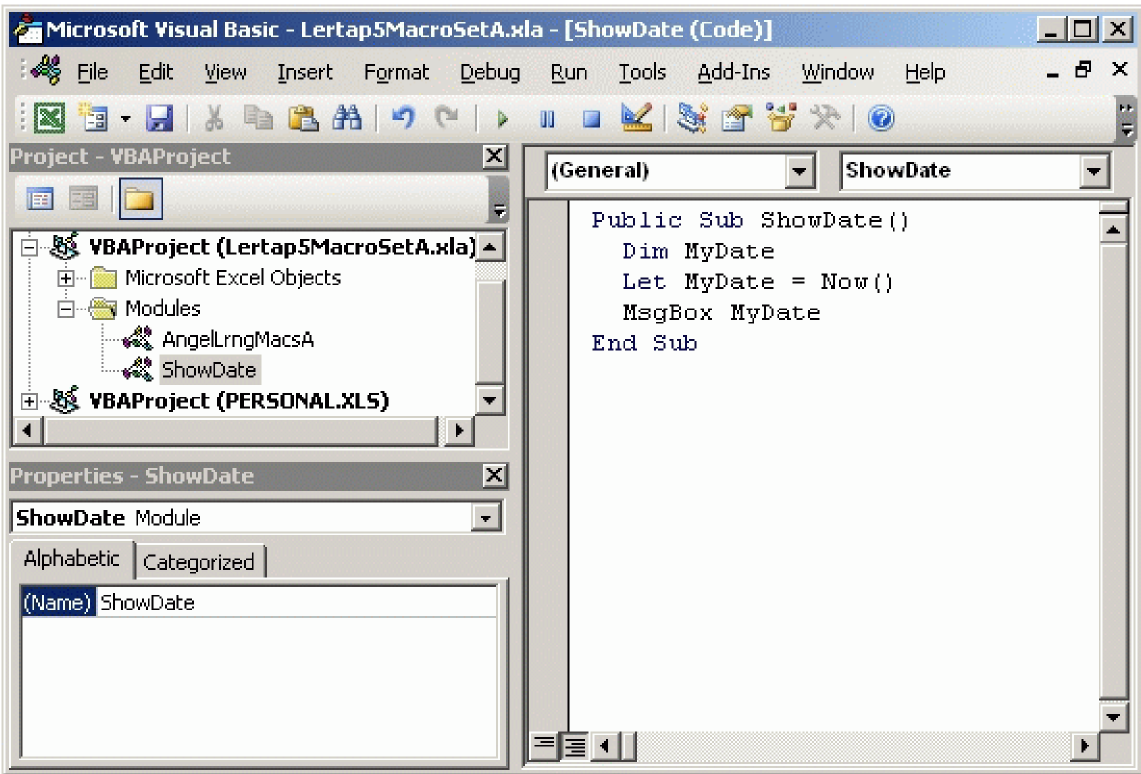
Please note that making changes in rows 101 to 110 of the System worksheet will not be effective until Lertap is saved, closed, and opened again. The Lertap5.xlsm workbook does not automatically save itself when the workbook is closed, or when Excel is closed.

4.7.10 Mac example

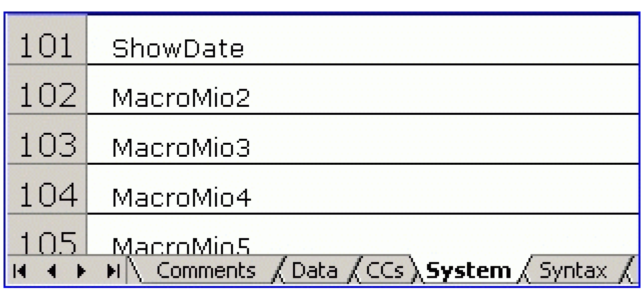
Please note: this example uses screen shots from Excel 2003.

Let's say that a user named Lee wants to create a macro which will display the date whenever it is activated.

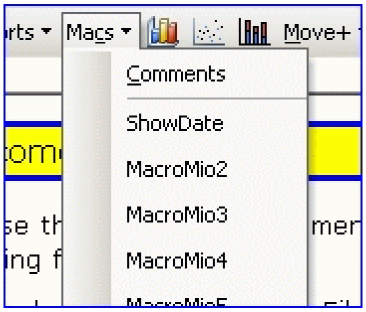
He opens Excel's Visual Basic Editor, inserts a new module in the Lertap5MacroSetA.xlam file, names the module ShowDate, and creates the following little subroutine. He saves the file.



Next, Lee puts the name of his new whiz-bang subroutine into Lertap's System worksheet:



He saves the Lertap5.xlsm workbook, closes it, and then re-opens it. When Lee taps on the Macs menu, here's what he sees:



With great anticipation, Lee clicks on the ShowDate option, and voila!, look:

So, there you go. With our friend Lee-ding the way, can you now add links to your own macros?

> *Postscript: your macros do not really have to be related to Lertap. If you keep the Lertap toolbar on screen, the Macs menu will always be available, even though you may not actually be Lertapping.*

4.7.11 Macro SAQs

SAQ1:

Q: *How many macro links can I have in the System worksheet?*
A: Ten (10).

SAQ2:

Q: *Do I have to have 10?*
A: No. When the Lertap5.xlsm workbook is opened, it populates the Macs menu by reading down the appropriate lines in the System worksheet, starting at row 101. Lertap stops reading these lines as soon as it encounters an empty cell in the first column, or when it has read information for ten macros. There may be more than 10 macros listed, but all those found after row 110 will not be used. (This correctly implies that the answer to SAQ1 is not exactly correct -- there may be more than 10 links, but only the first 10 are actually used.)

SAQ3:

Q: *I notice you've left two spaces at the start of each macro name in System worksheet. Why?*
A: Just because it makes the System worksheet a little neater. There's no real need for the two spaces.

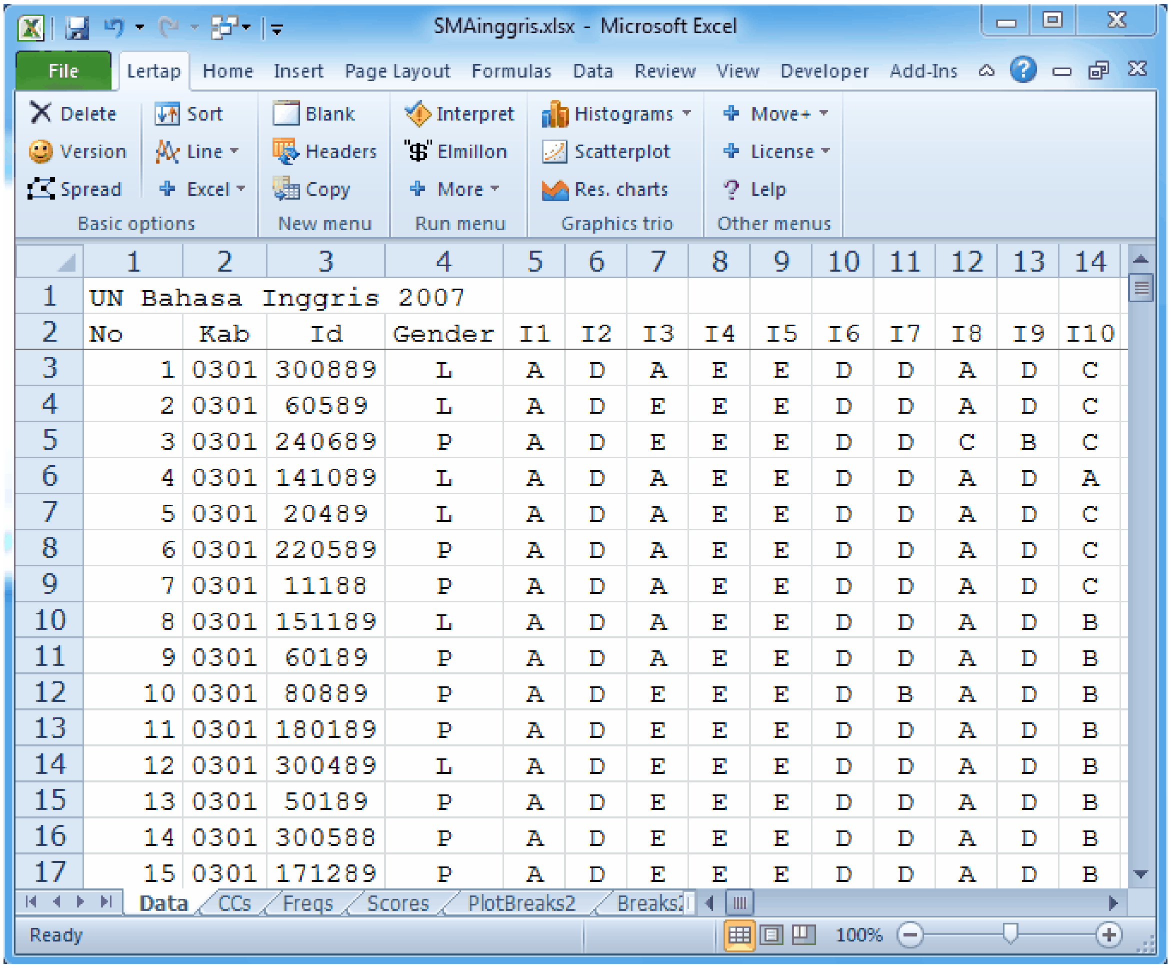
SAQ4:

Q: *How can I find out more about using the macros mentioned?*
A: Try them. They usually have a little explanation of what they do when they start up. If you have an interest in using Lertap with Angel Learning files, go back a few topics and click on the URL to the paper referenced way at the bottom of the topic, under 'Related titbits'.

(Thanks to phantom readers for sending in these Qs. Send yours, and if it's a good one we'll put it up here. Send them to lertap5@gmail.com)

4.7.12 NumericFilter2

Consider the following Data worksheet from an Indonesian data set:



The "Kab" field in the second column is a school district code; there were six districts, 0301, 0302, ..., 0306.

The *NumericFilter2* macro can be used to create a new workbook for each of the six districts, or for just some of them. It's a way of "breaking out" data records into new workbooks based on the codes found in any column of a Data worksheet.

In this example, the NumerFilter2 macro could also be used to breakout Gender data, creating two new workbooks, one with data only for boys, another with data only for girls. This would be done using the codes found in column 4 of this example, where L is the code for boys, and P the code for girls.

There are other ways to breakout, or select, only certain data records. See, for example, the Recode topic, and also the discussion of using a *tst "card" in the FilteringRecords topic.

And then, not to forget that there are ways to breakout results without Recodes, *tst, and NumericFilter2: there are options on the Run Menu to breakout test scores by groups, and also to breakout item responses by groups, where "groups" can be gender, school district, brand of beer preference, and so on. Interested? See the Breakouts and Ibreaks topics.

Note: using the NumericFilter2 macro is somewhat analogous to using "Select if" criteria in other software packages, such as SPSS. However, there is a difference: NumericFilter2, like a classic "select if" filter, will select only those records meeting the selection criterion, but it creates a physically distinct copy of those records in another workbook. (Classic select-if filters operate without creating a new subset of the data records.)

4.7.13 Omega1

This macro is designed to set things up for a 'McDonald' omega reliability analysis using the psych toolbox from CRAN, the repository of R packages. It creates a 'csv' data file suitable for use with applications such as JASP and SPSS, and small scripts for use with R and RStudio.

It also caters to users who have an interest in using the TAM package in R with an eye to IRT analyses (item response theory). Read more about this extended capability by jumping to this topic.

Note: it is also possible to get an estimate of omega by using the "Item scores and correlations" option within Lertap, an option that produces the "IStats" report. These reports will have omega output towards the end. Examples are found in this working paper.

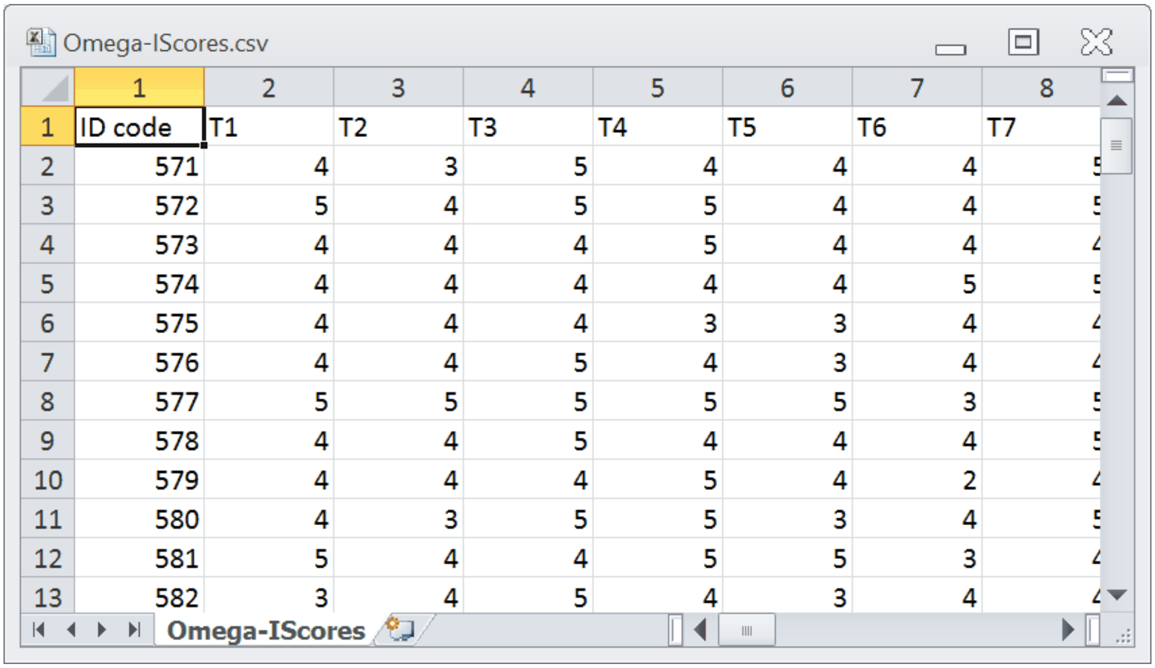
The procedure mentioned below was followed in the production of this little paper comparing coefficients alpha and omega.

The macro begins by looking in the currently open Excel workbook for an IScores worksheet. If it can't find one, it then looks for an IStats worksheet and uses it to create an IScores sheet. *Note that the workbook should have been saved before the macro is used.*

The macro then moves on; it'll create a special 'csv' data file based on the IScores worksheet, and a small text file with several lines of R code designed for use with the psych toolbox. Both of these files will be saved in the same folder the workbook itself is in. If an internet connection is active, the macro will also download an "Rmd" file to be used with RStudio (please refer to this paper).

The csv file will always be called Omega-IScores.csv. Excel will of course open this file, but it may also be opened using a text editor, such as the "Notepad" app found in Windows, or "TextEdit" on a Mac. JASP and SPSS will open it too.

An example of the csv file as displayed in Excel is shown below:



	1	2	3	4	5	6	7	8
1	ID code	T1	T2	T3	T4	T5	T6	T7
2	571	4	3	5	4	4	4	5
3	572	5	4	5	5	4	4	5
4	573	4	4	4	5	4	4	4
5	574	4	4	4	4	4	5	5
6	575	4	4	4	3	3	4	4
7	576	4	4	5	4	3	4	4
8	577	5	5	5	5	5	3	5
9	578	4	4	5	4	4	4	5
10	579	4	4	4	5	4	2	4
11	580	4	3	5	5	3	4	5
12	581	5	4	4	5	5	3	4
13	582	3	4	5	4	3	4	4

Note the columns above: "ID code", followed by columns with the item scores for each item -- in this case, the items were labeled "T1" through "T39". (The data come from "Mente", one of the Lertap sample datasets.)

The first column above is actually not wanted and *should be deleted* before continuing. If it's not deleted, "ID code" will be processed as an item scores field, and the output from the omega routine in the toolbox will be in error.

It is very common for the first column in Lertap data sheets to contain ID information for each record, so care must be taken -- *delete the ID column if it exists, then save the workbook again (sometimes it's necessary to ask Excel to save this csv workbook twice, depending on the version of Excel).*

As mentioned above, the macro will produce scripts for use with R and RStudio, saving them as small files in the folder where the workbook itself is located. Use of these scripts is exemplified in the "test drives" found towards the bottom of this document. They're quick and easy to use -- examples of the output produced by the scripts may be admired here (the link leads to a Word document that will download to your computer).

Related tidbits:

A detailed discussion / description of using R in conjunction with the Omega1 macro is available here.

Refer to this document for a discussion of issues that may arise when creating the Omega-IScores.csv file.

The Rmd file mentioned above is called "Omega-From-IScores.Rmd"; it may be downloaded from here.

The R file, "Omega-IScoresProg.R", is also available for downloading.

Another Rmd file, added in early 2021, makes it possible to get an estimate of McDonald's omega using a closed-form method, much like that used in the latest version of SPSS. Refer to this working paper. The file itself is "ClosedFormOmega1.Rmd", available as a download here. Note that the fact that SPSS uses this method does not imply that it's the preferred way to calculate omega (see the working paper).

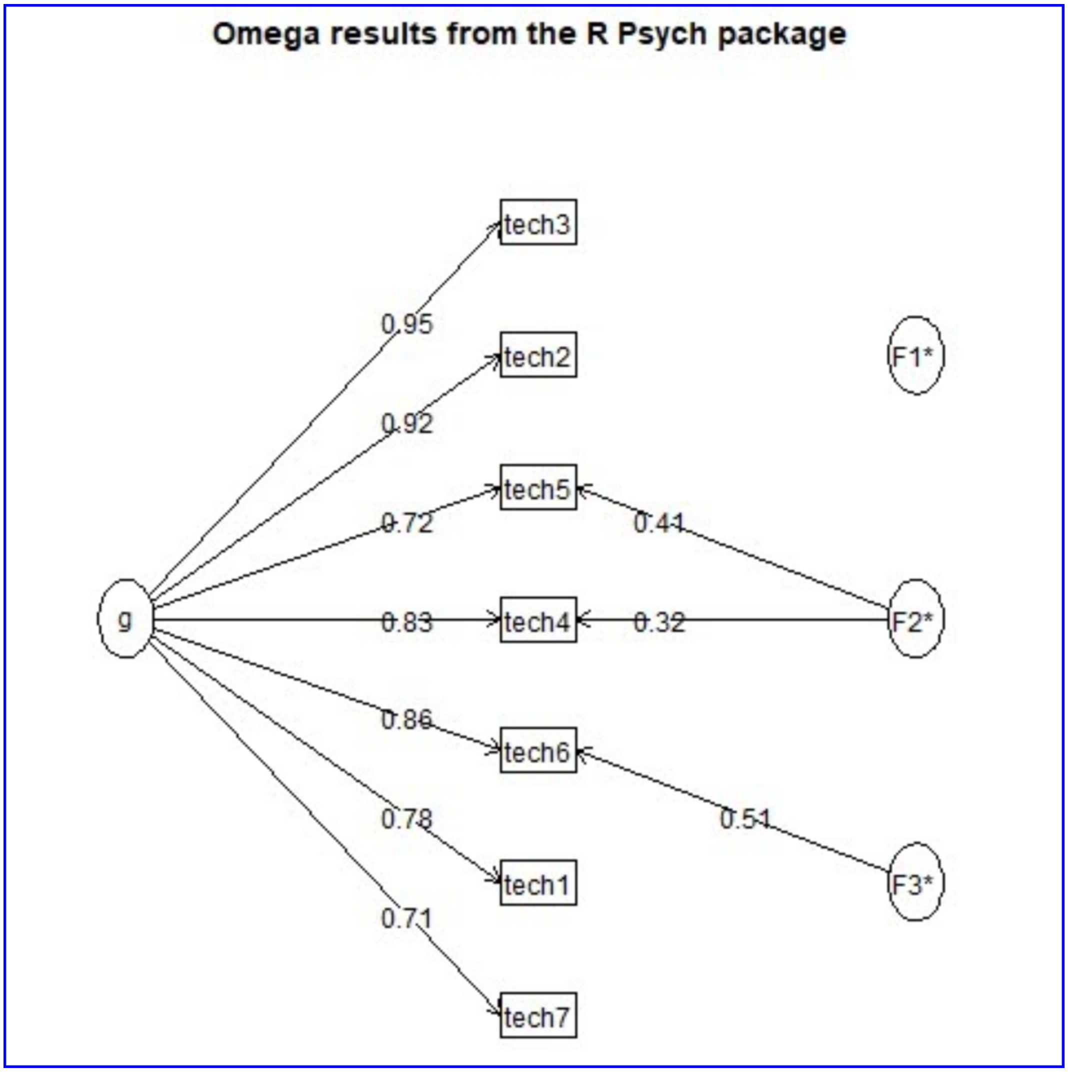
Be sure to have a look at this paper; it exemplifies the use of the omega routine in the Psych toolbox, and has links to sample output. (The next page also displays sample output.)

A general discussion of the use of Lertap's special macros is in this topic.

Help? Write to lertap5@gmail.com

4.7.13.1 Omega Factors Example

A sample of the factor structure graphs produced by the omega function in the Psych package is shown below for the SES dataset:



The omega routine's text output corresponding to the graph above is shown here:

```
[1] "Wed Nov 29 13:37:11 2017"
Omega results from the R Psych package
Call: omega(m = DunnItems, digits = 3, title = "Omega results
from the R Psych package")
Alpha:                 0.94
G.6:                   0.94
Omega Hierarchical:    0.93
```

```
Omega H asymptotic:    0.97
Omega Total            0.95

Schmid Leiman Factor loadings greater than  0.2
         g   F1*   F2*   F3*   h2   u2   p2
tech1 0.78                   0.62 0.38 0.99
tech2 0.92                   0.85 0.15 1.00
tech3 0.95                   0.91 0.09 1.00
tech4 0.83        0.32       0.79 0.21 0.86
tech5 0.72        0.41       0.69 0.31 0.75
tech6 0.86              0.51 1.00 0.00 0.74
tech7 0.71                   0.51 0.49 0.99

With eigenvalues of:
   g  F1*  F2*  F3*
4.80 0.01 0.27 0.28

general/max  17.42   max/min =   47.89
mean percent general =  0.9    with sd =  0.12 and cv of  0.13
Explained Common Variance of the general factor =  0.9

The degrees of freedom are 3  and the fit is  0
The number of observations was  193  with Chi Square =  0.29
with prob <  0.96
The root mean square of the residuals is  0
The df corrected root mean square of the residuals is  0.01
RMSEA index =  0  and the 10 % confidence intervals are  0 0
BIC =  -15.49

Compare this with the adequacy of just a general factor and no
group factors
The degrees of freedom for just the general factor are 14  and
the fit is  0.15
The number of observations was  193  with Chi Square =  27.61
with prob <  0.016
The root mean square of the residuals is  0.03
The df corrected root mean square of the residuals is  0.04

RMSEA index =  0.073  and the 10 % confidence intervals are
0.03 0.11
BIC =  -46.06

Measures of factor score adequacy
                                                 g   F1*   F2*
 F3*
```

```
Correlation of scores with factors            0.98  0.09  0.67
0.93
Multiple R square of scores with factors      0.96  0.01  0.46
0.87
Minimum correlation of factor score estimates 0.91 -0.98 -0.09
0.74

 Total, General and Subset omega for each subset
                                                 g   F1*  F2*
F3*
Omega total for total scores and subscales    0.95 0.94 0.85
0.87
Omega general for total scores and subscales  0.93 0.93 0.69
0.82
Omega group for total scores and subscales    0.02 0.00 0.15
0.05
```

4.7.13.2 IRT with TAM

When the Omega1 macro is used, it creates the two files mentioned earlier, Omega-IScores.csv and Omega-IScoresProg.R

It will always create those two files. If there is an active internet connection, it may download some additional files.

As of December 2021, files that might be downloaded included these possibles: Omega-From-IScores.Rmd, Rasch-Analysis-TAM.Rmd, and IRTmoduleUWO-1.Rmd

The first of these files is mentioned towards the end of this paper, a paper which should be a general read even for those users whose interests regard the application of the TAM package.

Rasch-Analysis-TAM.Rmd computes Rasch statistics and creates ICCs, item characteristic curves having an empirical overlay, making it possible to visually judge goodness of fit. It may be used by following the steps mentioned in the paper referenced in the paragraph immediately above -- where that paper references Omega-From-IScores.Rmd (page 11), substitute Rasch-Analysis-TAM.Rmd.

An example of the report produced by the Rasch-Analysis-TAM.Rmd script is shown here.

Other experimental work with related R packages is on-going; this includes, for example, a fairly comprehensive Rmd script for dichotomous IRT models based on work from the University of Western Ontario called "IRTModuleUWO-1.Rmd". (This UWO paper is a good reference.)

Related tidbits:

Rmd files contain "R Markdown" code. (Refer to the paper mentioned above.)

The "Omega-From-IScores.Rmd" file may be downloaded from here.

Rasch-Analysis-TAM.Rmd should be available here.

The IRTmodule "UWO-1.Rmd" file may be downloaded from here.

4.7.14 difR1

This macro is designed to set things up so that the 'difR' package may be used to get estimates of possible differential item functioning for cognitive test items (see difR at CRAN, the repository for R projects).

It creates a 'csv' data file, "difR-IScores.csv".

Read all about it here.

4.7.15 ExamDeveloper

The installer package for the Excel 2010/2013/2016 versions of Lertap 5 includes a special workbook template called "EDT0.xls". You will be able to find it in the folder where the main Lertap 5 files are (Lertap5.xlsm, Lertap5MacroSetA.xlam, and LertapRibbon.xlam). This folder is usually located under a user's standard "Documents" folder and will often have a name *similar* to "Lertap 5.10.7.4". This topic has a snapshot of the partial contents of such a folder.

The EDT0.xls template makes an attempt to be self-documenting. Open it and have a look. To understand it more completely, a visit to this webpage might be helpful. EDT stands for "Exam Developer Template".

The purpose of the ExamDeveloper1 macro is to copy the many statistics found in Lertap 5's "reports" over to a new workbook based on the EDT0.xls template. The macro assumes that you're looking at a Stats_f report for a cognitive subtest before the macro is opened via the Macs Menu.

Suppose, for example, that the M.Nursing.xlsx workbook has been opened, and that the Interpret and Elmillon options have been run, resulting in new worksheets called "Stats1f", "Stats1b", and "Stats1ul".

Suppose that we're looking at the Stats1f sheet, and that we then open the Macs Menu and select "ExamDeveloper1".

The macro will make a new workbook called ED_M.Nursing.xlsx, a workbook based entirely on the EDT0.xls template.

Try it and see for yourself. Before doing so, make sure that ExamDeveloper1 is showing as one of the "top ten" macros by following these instructions. Assuming that you have indeed downloaded a copy of the M.Nursing dataset, and have used Interpret and Elmillon so that a Stats1f worksheet is available, you may expect to have the ExamDeveloper1 macro make the just-mentioned ED_M.Nursing.xlsx workbook. Open it and have a look -- it's ready for use with Pearson's Exam Developer system.

The EDT0.xls template is designed for use with cognitive test items having no more than six options {A B C D E F}. To use it with items using other options, such as {1 2 3 4 5}, write to lertap5@gmail.com and ask for assistance.

Also see: the PearsonVUE topic.

4.7.16 PearsonVUE

"PVueExamSeries1" is a macro created for users who obtain item response data from Pearson VUE in spreadsheet format, for example, as an Excel workbook.

A not-too-technical document which describes the macro is available as a PDF document here.

We have also put together a much more extensive stand-alone suite of macros which accept and process data collected using Pearson VUE's Client Data Interchange model and services. In the main, these services are designed for large scale testing programs involving the administration of many tests and test forms. The files which commonly accompany the CDI-based testing model often feed into a client relational database of some sort.

See also the Exam Developer topic.

Need more information? Drop us a note: lertap5@gmail.com.

4.7.17 Quantile Shader 1

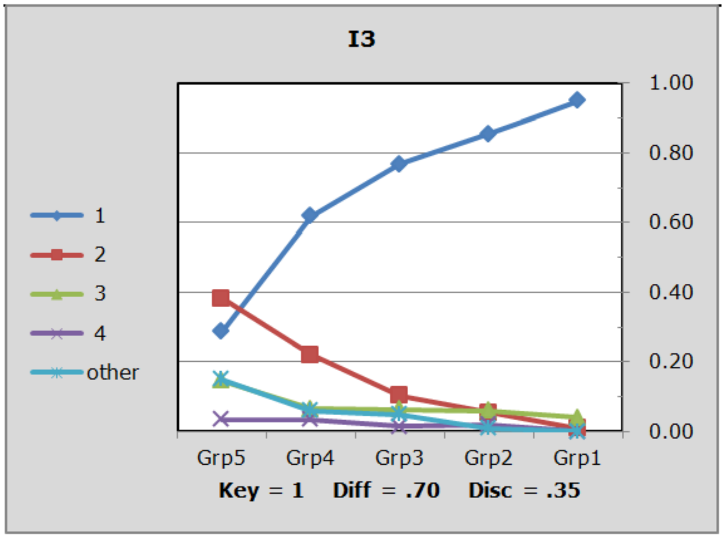
This simple little macro was added to the Lertap5MacrosSetA.xlam workbook early in 2016.

It has some relationship to a setting in row 98 of the System worksheet:

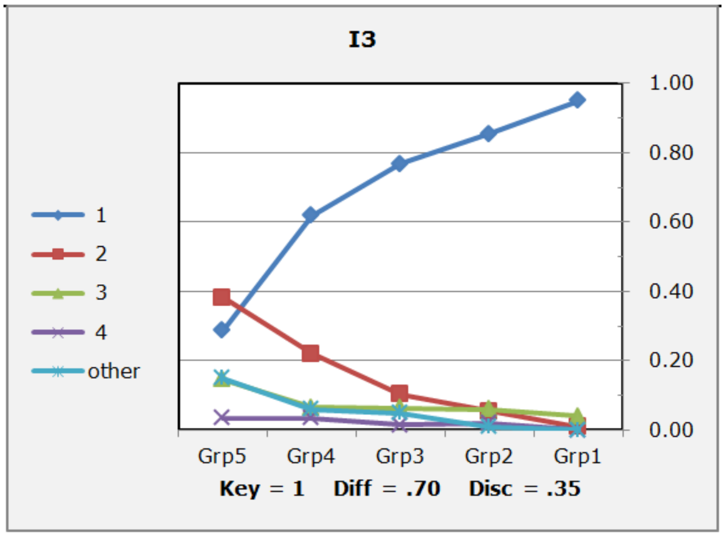
95	Use page margin settings in rows 64 - 69 above?	yes	yes/no	yes
96	Remove the series markers from the plot lines?	no	yes/no	no
97	Present black & white option for the plots?	yes	yes/no	no
98	Use very light backgrounds in the chart frames?	no	yes/no	no

If row 98's setting is "no", the QuantileShader1 macro may be used to lighten up the frames found in quantile plots -- it's an alternative to changing row 98's setting to "yes".

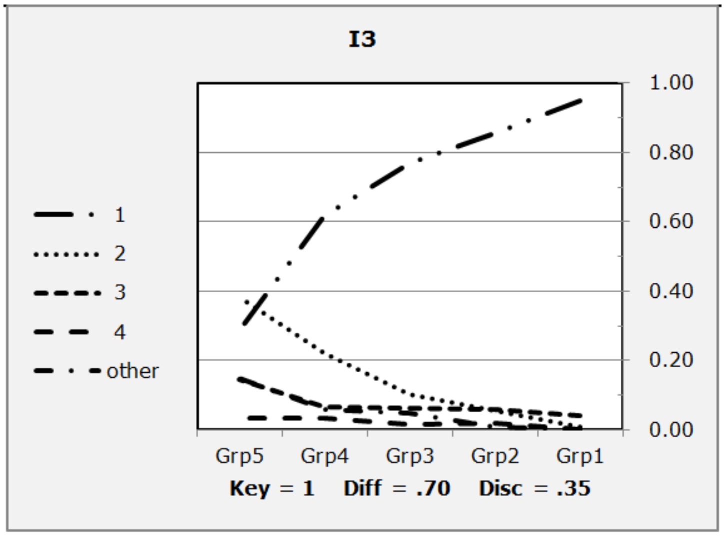
If for example a quantile plot looks like this:



It will look like this after using the macro:



If row 97 is set to "yes", the macro will offer to make black and white copies of the plots:



Note: row 96's setting is ignored in the present version of Lertap 5. "Markers" are the little symbols used in plot lines to help distinguish one line from another. Of the three sample plots shown above, the first two use markers, five per line. The third plot, the black and white one, does not; the black and white plot relies on unique line styles to enable viewers to distinguish one line from another. (These black and white plots are easier to interpret when the chart is expanded by selecting it and then tugging at its edges.)

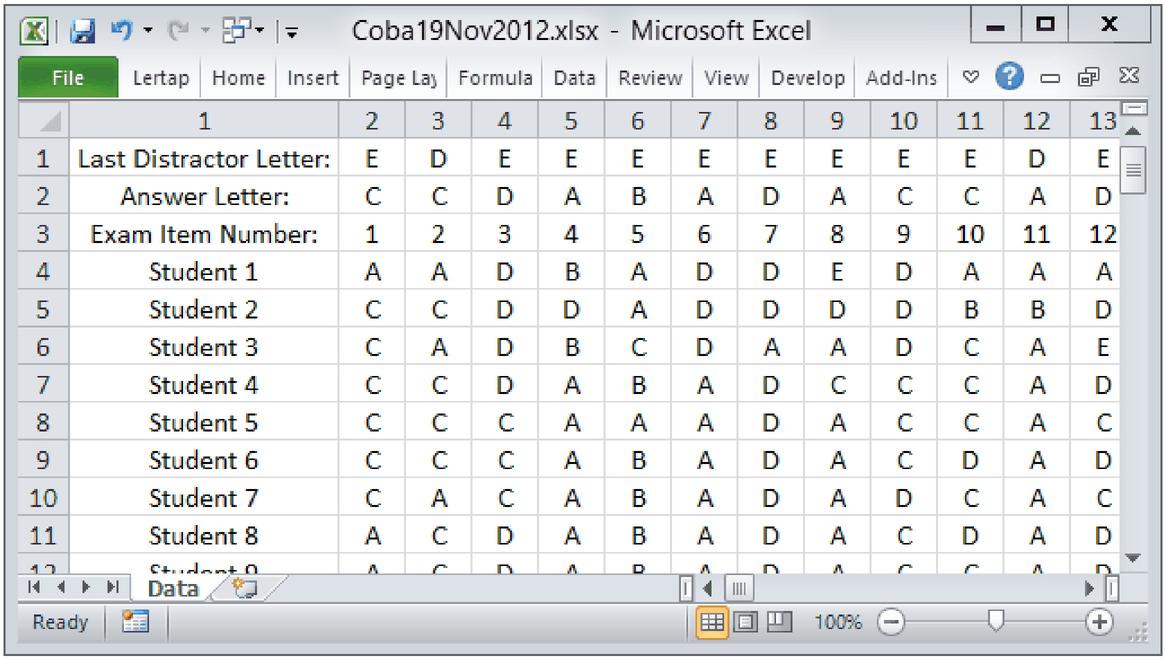
Another way to remove colors from quantile plots is via the use of a standard Excel color scheme called "Grayscale". Branch to this topic to read about it.

4.7.18 StatsbSortAZ

This handy, simple little macro will make a copy of a Stats_b report, such as Stats1b, and sort it on the basis of either item difficulty or item discrimination. It's one of those little macros which has no frills and works rapidly. To find out how to use it, just start it up by selecting it from the Macs menu. It'll give instructions. Don't be afraid to simply try it; it doesn't alter anything you may have already done.

4.7.19 ScrunchBoss

Look at this workbook sent in by a user:



	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Last Distractor Letter:	E	D	E	E	E	E	E	E	E	E	D	E
2	Answer Letter:	C	C	D	A	B	A	D	A	C	C	A	D
3	Exam Item Number:	1	2	3	4	5	6	7	8	9	10	11	12
4	Student 1	A	A	D	B	A	D	D	E	D	A	A	A
5	Student 2	C	C	D	D	A	D	D	D	D	B	B	D
6	Student 3	C	A	D	B	C	D	A	A	D	C	A	E
7	Student 4	C	C	D	A	B	A	D	C	C	C	A	D
8	Student 5	C	C	C	A	A	A	D	A	C	C	A	C
9	Student 6	C	C	C	A	B	A	D	A	C	D	A	D
10	Student 7	C	A	C	A	B	A	D	A	D	C	A	C
11	Student 8	A	C	D	A	B	A	D	A	C	D	A	D

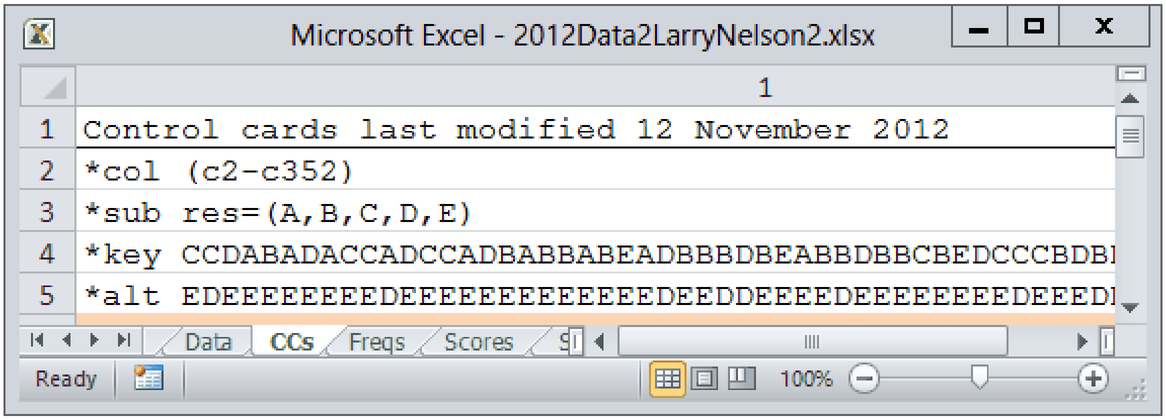
There are several things that need to be done before this Data worksheet is ready for Lertap processing. (For a reminder of how a Data sheet should be structured, please have a gander at this topic.)

As far as this topic goes, let us draw your kind attention to the first two rows, "Last Distractor Letter", and "Answer Letter".

Consider the "Answer Letter" row. It contains the correct answer for each item, that is, it provides us with the item "keys", the correct answers. There is one key per item (of course) and, above, each key has been recorded in a single column.

These keys will go into the *key line used in the CCs worksheet. But does a *key line have the correct answers spread over columns, as they are above? No. No indeed.

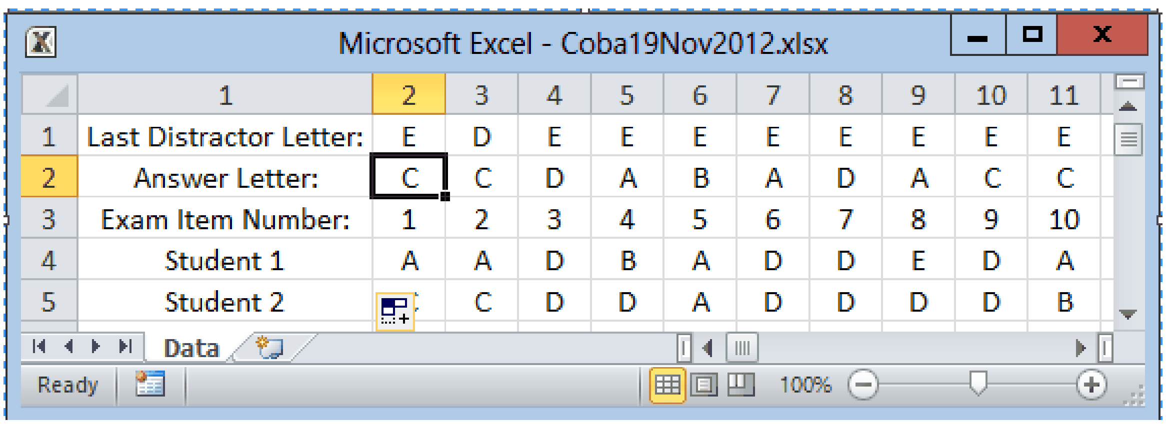
We want to see something like the *key line seen here:



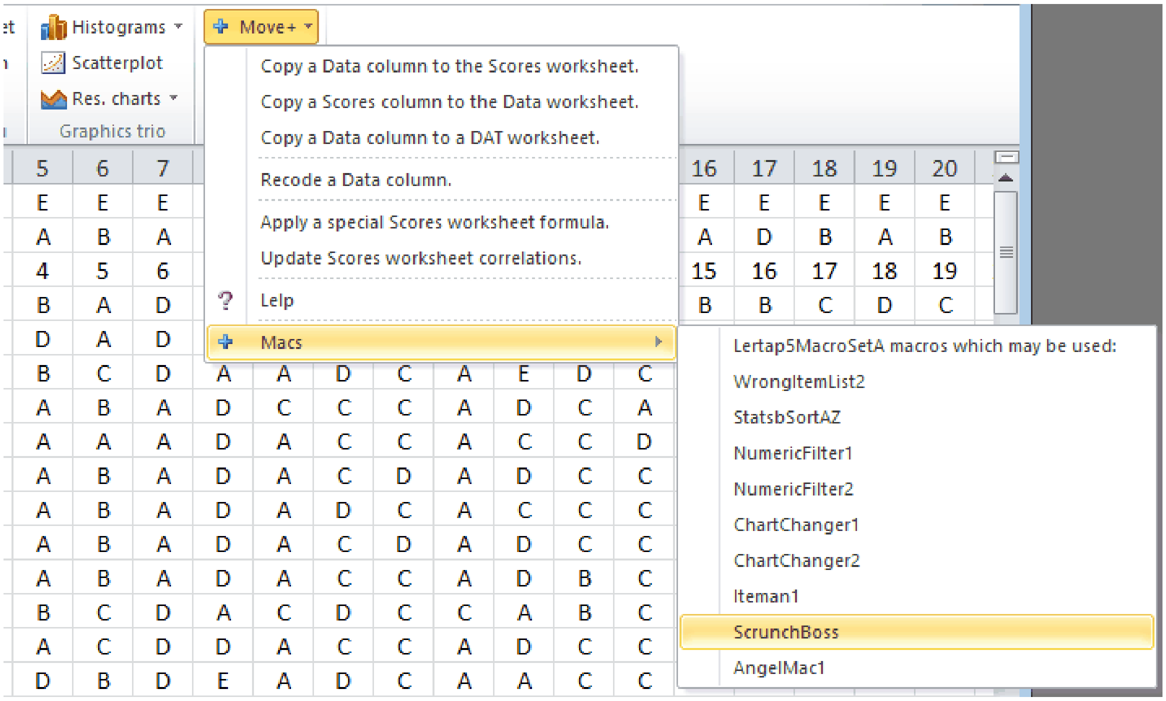
What do we need to do? Concatenate the columns with the keys. Naturally! Concatenate means to join. In the very technical terms sometimes found in Lertap documentation, we could say that we need to "scrunch" the columns with the keys.

Enter the "ScrunchBoss" macro.

Here's how to use it.



Start by selecting the first key. In this case it's found in row 2, column 2, as indicated above.



Next, go to the Move+ menu, click on Macs, then click on ScrunchBoss.

That should do it! The ScrunchBoss will join all the correct answers together and make what's called a "string". This string will be found in the column immediately to the right of the last correct answer.

Once the string has been made it may be copied and pasted into the appropriate row in the CCs worksheet.

What about that row seen above called "Last Distractor Letter"; it later appears, in scrunched form, in the CCs sheet seen above, as the *alt line.

The *alt line indicates the last option used by an item. In this case, the res=(A,B,C,D,E) part of the *sub line says that at least one item makes use of five possible options, from A to E. But here, in this example, not all items use all of these options. The second item's last option is D; the 11th item's last option is also D.

To read more about *alt, why not see if your mouse (or finger?) will click here?

Some readers might be wondering what happens when an item has more than one correct answer? In Lertap, such items are said to be "multiply-keyed", or (often) "double-keyed", and *mws lines are used in the CCs sheet to define them. Please refer to Example C12 in this topic.

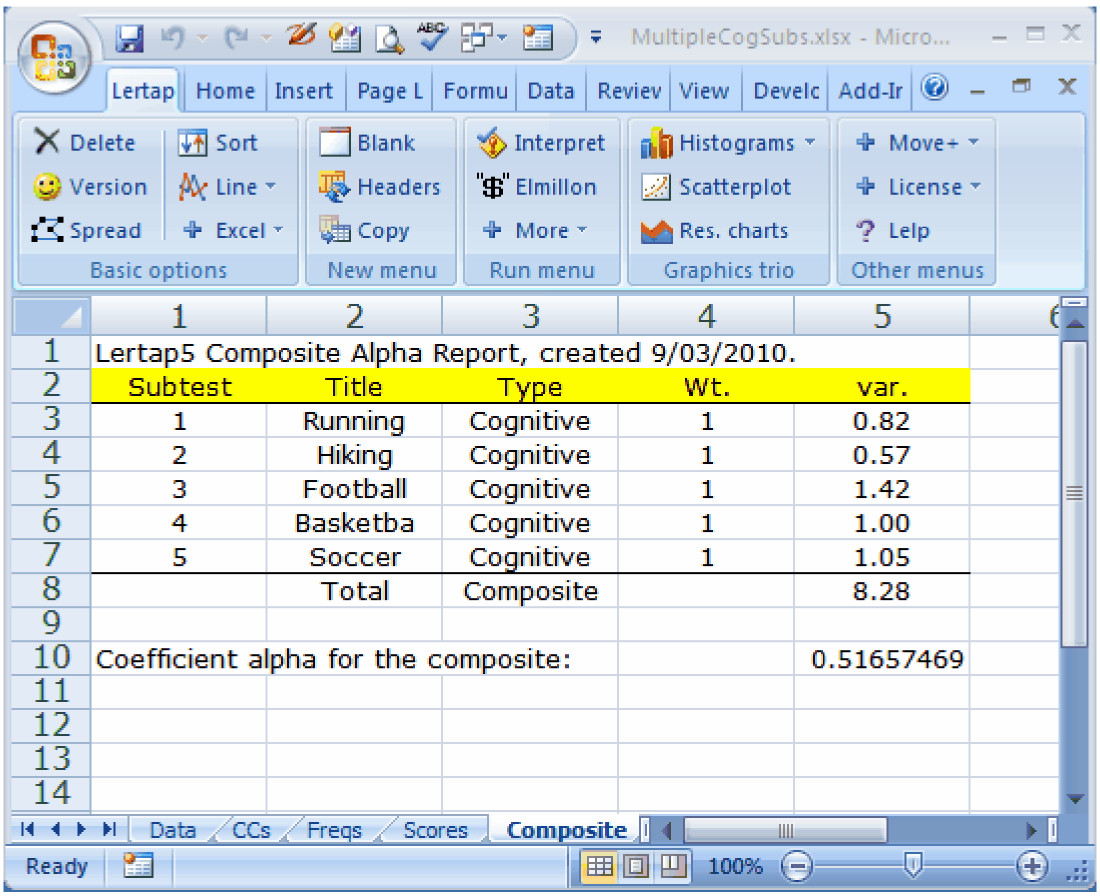
Finally, other users might ask about the practical effects of using the *alt line. It's a question which is briefly discussed in the manual, but we might do well to answer it again.

If Lertap knows that an item does not make use of all of the options used by other items, it will make some adjustments to the Stats1b and Stats1f reports. For example, the Stats1b report has a ? column. One of the things this column does is "flag" options which are not taken by anyone. If an option is in fact not used by the item, the *alt line will effectively inform the Stats1b report that it need not (should not!) flag the option.

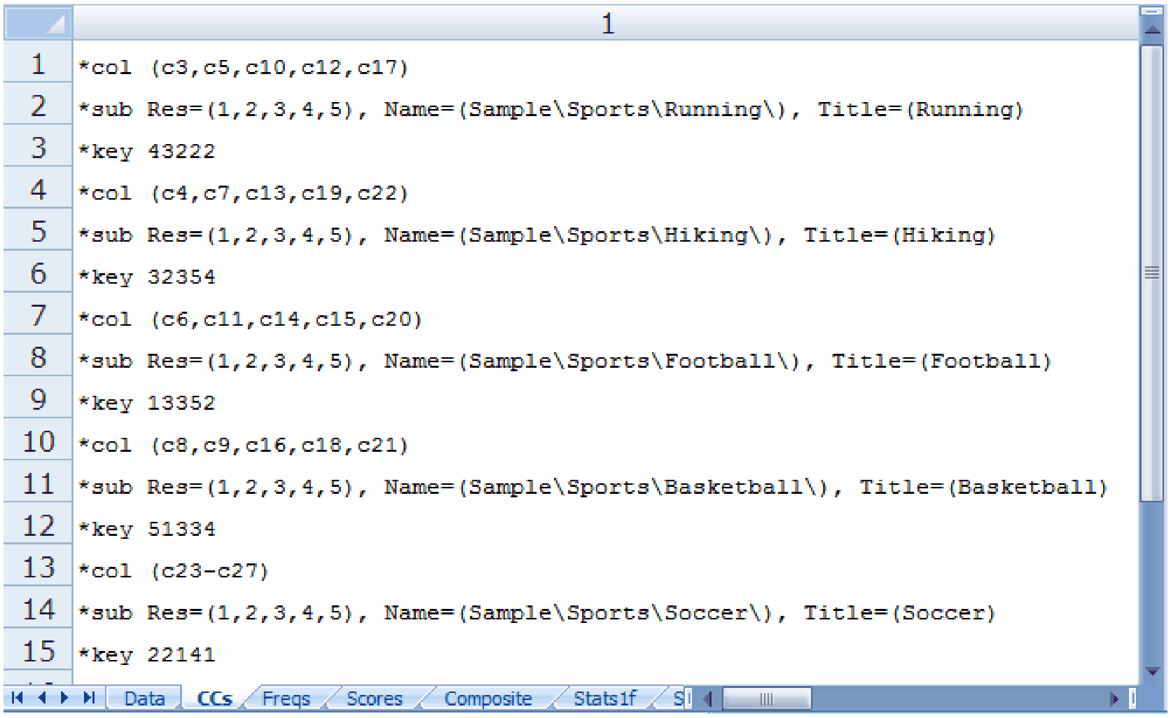
In the Stats1f report, options which are not used by an item will not be included as part of the report.

4.7.20 TotalTest1

The *TotalTest1* macro produces output exemplified here when running Lertap 5 with Excel 2007:



The CCs lines corresponding to this output are shown here:



	1
1	*col (c3,c5,c10,c12,c17)
2	*sub Res=(1,2,3,4,5), Name=(Sample\Sports\Running\), Title=(Running)
3	*key 43222
4	*col (c4,c7,c13,c19,c22)
5	*sub Res=(1,2,3,4,5), Name=(Sample\Sports\Hiking\), Title=(Hiking)
6	*key 32354
7	*col (c6,c11,c14,c15,c20)
8	*sub Res=(1,2,3,4,5), Name=(Sample\Sports\Football\), Title=(Football)
9	*key 13352
10	*col (c8,c9,c16,c18,c21)
11	*sub Res=(1,2,3,4,5), Name=(Sample\Sports\Basketball\), Title=(Basketball)
12	*key 51334
13	*col (c23-c27)
14	*sub Res=(1,2,3,4,5), Name=(Sample\Sports\Soccer\), Title=(Soccer)
15	*key 22141

This example involves five cognitive subtests.

The information displayed in the Composite Alpha Report is derived from various sources. The Titles are taken from the CCs lines, as are the values seen in the Wt. column (a subtest's Wt. value will equal 1 (one) unless Wt= assignments are made on the *sub line). "var." values are taken from the Scores worksheet.

The Total score is a composite:

Total = Sum (Wt(j) * X(j))

where Wt(j) is the 'weight' assigned to subtest 'j', and X(j) is a student's score on subtest 'j'. The Sum is taken over all J subtests whose Wt. is not equal to 0 (zero).

Let V(j) be the variance (var.) of subtest j, and V(T) the variance of the Total score.

Then

Let P1 = Sum ($Wt(j)^2$ * V(j))

Let P2 = J / (J-1)

Let P3 = P1 / V(T)

then

alpha = P2 * (1 - P3)

In this example, alpha = 0.51657469

To keep a subtest from entering the Total score, set Wt=0 on the subtest's *sub line.

```
*col (c3-c27)
*sub Res=(A,B,C,D,E,F), Name=(Knowledge of LERTAP2), Title=(Knwldge), Wt=0
*key AECAB BEBBD ADBAB BCCCB BABDC
*alt CEDBC CEDFD EDCBD DCCDD DEEDF
*col (c28-c37)
*sub Aff, Name=(Comfort with using LERTAP2), Title=(Comfort), Wt=0
*pol +---- ++--+
```

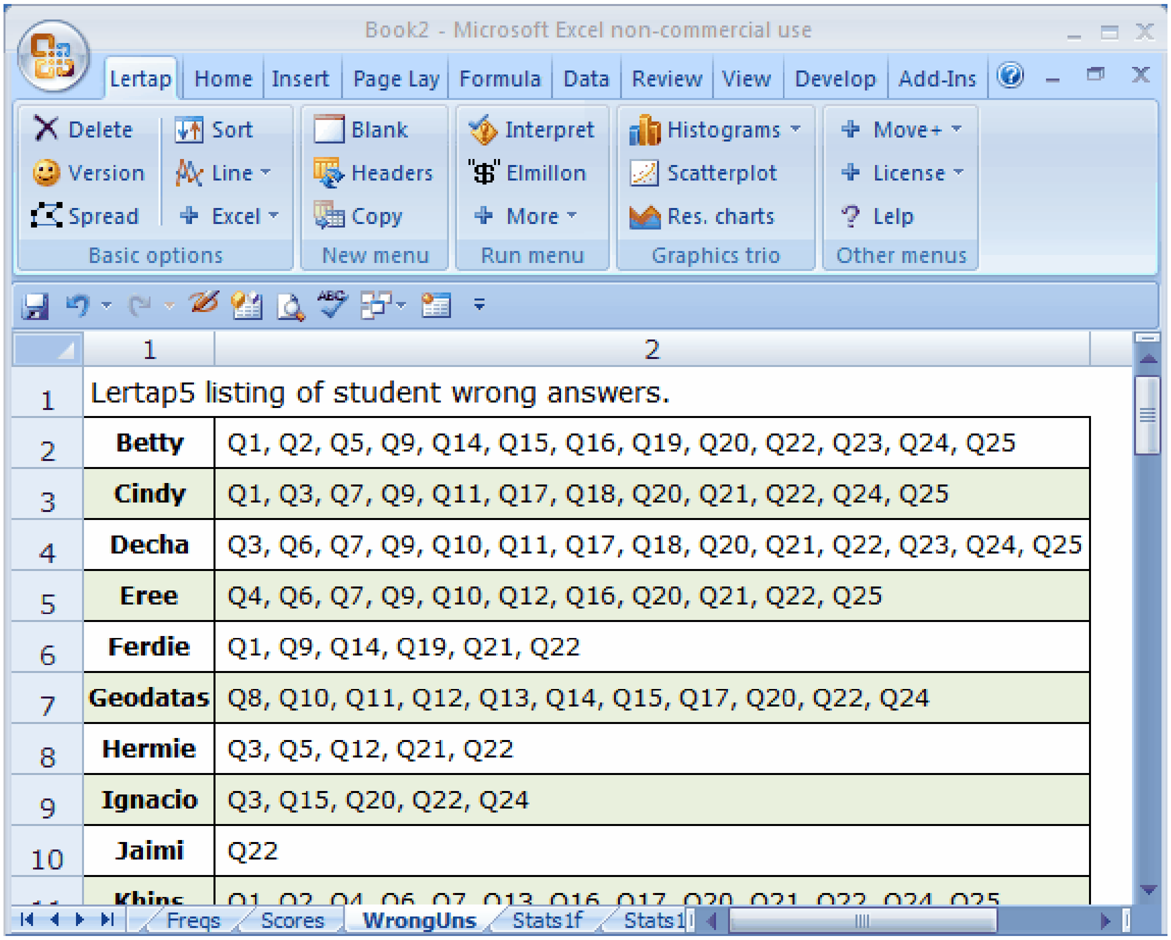
In this example, all subtests (J=2) have Wt=0, and no Total score will be made.

The *TotalTest1* macro will only work when two or more subtests have non-zero Wt. values. The subtests do not have to be of the same type, that is, some may be cognitive while others are affective.

(This macro was created at the request of a veteran Lertap user who knew that composite alpha used to be computed by the mainframe version of Lertap, Lertap 2. The P1, P2, and P3 terms are as seen on pp. 278 - 281 of the Lertap 2 user guide (Nelson, 1974).)

4.7.21 WrongItemList2

This macro creates the *WrongUns* report. A typical *WrongUns* list looks like this:



In this example, Jaimi missed only one question, Q22. Decha missed the most questions. Q24 seems to have been hard -- about half the students listed in the report got it wrong.

When there are two or more subtests, results are grouped, as seen here:

Lertap5 listing of student wrong answers.

2	Anton	(Running Q8) (Hiking Q2,Q17,Q20) (Football Q4,Q9,Q12,Q18) (Bsktball Q6,Q16,Q19) (Soccer Q21)
3	Bisekel	(Running Q10,Q15) (Hiking Q17,Q20) (Football Q4,Q13,Q18) (Bsktball Q6,Q14,Q16) (Soccer Q21,Q25)
4	Chu	(Running Q15) (Hiking Q17,Q20) (Football Q9,Q18) (Bsktball Q16) (Soccer Q25)
5	Dipak	(Running Q15) (Hiking Q17,Q20) (Football Q4,Q18) (Bsktball Q16) (Soccer ok)
6	Emil	(Running Q3) (Hiking Q11,Q17,Q20) (Football Q9,Q12,Q18) (Bsktball Q7,Q14,Q16) (Soccer Q22,Q24)
7	Fatima	(Running Q15) (Hiking Q17,Q20) (Football Q9,Q12,Q13) (Bsktball Q6,Q16) (Soccer Q21,Q24,Q25)
8	Greg	(Running Q15) (Hiking Q17,Q20) (Football Q9,Q18) (Bsktball Q6,Q14,Q16) (Soccer Q21)
9	Irawadi	(Running ok) (Hiking Q17,Q20) (Football Q18) (Bsktball Q6,Q16) (Soccer Q21,Q22)

When a student has no incorrect answers, 'ok' appears in the list (Dipak got all the Soccer items correct; Irawadi had no problem with the Running items, something we might expect of someone named after a major river; on the *not*-ok side of the ledger, results suggest that we may not want Anton on our football team).

Note: if a subtest has Wt=0 on its *sub line, it will not be included when this macro goes about its business. To read more about this Wt=0 thing, zero in on the previous topic.

SAQ: *What happened to WrongItemList1?*

Very good question. Nate Thompson, Assessment Systems Corporation, wrote WrongItemsList1 as a special macro for a Chicago-based Lertap site. They're still using it in the Windy City, and you can use it, too: it's in the Lertap5MacroSetA.xlam file, where it's simply called WrongItemList. It makes some assumptions regarding the nature of Lertap subtests, and may fail if your subtests use complex *col lines in CCs worksheets. For more assistance write: lertap5@gmail.com.

4.7.22 ZoomAllSheets

This macro does two things: it selects all the present worksheets in a workbook, and then zooms them (all of them) to the setting found in row 57 of the System worksheet. This topic mentions other ways to zoom sheets.

4.8 License Menu

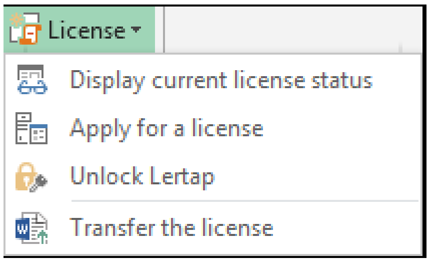
This topic assumes that you've installed a version of Lertap5 on a computer which is running Excel 2010, Excel 2013, Excel 2016, Excel 2019, Excel 365. Related background information is presented in other topics in this document: Requirements, How to get Lertap, and How to run it.

Lertap5, when first installed on a computer running the Windows operating system, will run as the so-called "Mini" version.

The "Mini" version never expires; it has no end date. But it's limited to processing no more than 250 data records. "Mini" may be upgraded to a standard version by purchasing a license from the Lertap e-store. The standard version will process thousands of test and survey results. (There is no limit to the number of items which may be processed.)

The License menu may be used as a means of starting the process of purchasing a license, but more often it is used by folks who have already obtained a license, and want to activate, or "unlock", their copy of Lertap5.

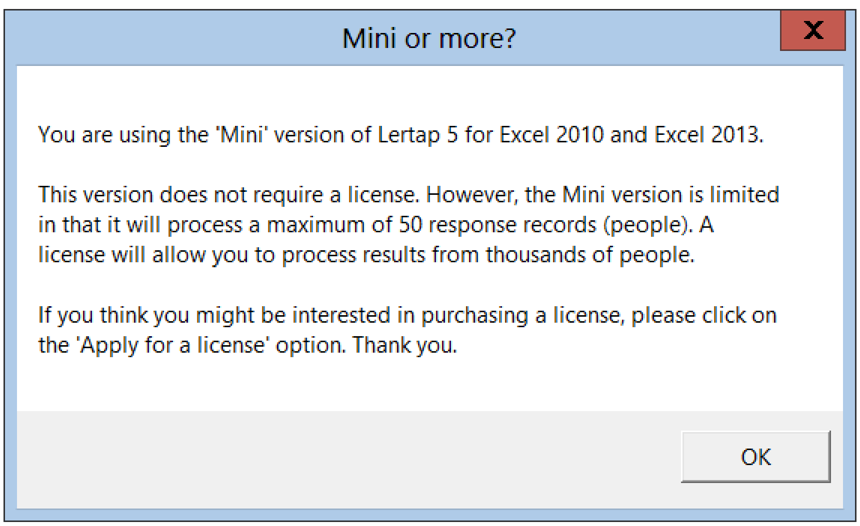
The menu format varies a bit depending on the version of Lertap5 installed, but it will look something like this:



Page forward to find out what the various options on the License menu do.

4.8.1 Display current status

If your computer does not yet have a license to use Lertap5, a message *similar* to the one below will appear when you click on "Display current license status".

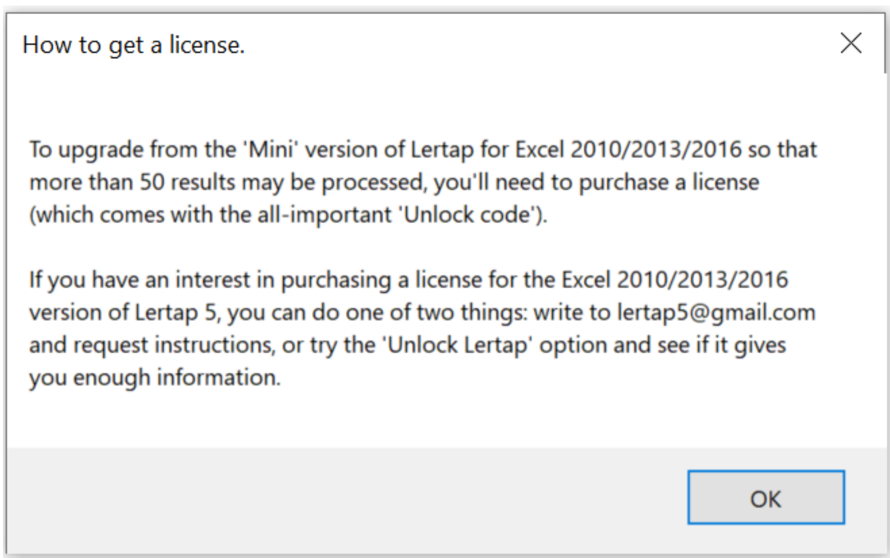


On the other hand, if your computer has a valid license to use Lertap5, you should see a message *similar* to this one:

Once you have obtained a Lertap5 license, it is possible to transfer the license to another computer. Information on how to do this may be found under the "Setup the license transfer process" option.

4.8.2 Apply for license

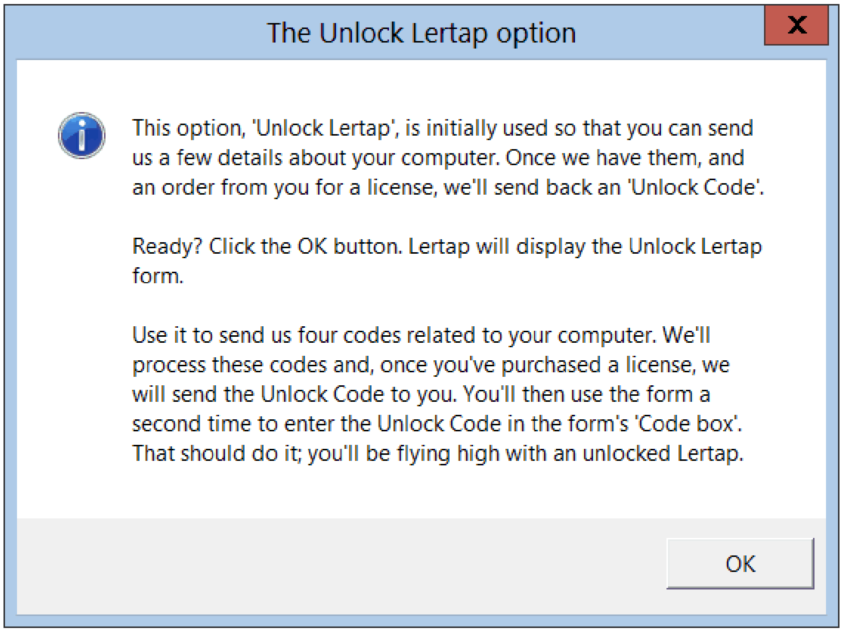
Click on the "Apply for a license" option, and a screen *similar* to the following should appear:



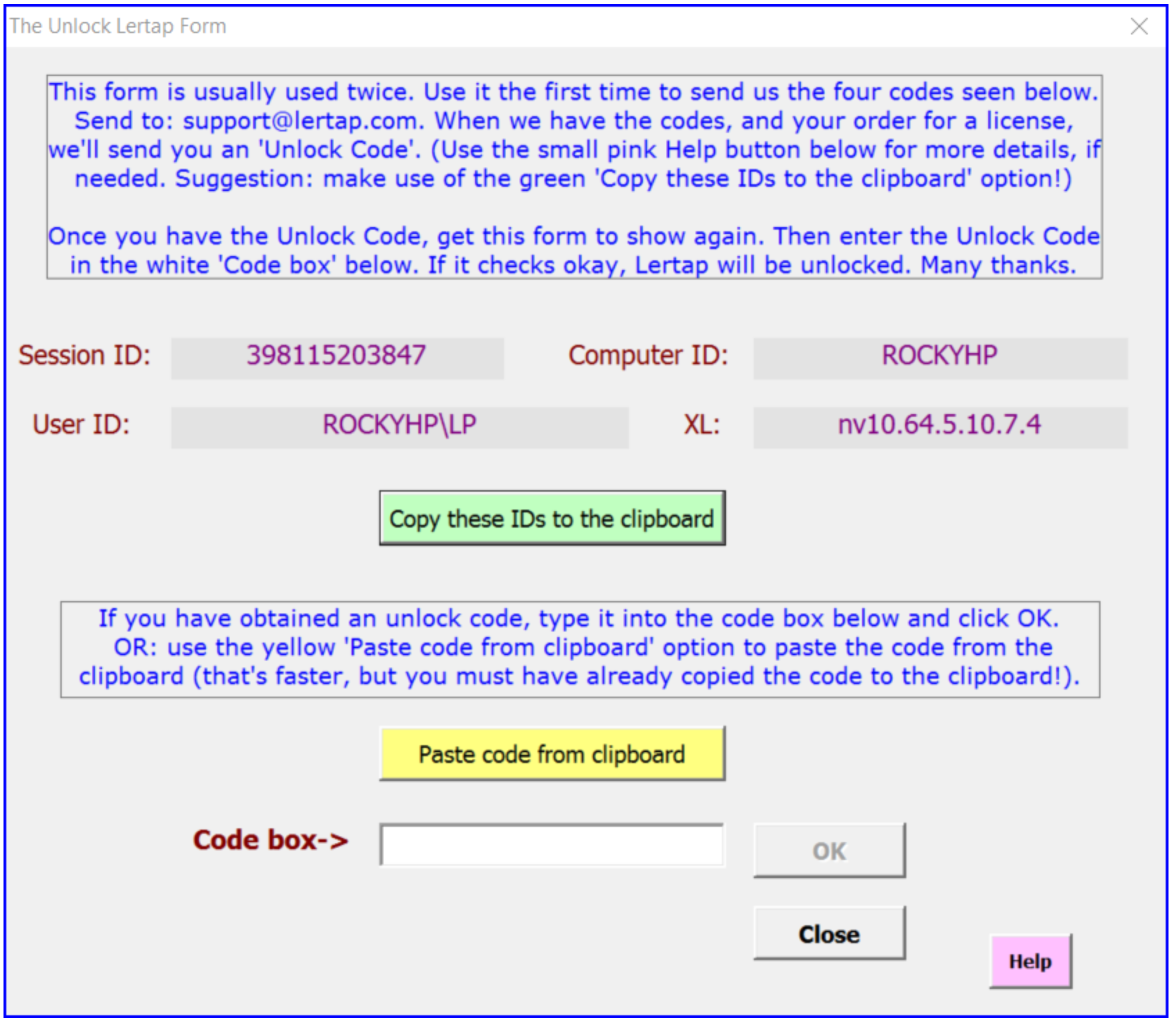
Licenses are purchased from the Lertap e-store.

4.8.3 Unlock Lertap

A message somewhat like the one below may be expected when the "Unlock Lertap" option is chosen from the License Menu:



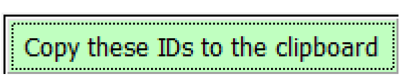
Click on the OK button, and a form similar to this one will be displayed:



Lertap 5 licenses are sold by the Lertap e-store. (Also see the How to get Lertap topic.)

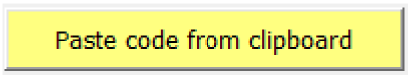
The text in the two boxes above mentions the steps to follow in order to "unlock" (or, activate) Lertap 5 on your computer. You are asked to send an email message to lertap5@gmail.com, including the four codes seen above: Session ID, Computer ID, User ID, and XL.

If you have already purchased a license, mention the invoice number, or the PayPal receipt number, when you send your message. If you have not yet obtained a license, please write to lertap5@gmail.com for assistance.



A button similar to this shows on the form above. Its purpose is to save you having to copy down the four codes. If you click this button, the codes will be copied to the Windows clipboard; you can then use the Paste command in your email program to have the codes automatically inserted in your message.

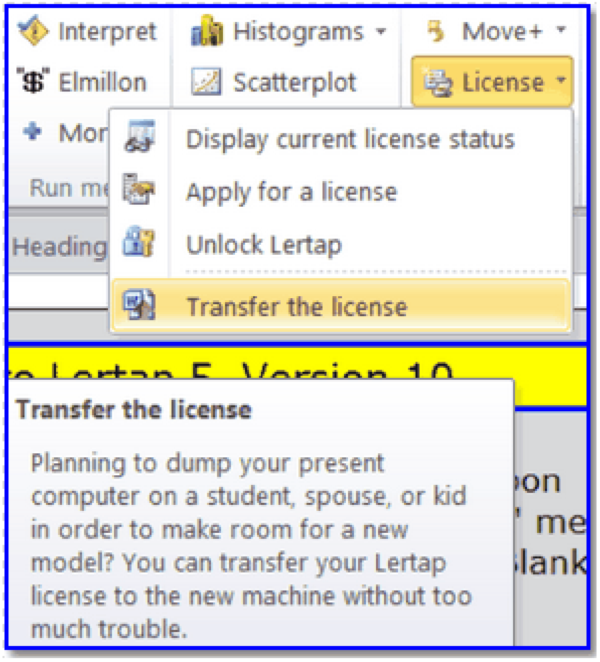
Lertap 5 is activated by entering the unlock code in the 'Code box->' seen towards the bottom of the form. The yellow box makes this process easier if the unlock code has first been copied to the computer's clipboard. (This document has more comments, if needed.)



Click OK. If the unlock code is correct, a message similar to this one should appear:

If a problem is detected with the code, a message will be displayed with a special error "string" which should be sent, along with a request for help, to lertap5@gmail.com. Please be sure to include the invoice number from your order when you write (or the PayPal receipt number).

4.8.4 License transfer



License holders sometimes wish they could shift their base of operations, getting Lertap to work on another computer without having to purchase another license. Can do? Yes. Certainly. However, please be advised that the transfer process can take more than a day or two as it involves some email exchanges, and there is a small cost involved (mentioned below).

The "Transfer the license" option involves a brief interactive session which ends up asking you to send an email with important data from your computer. The text for the email is created by Lertap and copied to your computer's clipboard; all you have to do is open a new email message addressed to lertap5@gmail.com with a subject line of "De-activate Lertap", and then paste the clipboard's content into the message area of the email.

Here's an example of the text that will be copied to the clipboard:

```
To my friends at <lertap5@gmail.com> :

I'm sending this message because I want to transfer Lertap 5 to another
computer.
Details about my present computer are shown below:

Session ID: 3665117049426
Enviado ID: 8664012428114563
Computer ID: MEGUSTA
User ID: MEGUSTA\NELSON2
XL: nv10.64
```

```
My computer's copy of Lertap 5 has been reset to 'Mini' mode.

Please acknowledge receipt of this message as soon possible so that I will
be able to activate Lertap on another computer.

Thank you.
```

After you've sent your message, Lertap 5 will still be running on your computer, but now it will be a "Mini" version, limited to processing 100 data records.

Upon receipt of your message, we'll acknowledge that you're now okay to install Lertap 5 on another computer, with the right to unlock it (or "activate" it) without again having to pay the full license fee. Note that Lertap headquarters are located in Western Australia, in the same time zone as Hong Kong. There's a big time difference between us and the Americas -- be patient -- we will acknowledge your message, just allow a day or two please.

License transfers are done on a cost-recovery basis, there's a small fee involved (as of January 2023 it was USD$39).

Important note: at one point the transfer process will display a code number for you to COPY DOWN to paper. This will be a fairly long code ending in a small letter such as u, v, w, x, y, or z. You should make a note of this code as it's important; should something go wrong with the email message, we are likely to ask you for the code (maybe take a picture or screen snapshot of the code if paper and pencil are not at hand).

4.9 Advanced Toolbar

THIS IS AN OLD TOPIC! It does not apply to the Excel 2007/2010/2013 versions of Lertap, but has been left here as the manual refers to it. Note that the Excel 2003 version of Lertap is still available from www.assess.com, and Excel 2003 will work fine even under Vista and Windows 7, Microsoft's new operating system. If you need the features offered by the advanced toolbar, consider using the Excel 2003 version of Lertap.

Other versions of Lertap, such as those running under Excel 2002, 2003, and 2011 (a Mac version), have what's referred to as an "advanced toolbar".

Here's a picture of it, as seen on the Excel 2003 version in January, 2005 (the one you see on your computer may differ a bit, that is, if you try to activate it -- please note that Macintosh Excel, Excel 2011, sometimes has a problem with this toolbar, and may fail to display it -- write to lertap5@gmail.com for assistance):

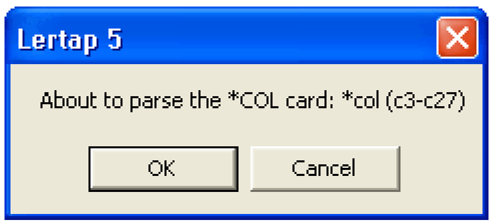
This toolbar has two special icons immediately to the right of Run. The first one, the magnifying glass, unhides and hides the Sub worksheets. Of course, it's always possible to hide and unhide Excel worksheets; this little button simply makes the job easier.

The next icon to the right, the bell, is used to run Elmillon for just a single subtest. This icon is referred to as the "Liberty Bell" in the manual's Chapter 10. (Elmillon is usually accessed via the Run menu, where it's advertised as "Elmillon item analysis". When you activate Elmillon via the Run menu, all subtests are processed.)

Before leading into why these two new icons are sometimes useful, to some, we should mention that the behaviour of the first icon on the toolbar, the X to the left of the yellow smiley face, is changed from normal. This is the "delete worksheets" icon.

Click here to be reminded of how the X normally behaves. When used from the advanced toolbar, the X does not delete the Freqs, Subs, and Scores worksheets.

We should also mention that the behaviour of the Run menu's "Interpret CCs lines" option is different when it's accessed from the Advanced Toolbar. You have the chance to skip the subtest corresponding to each *col line, or "card", as shown below:



A click on the Cancel button will cause Lertap to skip to the next *col card, without doing any processing.

Please read on into the next topics to get a grasp of why some people use the advanced toolbar (some of the time), and how they get it to show.

4.9.1 Advanced toolbar: why?

Perhaps one of Lertap's greatest strengths lies in its ability to fairly easily attach weights, or points, to each one of an item's responses. The *mws line is the way multiple weights are usually applied; mws stands for "multiple-weights specification". *mws "cards" live on CCs worksheets.

What many people do not realise, no matter how often they pull the Lertap manual down from their bedside table for a relaxing read before sleep, is that it's possible to change item weights without changing the control lines in the CCs worksheet.

Let's say this is what you've done: you've set up a nice CCs worksheet, and used the "Interpret CCs lines" option from the Run menu, after which you run "Elmillon item analysis", again from the Run menu.

You look at the output and realise that there's a need to re-weight some of the items. In fact, your need is so special, you're not even sure how you'd go about creating the *mws lines which seem to be required.

What you need to do is roll up your sleeves, perk some fresh coffee, and get into the depths of a Sub worksheet. You can poke any weights anywhere -- once you have a look at a Sub worksheet, it will (hopefully) be obvious what to do.

Here's a snippet from a Sub worksheet:

A	B	C	D	E	F					other
1.00	1.00	1.00	1.00	1.00	1.00					1.00

What we're looking at here is the weights array corresponding to an item from a subtest having Res=(A,B,C,D,E,F). The item could be an affective one; it could be a cognitive one -- the format of the array is the same. The numbers in the boxes, all 1.00 in this case, correspond to the number of points a person will get if s/he chooses one of the options, A through F.

There are four empty boxes for this item -- they'd also have weights (points) if the item used more response codes.

What's "other"? It's the number of points a person gets if her/his response to the item is not A, B, C, D, E, or F. The person may not have responded at all, in which case there might be a blank in the Data worksheet for this person on this item. Many times scanners will record unanswered items, or "funny" answers, as an asterisk (for example, when a person has shaded in more than one bubble on the answer sheet). Blanks and asterisks are caught by Lertap's other category -- in fact *anything* which is not one of the six recognised response codes for this item, A through F, will be classed as "other" in Lertap.

Consider again the weights array above. It's essentially saying that a person will get 1.00 points, no matter how s/he responded, or non-responded, to this item. Even unanswers get a point. Crazy? No, not exactly. There are times in the life of test scorers when it's necessary to do this.

Put this "crazy" scoring aside -- it's unusual, yes -- but what we want to point out is that you get can into a Sub worksheet, and poke away at the item weights. Any number, positive or negative, may be placed in any of the weights boxes.

True: *mws cards may be used to accomplish the same thing. In Lertap version 5.25, released 8 August 2003, the power of the *mws card was increased so that it could be used to get weights into the "other" category, as exemplified here:

```
*mws c12, 1, 1, 1, 1, 1, 1, other=1
```

An *mws line such as this will result in the same weights array seen in the boxes above. (For more examples of other= in action, use the Index, or Search for other=.)

This being the case, it is nonetheless true that some people have a preference for launching their response weights from within a Sub worksheet. The advanced toolbar is for such users.

4.9.2 Advanced toolbar: how use?

To review, the advanced toolbar has two more icons than the normal toolbar. They appear immediately to the right of the Run menu, as seen here:

The magnifying glass unhides or hides the Sub worksheets. It's called a "toggle"; it's an on-off switch.

The Liberty Bell calls in Elmillon, the item analysis program. To use it you must have a Sub sheet open.

Say, for example, that you've been changing weights in a Sub worksheet called Sub3. You're a good scout -- you've gone to Excel's File menu, and saved Sub3 after making the changes. Now you click on the Liberty Bell. In comes Elmillon; out come your results: a new column is added to the Scores worksheet, and then those lovely Stats reports are created, Stats3f, Stats3b, and, if Sub3 corresponds to a cognitive test, probably Stats3ul too. (If these Stats3 reports existed before, they will be overwritten.)

You bewdy.

You look at the results. Hmmm ... maybe give more points to option D on item 21. You return to Sub3, scroll down to item 21, and increase the number of points corresponding to option D. You save the worksheet. You tickle the Liberty Bell. Another column is added to the Scores worksheet -- you can compare the new scores with the last ones. And once again you get all the Stats3 reports.

Another way to go about this: make a copy of Sub3. Call the copy (say) Sub3B. Make the changes in Sub3B. Maybe even change the Subtest Title at the top so the new score will have a new moniker.

Then, with the Sub3B sheet in focus, ring the Liberty Bell. You'll Scores again (onya!), and now you'll have Stats3Bf, Stats3Bb, and Stats3Bul (?) reports to look at. (The Stats3f, 3b, and 3ul reports from the last run will remain unchanged.)

You bewdy Newk!

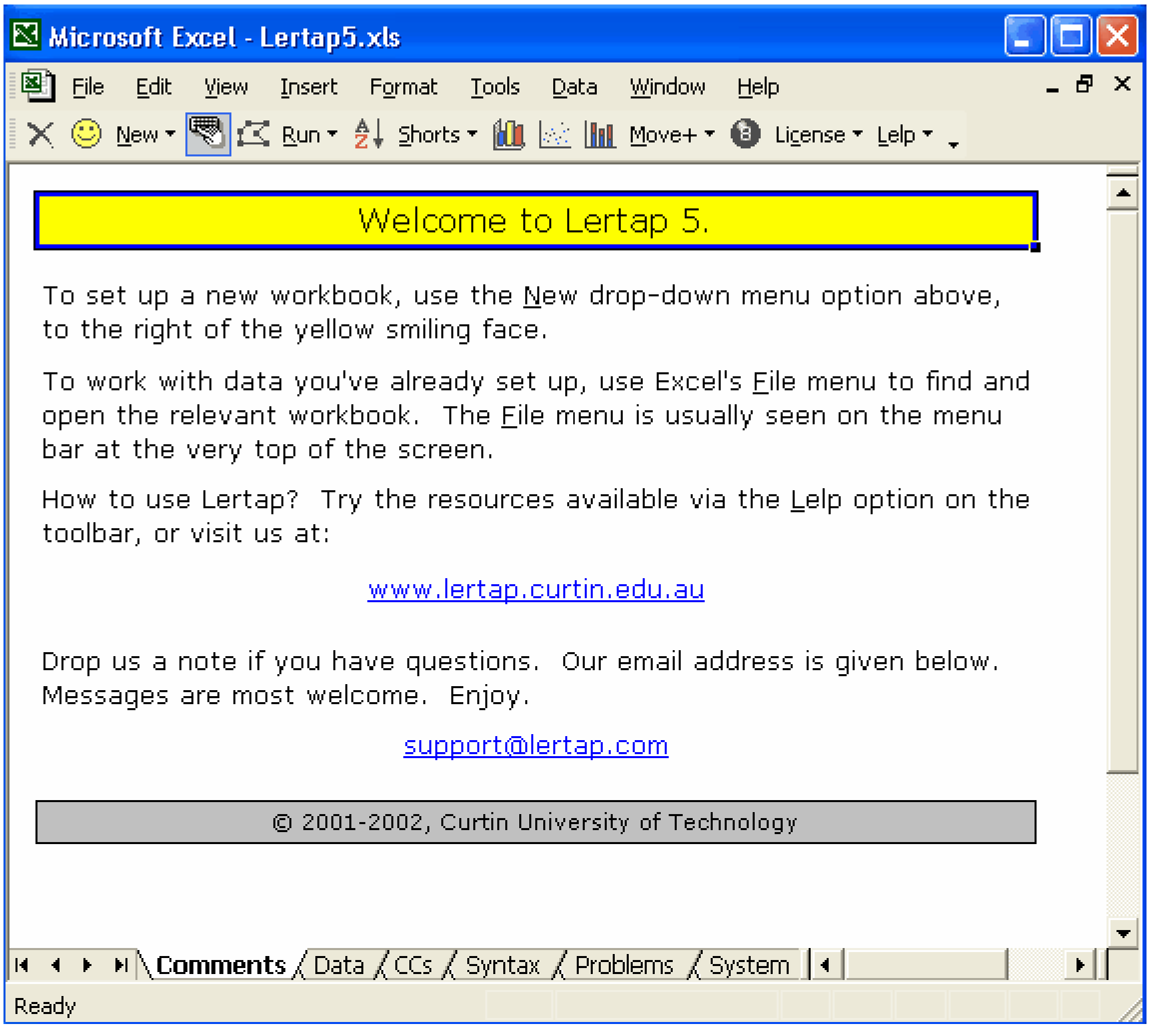
4.9.3 Advanced toolbar: how show?

THIS IS AN OLD TOPIC! It does not apply to the Excel 2007/2010/2013/2016/2019/365 versions of Lertap, but has been left here as the manual refers to it. Note that the Excel 2003 version of Lertap is still available from www.assess.com, and Excel 2003 will work fine even under Vista, Microsoft's new operating system for Windows. If you need the features offered by the advanced toolbar, consider using the Excel 2003 version of Lertap.

How to get the advanced toolbar to show? Go to the System worksheet. Change the UserLevel setting in row 7 to 2. Save. Close Lertap. Reopen Lertap. The advanced toolbar should be there. Write to us if you have probs: lertap5@gmail.com; ring the bell if you don't.

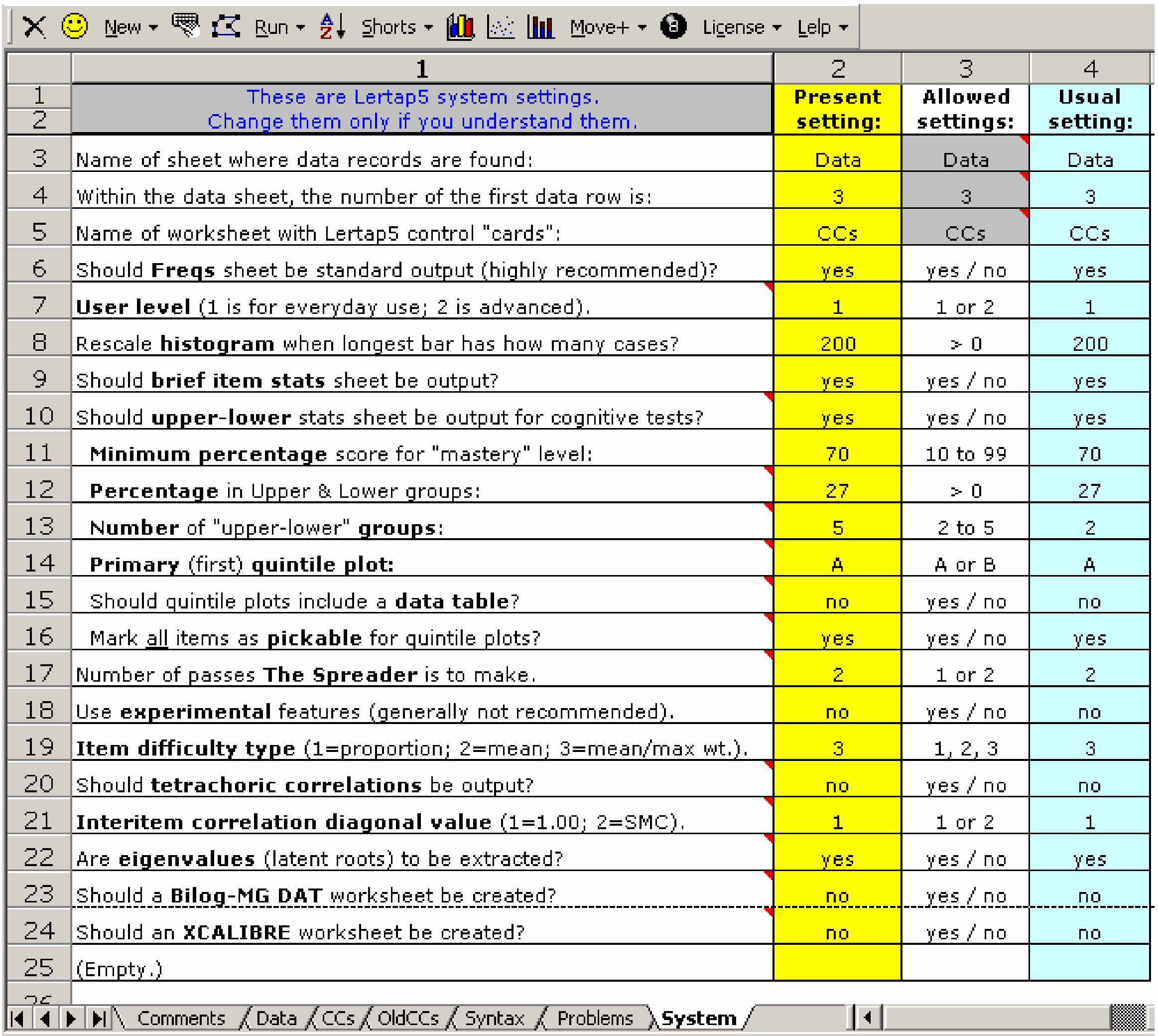
4.9.4 System Worksheet

The System worksheet is one of the worksheets found in the Lertap5.xlsm workbook -- you can see its tab down towards the bottom of this screen snapshop:



The Comments worksheet is displayed above. To see the System worksheet, just click on its tab.

The System worksheet looked like this as of February, 2005:



New Run Shorts Move+ License Lelp

	1	2	3	4
1	These are Lertap5 system settings.	Present	Allowed	Usual
2	Change them only if you understand them.	setting:	settings:	setting:
3	Name of sheet where data records are found:	Data	Data	Data
4	Within the data sheet, the number of the first data row is:	3	3	3
5	Name of worksheet with Lertap5 control "cards":	CCs	CCs	CCs
6	Should **Freqs** sheet be standard output (highly recommended)?	yes	yes / no	yes
7	**User level** (1 is for everyday use; 2 is advanced).	1	1 or 2	1
8	Rescale **histogram** when longest bar has how many cases?	200	> 0	200
9	Should **brief item stats** sheet be output?	yes	yes / no	yes
10	Should **upper-lower** stats sheet be output for cognitive tests?	yes	yes / no	yes
11	**Minimum percentage** score for "mastery" level:	70	10 to 99	70
12	**Percentage** in Upper & Lower groups:	27	> 0	27
13	**Number** of "upper-lower" **groups**:	5	2 to 5	2
14	**Primary** (first) **quintile plot:**	A	A or B	A
15	Should quintile plots include a **data table**?	no	yes / no	no
16	Mark all items as **pickable** for quintile plots?	yes	yes / no	yes
17	Number of passes **The Spreader** is to make.	2	1 or 2	2
18	Use **experimental** features (generally not recommended).	no	yes / no	no
19	**Item difficulty type** (1=proportion; 2=mean; 3=mean/max wt.).	3	1, 2, 3	3
20	Should **tetrachoric correlations** be output?	no	yes / no	no
21	**Interitem correlation diagonal value** (1=1.00; 2=SMC).	1	1 or 2	1
22	Are **eigenvalues** (latent roots) to be extracted?	yes	yes / no	yes
23	Should a **Bilog-MG DAT** worksheet be created?	no	yes / no	no
24	Should an **XCALIBRE** worksheet be created?	no	yes / no	no
25	(Empty.)			

Comments / Data / CCs / OldCCs / Syntax / Problems / **System**

If you're looking at this page using an interactive version of Lelp (not a printed PDF version), you should find that some of the rows above may be clicked on -- many of the rows hyperlink to corresponding Lelp topic pages.

If the row of interest to you does not hyperlink, you might try referring to the manual for a suitable discussion; as an example, Row 12, the percentage of N found in upper and lower groups, is not discussed in Lelp: see Chapter 10 of the manual (look at page 166 if you have the good fortune of having a printed copy of the manual).

We used to ship Lertap with the System worksheet hidden in order to discourage less experienced users from fooling around and getting into trouble. Trouble is especially likely to result if the settings in Rows 3, 4, and 5 are altered.

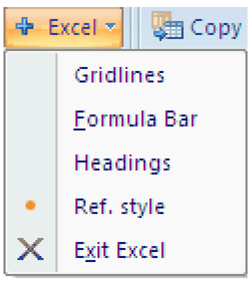
If one of the Present Settings is changed, sometimes the Lertap5.xlsm workbook has to be closed and then re-opened before the change will be effected. However, numerous settings take effect immediately, and do not require closing Lertap. Among these are the settings seen above in Rows 12 through 17, and in Rows 19 through 24 -- if you alter the column 2 setting in one of these rows, the change will take effect without having to close Lertap and then re-open it. (But don't read too much into this. If, for example, you change the item difficulty calculation method via row 19, the item difficulties you've already obtained will not change until you once again use the Run menu.)

Note 1 inserted May, 2005: a "production mode" capability was added to Lertap, with relevant options set in rows 26 through 29 of the System Worksheet. A simple click here will take to you the corresponding topic. (The row-position of these options changed after the following was inserted.)

Note 2 inserted July, 2005: support for response-similarity checking was added, with relevant options set in rows 25 through 27 of the System Worksheet. A simple click here will take to you the corresponding whatchamadoodle.

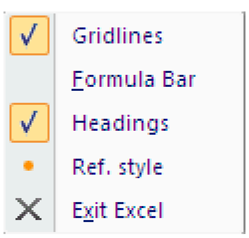
5 Excel shortcuts

See this topic for an updated version of this page.



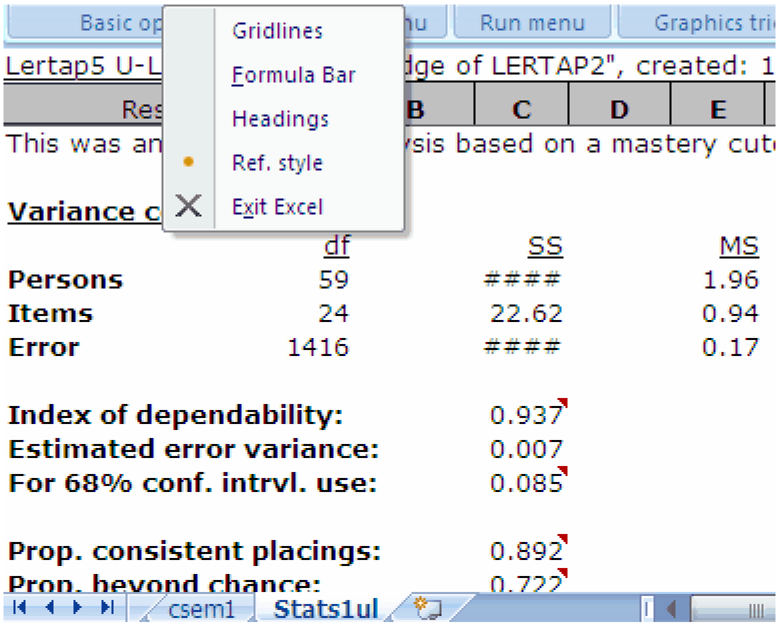
The snapshot above shows the Excel shortcuts available with the Excel 2007/2010/2013 and later versions of Lertap 5.

Experiment with these options. They are 'toggles', turning on and off corresponding options.



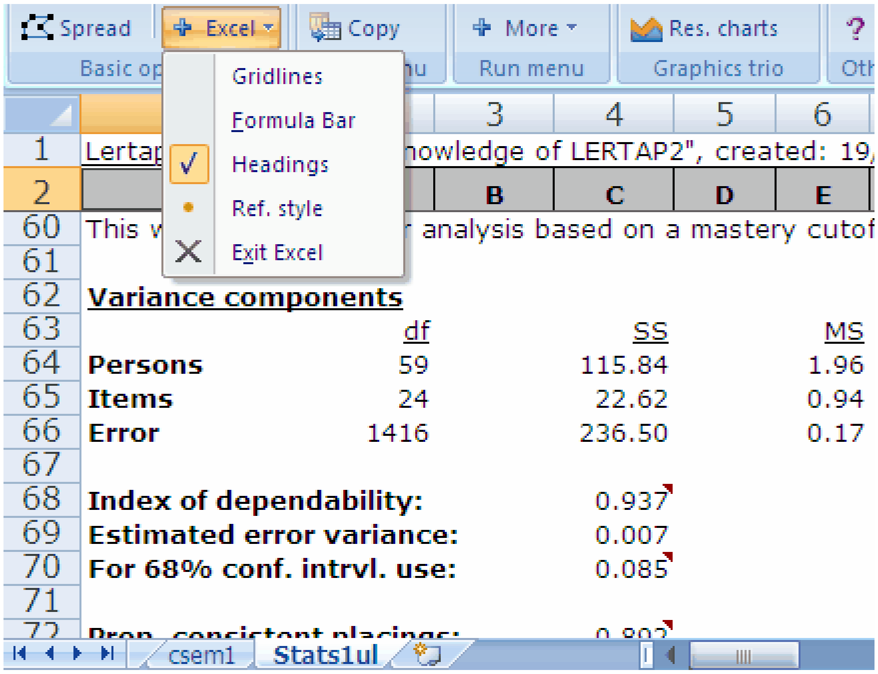
When an option has been turned on, a tick mark will appear next to it. Here the Gridlines and Headings options are on.

Lertap will commonly turn these options on and off as it thinks best, depending on the worksheet in view. However, this is not always a good thing. Look, for example, at the Stats-ul report below:

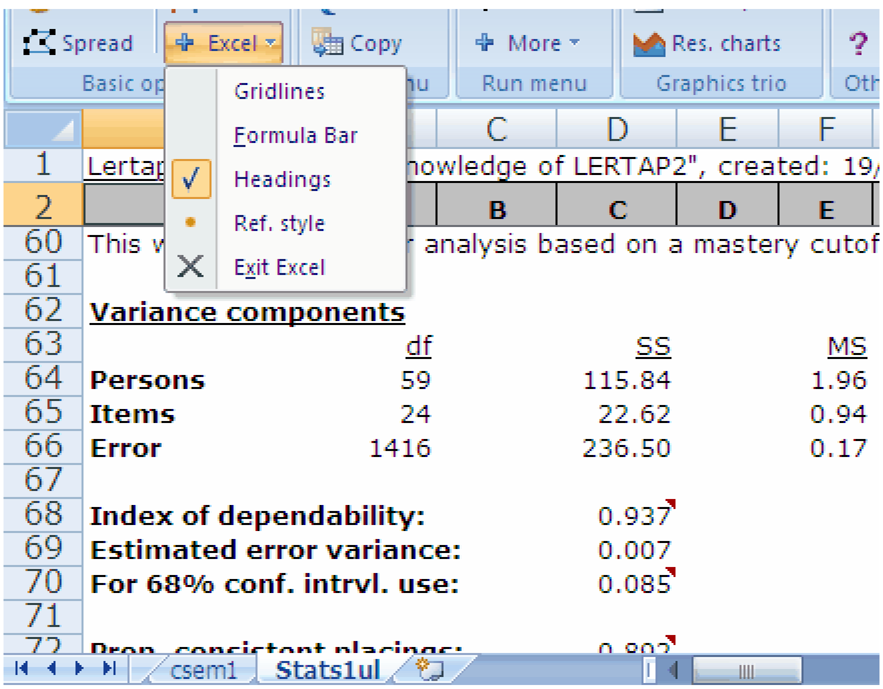


Lertap's Stats-ul reports always have their Gridlines and Headings options set to off, and what has happened above is not uncommon: the column with SS data is not wide enough to display all values. Whenever this happens, Excel's reaction is to fill fields with #### symbols.

This problem is fixed by turning the Headings option on, and then widening the corresponding column (column 4 in this case):



The 'Ref. style' option turns R1C1 referencing on and off. When R1C1 is on, column heading are numbers, as displayed above. When it's off, column headings are letters, often called the A1 referencing style:



Now the column with SS figures is column D; if the 'Ref. style' option were clicked again, the column with SS figures would be denoted as column 4.

Which referencing system is best? Whichever you like -- the 'Ref. style' option makes no difference to the format of a worksheet, nor does it affect how Lertap runs. Lertap prefers the R1C1 style, and will always try to turn it on if it can.

6 Input & Output

Data analysis systems ask users to do at least three things: input information, describe how the information is to be analyzed, and, when ready, signal that the analysis should start.

There are always some constraints on how things are to be done. Lertap 5 is no exception; it uses Excel to accomplish much of its work, but not just any Excel workbook will meet Lertap's requirements.

Lertap wants its Excel workbook to have a worksheet named Data. This is where the answers respondents have given to test or survey items are recorded. Lertap wants its description of how the information is to be analyzed to be expressed as lines in another worksheet, one named CCs.

The Data and CCs worksheets are referred to as "primary" worksheets. It's the information from these two worksheets that enables Lertap to go about the business of creating its output: the various reports found in worksheets such as Stats1f, Stats1b, and Scores -- these worksheets are referred to as "secondary" worksheets. Users create the primary worksheets; Lertap makes the secondary ones.

The following topics get into some of the specifics of Lertap's primary and secondary worksheets.

Related tidbits:

For the definition of a Lertap workbook, please click here.

For information on how to create a new Lertap workbook, simply click here and you'll be whisked away to a discussion of the Lertap toolbar's New menu.

6.1 Lertap workbook def.

A Lertap 5 workbook is an Excel workbook with the particular features mentioned here.

A Lertap 5 workbook will *always* have at least two primary worksheets: one of these is named Data, the other is named CCs.

The Data worksheet's top two rows are reserved for titles. The first row may contain any text; it's used as a means of briefly describing the data found in the worksheet. For example, the first row might say "Data collected 10 March 2023 in Psych 501".

The second row of the Data worksheet contains column headers. For example, if the first column of each data record is some sort of ID field, then the contents of the worksheet cell formed by Row 2, Column 1 might be "ID Number". If responses to the first item are found in column 5 of each data record, then the contents of the worksheet cell formed by Row 2, Column 5 might be "Item5".

Data records begin in Row 3 of the Data worksheet.

If the first column of any row in the Data worksheet is empty, or contains a zero, then that is considered to be the end of data. (Users sometimes use this fact when they're testing their CCs lines to make the test go faster, a blank line may be inserted after, say, 10 data records -- this stops Lertap from reading all the data records.)

If the first column of a row in the CCs worksheet is blank, then that is considered to be the end of the CCs lines -- Lertap will not read beyond this line. (This is useful when a user only wants to have the Freqs worksheet produced -- in this case the first CCs line will be *col, and the second line will be blank -- Lertap will produce its Freqs worksheet, and nothing else.)

Ideally, a Lertap 5 workbook has its default font set to Verdana, with the CCs worksheet being an exception in that it may at times make use of the Courier New font.

Lertap workbooks may have a number of secondary worksheets in addition to the primary ones. Examples of secondary worksheets are Freqs, Stats1f, Stats1b, Scores, and so on. Secondary worksheets are usually the result of applying a Lertap or Excel function; for example, the "Elmillon item analysis" option on the Run menu reads data records from the Data worksheet, a primary worksheet, and produces such secondary worksheets as Stats1f and Stats1b.

(For a related topic, see Deleting secondary worksheets.)

Lertap workbooks may also have other user-created worksheets. For example, in data sets with more than one subtest, or scale, users will sometimes create a codebook worksheet which keeps track of the location of the items comprising the subtests or scales.

Note that a Lertap workbook does *not* include the Lertap toolbar. It bears mentioning that the Lertap5.xlsm file is indeed a Lertap workbook, but, if someone uses the term "Lertap workbook", they are not necessarily referring to Lertap5.xlsm. We could say that the Lertap5.xlsm workbook is a very special Lertap workbook. Why? Because the Lertap5.xlsm workbook includes the Lertap toolbar, something no other Lertap workbook will have.

6.2 Data sheet

The name of the Excel worksheet where data are recorded for Lertap analysis has to be "Data". The first two rows of the Data worksheet are for header information, as described in the definition of a Lertap workbook.

Have a look at the top of a typical Lertap Data sheet:

	1	2	3	4	5	6	7	8	9
1	Data collected from principals and teachers, April 2003.								
2	ID	Position	Experience	Gender	Q1	Q2	Q3	Q4	Q5
3	A	3	2	2	4	3	4	4	4
4	B	3	1	2	2	1	3	4	3
5	C	1	3	2	4	2	4	4	4
6	D	1	2	2	3	3	3	3	3
7	E	1	3	1	3	2	4	4	4
8	F	1	3	1	2	2	3	4	4
9	G	1	2	2	4	3	2	4	3
10	H	1	3	2	4	3	3	4	4

Row 1 is a general "header", or title, which can contain any information you wish, including nothing at all. Whatever is typed in this row will not appear anywhere else; Lertap doesn't read this row. This row is for your own use -- we use it to provide a brief reminder of the information contained in the Data sheet.

Row 2 also has header information. Each column has been given a header, or label: ID; Position; Experience; Gender; Q1; Q2; and so on.

For your information, the CCs sheet corresponding to this workbook had these two lines:

```
*col (c5-c64)
*sub aff, res=(1,2,3,4)
```

Item responses begin in column 5 of the Data sheet, and continue through column 64.

Lertap will use the labels found in row 2, columns 5 through 64, as item IDs. That is, the ID for the first item will be Q1; for the second item Q2; ... and Q60 for the last item (not shown above).

Item IDs can be anything, and in theory can have any length. However, we strongly suggest that items IDs be short -- not greater than 8 characters in length. Valid examples of item IDs: Item1; Preg.2; Soal3; Ques2b; SD204; Likrt17a. Having short item IDs makes parts of Lertap's output easier to read; for example, the Stats1f report has a section which looks like this:

item difficulty bands

.00: Q22
.10:
.20:
.30:
.40: Q1 Q2 Q9 Q11 Q14 Q18 Q19 Q20 Q21 Q25
.50: Q3 Q4 Q6 Q7 Q10 Q12 Q15 Q17 Q24
.60: Q8 Q13 Q16 Q23
.70: Q5
.80:
.90:

The item IDs play a prominent role in tables such as that seen above; the longer the item IDs, the more cluttered the tables look.

If item labels are not found in row 2 of the Data sheet, Lertap will automatically create item IDs of this sort: Item1, Item2, Item3, ...

If it's desired to include ID information for the respondents, such information may be recorded in any column of the Data worksheet (but: Lertap versions dated before July 2004 have to have the ID in either the first or second column). Lertap will use the IDs to label the scores found in its Scores report, providing the respective column header begins with the letters "ID", or "id". Click here for more about this.

Excel has two reference styles used to refer to the rows and columns of its worksheets. Lertap uses what's called the "R1C1" style. In R1C1 notation, the columns of an Excel worksheet are numbers. In the other style, called the "A1" style, columns are labelled alphabetically.

Excel's default referencing style is A1. When Lertap starts up, it automatically changes this to R1C1. Later, when Lertap is closed, it will set the style back to A1 if that's what was in use before Lertap was started. (The referencing style may be manually set by Excel's Tools / Options / General tab.)

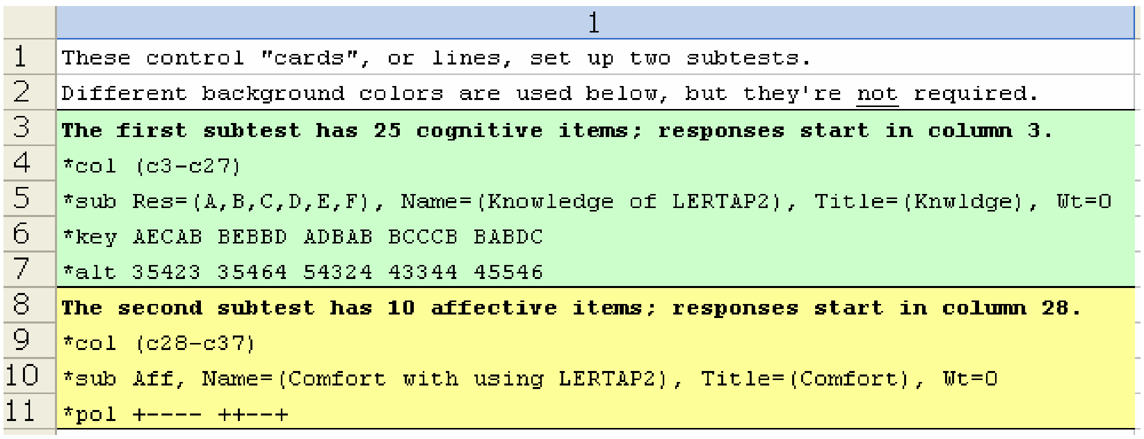
How does Lertap find the end of the Data records? It thinks it's come to the end when it finds a row whose first column is empty, or whose first column contains a blank, or whose first column contains a zero. Because of this, it is generally a good idea to see that the first column of the Data worksheet is used for something other than an item response. We say this as non-responses to items are often recorded as a blank -- try to keep blanks out of the first column.

At Lertap central, when in the process of testing out large new data sets, we often insert a blank row in the Data worksheet after row 52. This effectively fools Lertap into thinking there are just 50 respondents (remember: the first two Data rows are for header information); in turn, this lets us test our CCs cards faster, enabling us to quickly see if we've set up the cards required to get the analyses we wanted.

6.3 CCs sheet

The formatting of the CCs worksheet is relatively "ad hoc" when compared to the formatting of the Data worksheet. For example, there's no requirement to have rows with header information.

We almost always include comments in our CCs sheets, rows which remind us what we've done, and when. We sometimes even add some sort of colour coding to our CCs worksheets, as shown below:



	1
1	These control "cards", or lines, set up two subtests.
2	Different background colors are used below, but they're not required.
3	**The first subtest has 25 cognitive items; responses start in column 3.**
4	*col (c3-c27)
5	*sub Res=(A,B,C,D,E,F), Name=(Knowledge of LERTAP2), Title=(Knwldge), Wt=0
6	*key AECAB BEBBD ADBAB BCCCB BABDC
7	*alt 35423 35464 54324 43344 45546
8	**The second subtest has 10 affective items; responses start in column 28.**
9	*col (c28-c37)
10	*sub Aff, Name=(Comfort with using LERTAP2), Title=(Comfort), Wt=0
11	*pol +---- ++--+

The CCs sheet above has four rows with comments: rows 1, 2, 3, and 8. Any row which does not begin with an asterisk is considered to be a comment.

We also like to use a fixed-pitch font, such as Courier New, with the CCs worksheet. This makes the information in the CCs cards line up, as seen above in the *key and *alt cards.

Lertap only looks at the first column of the CCs sheet; any information found in subsequent columns is ignored.

When Lertap finds a CCs row whose first column is empty, it thinks it's come to the end of the CCs lines, and reads no more. This can be handy -- we sometimes enter a single *col card in the CCs worksheet, followed by a blank row. At other times, we'll step into a CCs sheet, and insert a blank row after the *col card. We do this as we know that this will get Lertap to make its Freqs report, but nothing else. Maybe with time you'll come to like the Freqs report as much as we do: it provides a quick, no-frills look at our data. We use it to rapidly get a glimpse of how people answered our questions, and to see if some errors may have arisen whilst processing the data. For example, if the items in our test used the default response codes of (A,B,C,D), we'd be surprised to find Freqs reporting it found an E as one of the item responses.

About here we again insert a critical note regarding response codes (this same message may be found under the Cognitive CCs topic). As all readers know, the Data worksheet contains item responses. In the case of cognitive tests, it is common for item responses to be coded as letters, such as the set {A,B,C,D}. For affective items, {1,2,3,4,5} is a popular response code set. If we look down the columns of the Data worksheet, these response codes are what we see -- *"but of course"*, you might say. And well enough. But: what sometimes happens is that users mis-match the codes found in the Data worksheet with the codes found in a *sub card's Res= declaration. For example, if the Data worksheet shows responses as being from the set {A,B,C,D}, and if Res=(a,b,c,d), there will be a crash -- the Res= declaration is wrong -- the Data worksheet uses upper-case letters.

Now, item responses can be just about anything. The answers to cognitive items can be coded as digits; the answers to affective items may be coded as letters. If letters are used, they may be upper case, or lower case. But in all cases, the Res= declaration has to be "fair dinkum"; if the Data sheet uses lower-case letter, then so must the Res= declaration.

Finally, remember the default Res= assignments. For cognitive tests, the default is Res=(A,B,C,D). For affective tests the default is Res=(1,2,3,4,5). If a *sub card has no Res= declaration on it, these default settings are assumed.

6.4 Output

What are the steps in a Lertap analysis? Create a new Lertap workbook. Enter item responses in the Data worksheet. Make up the "control cards" for the CCs worksheet.

Go to the Run menu on the Lertap toolbar. Click on "Interpret".

What happens? Lertap has a squiz of the CCs cards, reading down the rows of the CCs sheet until it encounters a row whose first column is empty.

If an error is found in the CCs lines, Lertap stops and makes an effort to tell you what the error is. You'll need to fix the error, and then click once again on the Interpret option.

If the CCs lines appear to be error-free, Lertap then starts to read the records in the Data worksheet, going down the Data rows until it encounters a row whose first column is either empty or contains a zero.

This process initiates the production of Lertap output, the secondary worksheets often referred to as Lertap "reports". The first of these is called "Freqs"; simultaneously, Lertap makes the behind-the-scenes series of "Sub" worksheets. There will be one Sub worksheet for each subtest, that is, for each *col card found in the CCs worksheet.

It then hides the Sub worksheets, brings Freqs to the fore, and announces that it's ready for you to squiz the Freqs. If you're satisfied with the squiz, you return to the Run menu, and select the "Elmillon" option. Note that there are settings in the

System worksheet which get Lertap to roll from the Interpret option to Elmillon automatically, non-stop. Lertap is said to be in "production mode" when this happens.

Once Elmillon starts up, what happens? Quite a bit. Lertap returns to the Data worksheet and reads all its records, extracting the responses corresponding to the first *col card in CCs, forming item statistics, and making a subtest score for each respondent.

It writes the subtest scores "on the fly", that is, as it reads the Data records (you can sometimes see it doing this). Writes them to where? To the "Scores" worksheet, another new sheet which Lertap adds to the workbook.

Then Lertap usually creates its two main reports with item statistics. Each of these reports is a new worksheet. They're called "Stats1f" and "Stats1b", respectfully containing "full" and "brief" item statistics. If the subtest is a cognitive one, Lertap usually creates another new worksheet, "Stats1ul", with upper-lower discrimination and difficulty estimates. (In the process of making the Stats1ul report, Lertap creates a temporary worksheet called "Scratch". It deletes this worksheet on completing the Stats1ul report.)

What happens next? If there's more than one *col card in the CCs worksheet, Lertap repeats this process. Each *col card defines a Lertap subtest. For each and every subtest, Lertap adds a score to the Scores worksheet, and creates the appropriate series of Stats reports.

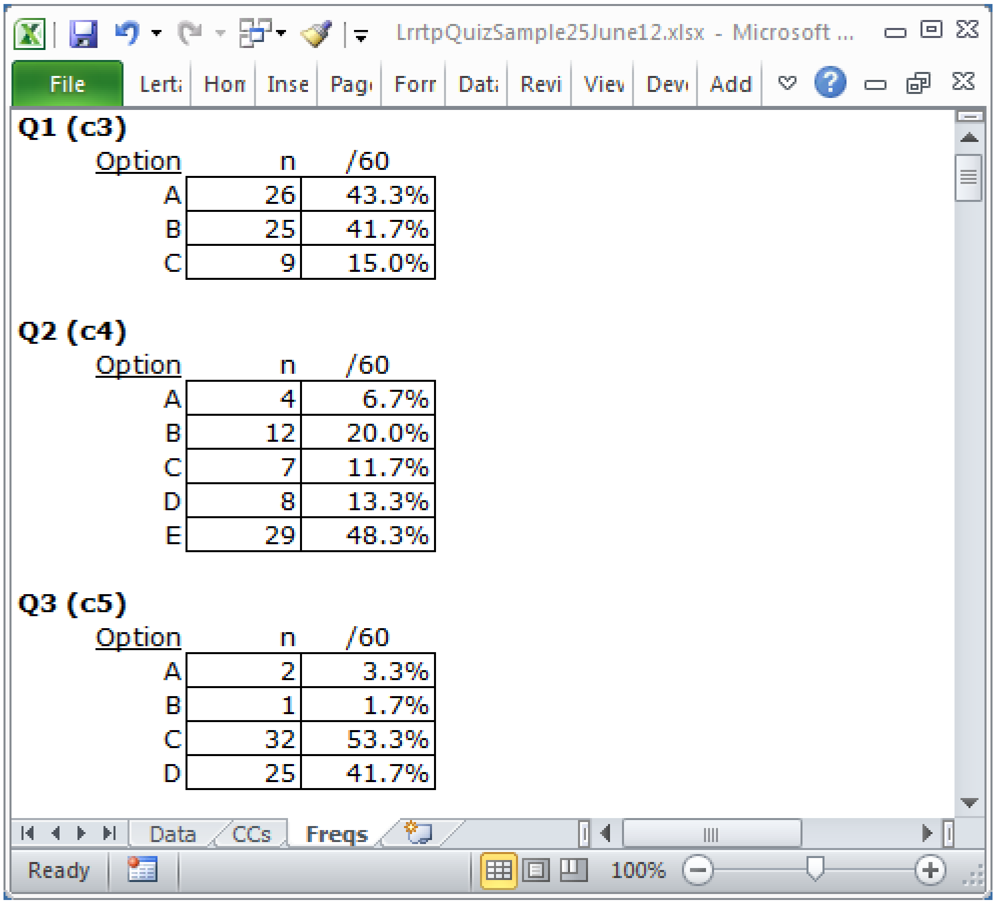
As to nomenclature, the Stats reports for the first subtest are Stats1f and Stats1b; for the second subtest they're Stats2f and Stats2b; ... and so on.

It is possible to control the number of reports made by Lertap. The Stats1f report is always standard, but the Stats1b and Stats1ul may be turned off. How? By making alterations in lines 9 and 10 of the System worksheet. The computational resources used to make the Stats1ul report are rather extensive; turning off this report will usually save a noticeable amount of processing time, especially when there are more than 500 records in the Data worksheet.

Freqs, Scores, and the suite of Stats worksheets comprise Lertap's main output. But it's possible to get Lertap to deliver more. For example, it'll make histograms, scatterplots, item response charts, a worksheet with item scores, subtest correlation matrices, and even another Stats report, ECStats1f (resulting from an external-criterion analysis).

6.4.1 Freqs

"Freqs" means "Frequencies". When you ask Lertap to "Interpret CCs lines", it does a bit more—it has a close look at the data columns referenced in *col lines, and makes a tally of response popularities, or "endorsements". Look:



Q1 (c3)

Option	n	/60
A	26	43.3%
B	25	41.7%
C	9	15.0%

Q2 (c4)

Option	n	/60
A	4	6.7%
B	12	20.0%
C	7	11.7%
D	8	13.3%
E	29	48.3%

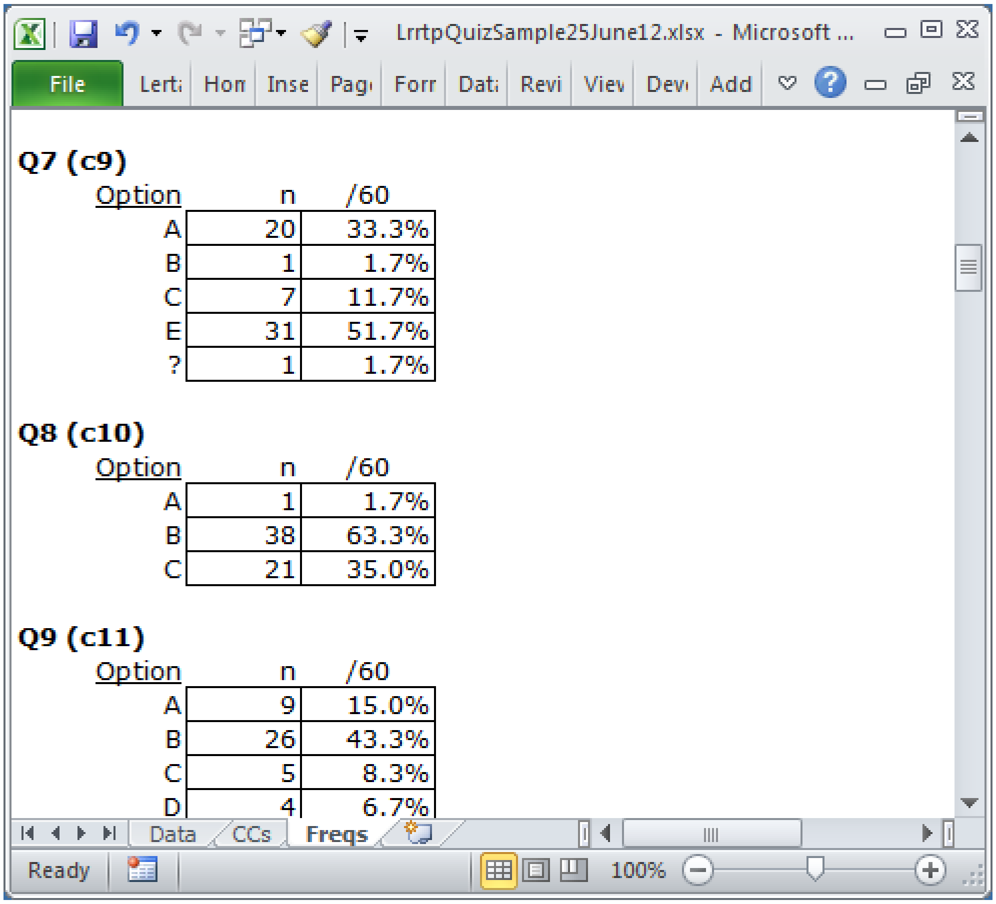
Q3 (c5)

Option	n	/60
A	2	3.3%
B	1	1.7%
C	32	53.3%
D	25	41.7%

For each item, or question, Freqs reports on the number of times letters or digits were found in the respective item column number in the Data worksheet. For Q1 there were 26 As, 25 Bs, and 9 Cs. The (c3) indicates that Q1 was found in column 3 of the Data worksheet (in this example).

The Freqs column headed "/60" indicates that a total of 60 data records were found, and gives the percentage associated with each frequency. For example, there were 26 As for Q1, which is 43.3% of the total number of 60.

As you scroll through your own Freqs sheets, you may come across some rows which have a ? mark on the left. For example:



Q7 (c9)

Option	n	/60
A	20	33.3%
B	1	1.7%
C	7	11.7%
E	31	51.7%
?	1	1.7%

Q8 (c10)

Option	n	/60
A	1	1.7%
B	38	63.3%
C	21	35.0%

Q9 (c11)

Option	n	/60
A	9	15.0%
B	26	43.3%
C	5	8.3%
D	4	6.7%

What Freqs is saying is that there was one response on Q7 which was not a letter or a digit. You can find out what it was by going back to the Data sheet, and browsing down the column with Q7 responses (column 9 in this case).

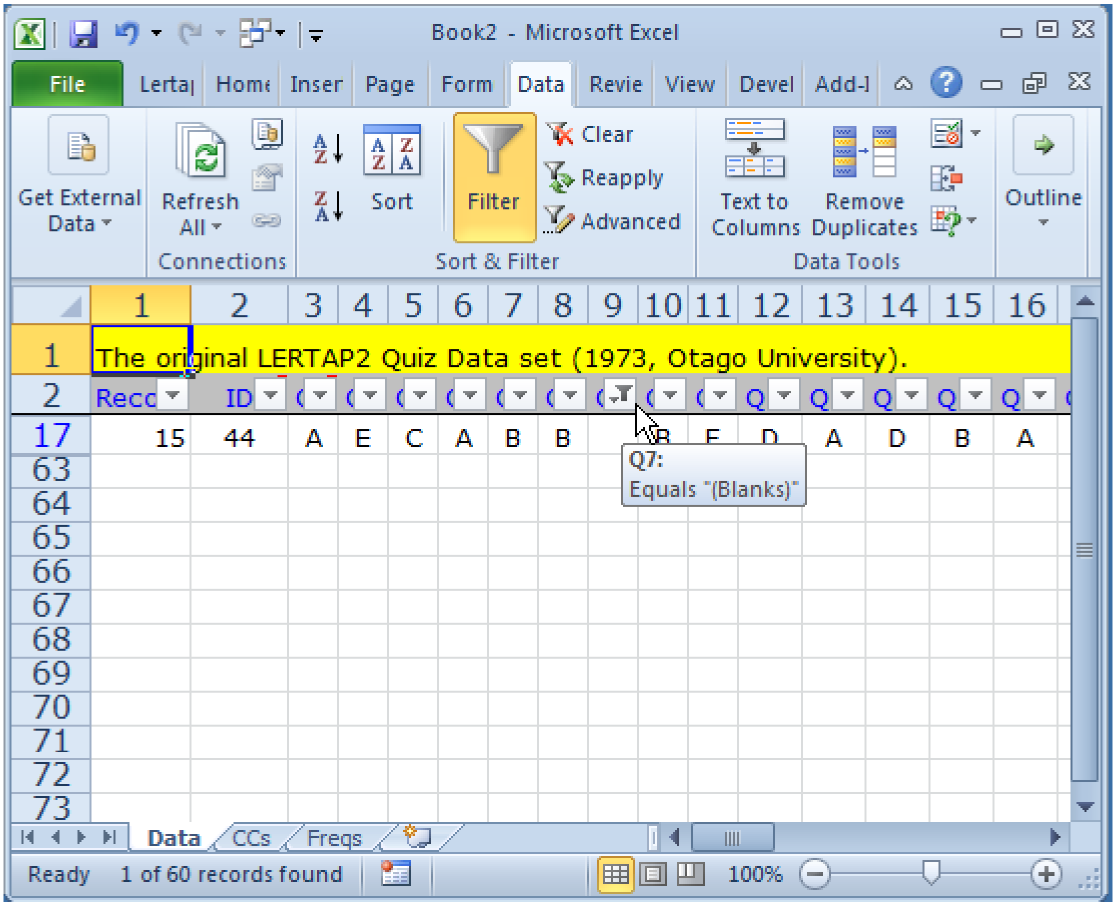
Why doesn't Freqs show a "D" for Q7? Because nobody chose that option.

While being mindful of Freqs' feelings, we can point out that it's a simple, no-nonsense summary of response frequencies. It has no pretenses; it does not claim to be sophisticated.

But it's useful, to be sure. It quickly summarizes what went on in the data set. And, very importantly, it's a sure-fire way to see if there are any weird responses in the data columns. For example, if the Q7 tally had included an X, that would be weird as only responses A though F were valid.

What if you do see weird results in Freqs, and want to quickly find the responsible records in the Data worksheet? Excel has a set of Data options, one of them is "Filter". This is a powerful little option which will let you rapidly find records which have "weird" results. (Keep in mind that Excel's on-line Help is there to assist you, should you have questions on how to use the Filter option.)

Here's a screen snapshot of the Filter option as seen when using Excel 2010:



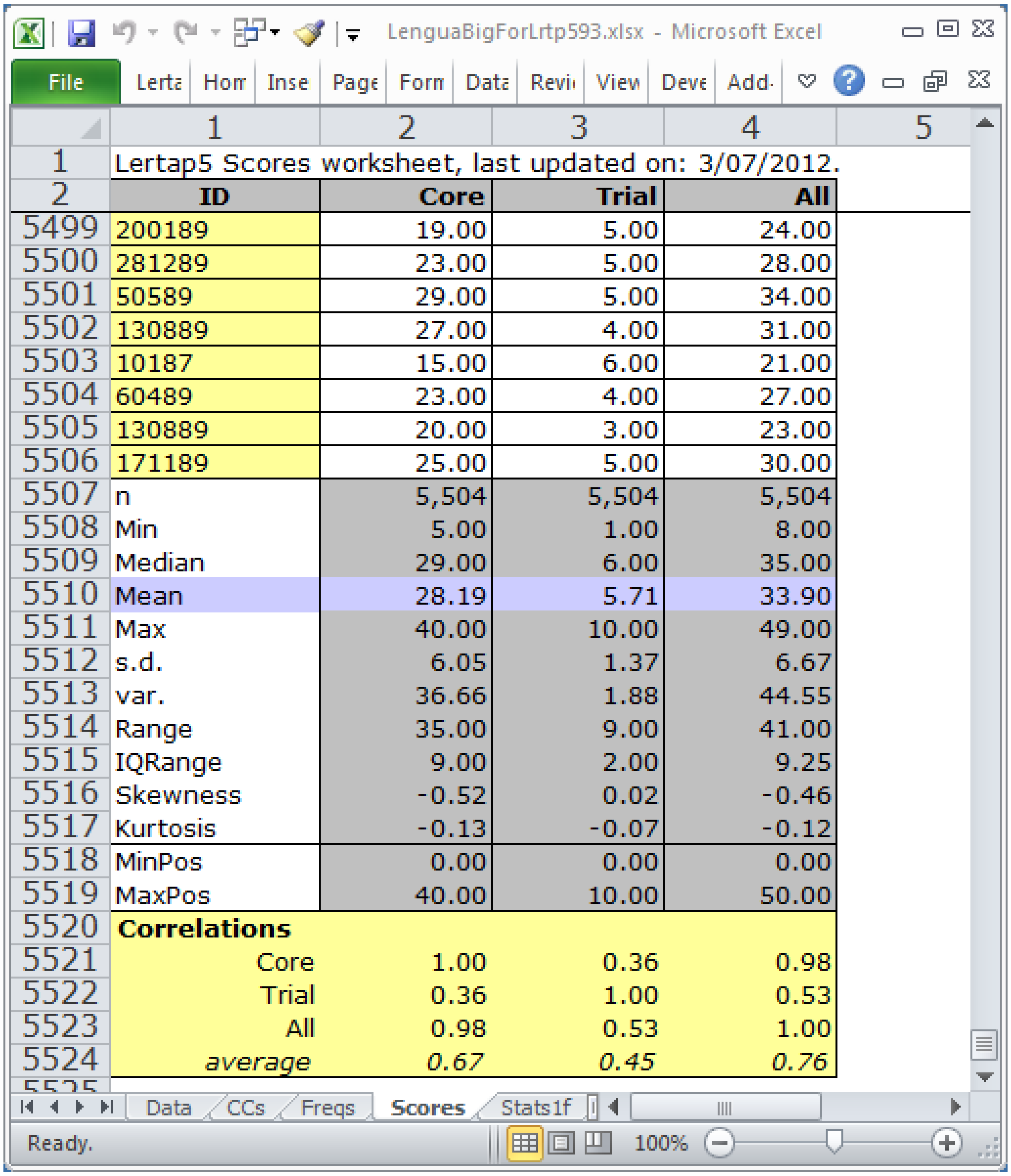
A practical example of Excel's filter in action may be seen at the bottom of this topic.

Freqs is a worksheet produced by Lertap's "Interpret" option. It's regarded as a crucial report, one which ought to be consulted before Lertap's other reports are examined -- data processing errors are not at all uncommon, and Freqs is, in Lertap, the best way to quickly spot them.

Note: for a practical example of using Freqs and looking for possible data processing errors, you might treat yourself to a read of this topic.

6.4.2 Scores

Here's an example of a Scores report. It's from the LenguaBIg dataset.



	1	2	3	4	5
1	Lertap5 Scores worksheet, last updated on: 3/07/2012.				
2	ID	Core	Trial	All	
5499	200189	19.00	5.00	24.00	
5500	281289	23.00	5.00	28.00	
5501	50589	29.00	5.00	34.00	
5502	130889	27.00	4.00	31.00	
5503	10187	15.00	6.00	21.00	
5504	60489	23.00	4.00	27.00	
5505	130889	20.00	3.00	23.00	
5506	171189	25.00	5.00	30.00	
5507	n	5,504	5,504	5,504	
5508	Min	5.00	1.00	8.00	
5509	Median	29.00	6.00	35.00	
5510	Mean	28.19	5.71	33.90	
5511	Max	40.00	10.00	49.00	
5512	s.d.	6.05	1.37	6.67	
5513	var.	36.66	1.88	44.55	
5514	Range	35.00	9.00	41.00	
5515	IQRange	9.00	2.00	9.25	
5516	Skewness	-0.52	0.02	-0.46	
5517	Kurtosis	-0.13	-0.07	-0.12	
5518	MinPos	0.00	0.00	0.00	
5519	MaxPos	40.00	10.00	50.00	
5520	**Correlations**				
5521	Core	1.00	0.36	0.98	
5522	Trial	0.36	1.00	0.53	
5523	All	0.98	0.53	1.00	
5524	*average*	*0.67*	*0.45*	*0.76*	

This Scores report has four columns and over five thousand rows. When Lertap creates Score reports, it automatically scrolls the report so that you can see the summary statistics and correlations right away -- they're always at the bottom of the report.

Have a look at the LenguaBIg dataset in order to get an idea of how the three test scores, Core, Trial, and All, were defined (see the dataset's CCs worksheet).

The statistics found at the bottom of a Scores report are explained in the table below.

n	The number of scores (corresponds to the number of students).
Min	The lowest score found.
Median	The median score. There will be 50% of the scores below the median and 50% above it. The median corresponds to the 50th percentile of the scores, and is often denoted as Q_2 in the literature.
Mean	The average of the test scores.
Max	The highest score found.
s.d.	The standard deviation of the scores. Computed as a "population" value. If you're familiar with the equations used to calculate this statistic, the one used here has "n" in the denominator (not "n-1").
var.	The variance of the scores, computed as a population value.
Range	Computed by finding the difference between the highest and lowest scores (that is, in this case, Max minus Min).
IQRange	The inter-quartile range, computed by finding the difference between the 75th and 25th percentiles of the scores. In the literature, the 75th percentile is often denoted as Q_3, while the 25th is Q_1. Thus, the IQRange is equal to Q_3 minus Q_1.
Skewness	This statistic will be zero when the distribution of scores is symmetric about their central value. In this case the Mean and Median will be the same. Negative skewness means that the scores have a tail which extends to the left of the distribution, while positive skewness indicates that the tail is to the right. Lertap uses Excel's SKEW function to compute this statistic.
Kurtosis	Kurtosis is an indicator of how peaked the scores are when compared to the normal (or Gaussian) distribution. The normal distribution has a kurtosis value of 3.00. Lertap uses Excel's KURT function to compute this statistic; KURT is actually an index of *excess kurtosis*, a comparative measure which indicates how the kurtosis of the scores compares to that of the normal curve. Negative values mean the scores are less peaked than the normal curve.

MinPos	The lowest possible score on the test. This will always be less than or equal to "Min", the lowest score found.
MaxPos	The highest possible score on the test. Always equal to or greater than "Max", the highest score found.
Correlations	These appear only when there is more than one score. Each value is the Pearson product-moment correlation between two scores. The *average* correlation is found by computing the mean of a score's correlations with the other scores. These correlation coefficients may range in value from -1.00 to +1.00. Lertap's Scatterplot option may be used to graph the relationship between any pair of scores.

6.4.3 Stats reports

Lertap makes two or three "Stats" reports for a subtest.

Every time you run Lertap's Elmillon option, Stats1f and Stats1b reports are created. If the subtest is a cognitive one, a Stats1ul report is also usually created. (It is possible to stop Lertap from creating Stats1ul reports by using a setting in the System worksheet.)

The f in Stats1f stand for "full". These reports have the most detailed information for test items. Statsf reports go back to Lertap's birth in 1973; they have changed a bit over the years, but not by very much.

Some users often find Stats1f to be excessively detailed, full of numbers and tables which make sense to experienced test developers but, well, couldn't there be something easier to read? Yes: Stats1b.

Stats1b reports are b for briefer. They are designed to be easier to understand whilst retaining the most important information from Stats1f.

Stats1f and Stats1b are similar in that they're based on the use of correlation coefficients to reflect how items have performed.

When cognitive items are involved, another time-tested way to summarize item performance is to use "upper-lower" (U-L) methods, sometimes referred to as "high-low" (H-L) methods. These methods date back to pre-computer days when educators endeavored to assess item performance by hand, without the help of any electronic aide. The idea is a simple one: to see if an item is working well, we'll look at results from the best test takers, those with the highest test scores. We will compare these results from those gathered by looking at how the weakest (lowest) students did. If the top (or upper) students got the item right, while the bottom (lower) students did not, we say the item has "discriminated"; such items are ones we'll use again in the future when we want to have a test which we know can identify the strongest and weakest students.

The Stats1ul report employs the upper-lower method to index item quality.

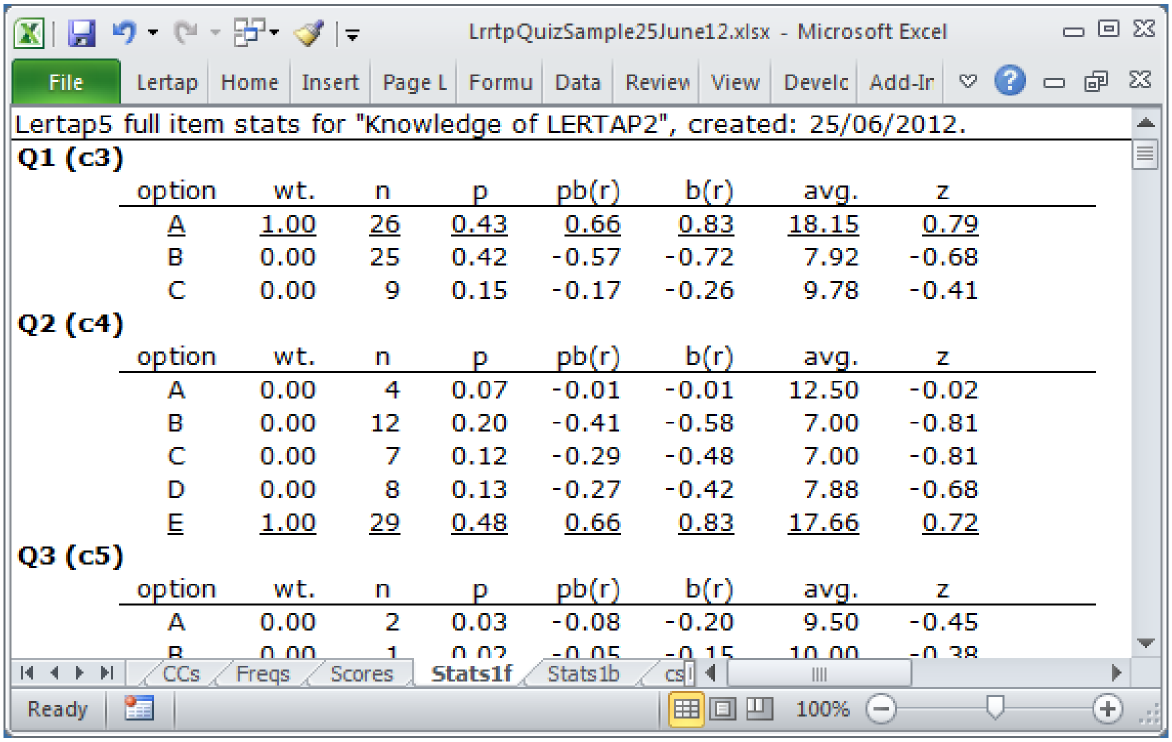
When more than one subtest is involved in a Lertap analysis, there will be more reports, two or three for each subtest. Thus we will have Stats2f and Stats2b, and maybe Stats2ul, corresponding to results for the second subtest, and, on a really lucky day with three subtests, we'll be able to bask in the light of Stats3f, Stats3b, and, well, surely you get the idea. The number in the report name refers to subtest number.

You'll now be keen to page ahead, getting more detailed information for each of Lertap's Stats reports. But before you do, note: there's a top-flight manual about this stuff, a real printed book whose battery lasts forever and ever, something which can even be read in the brightest of sunshine, withstanding coffee and tea spills without crashing. You can even drop it, or run over it with your cycle or the baby's perambulator. Old technology, but rock solid. And it doesn't assume you're full bottle on statistics -- it holds your hand with loving humor, gently nudging you along ever so gently and respectfully. Read more about the manual here, and do try and read it; it covers Lertap's various statistics reports in more detail, and has several practical examples.

6.4.3.1 Stats1f

Statsf reports have four sections.

By far the longest section is that which has the detailed item statistics. Here you will find a table of results for each test (or subtest) item, as exemplified below:



Lertap5 full item stats for "Knowledge of LERTAP2", created: 25/06/2012.

Q1 (c3)

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	1.00	26	0.43	0.66	0.83	18.15	0.79
B	0.00	25	0.42	-0.57	-0.72	7.92	-0.68
C	0.00	9	0.15	-0.17	-0.26	9.78	-0.41

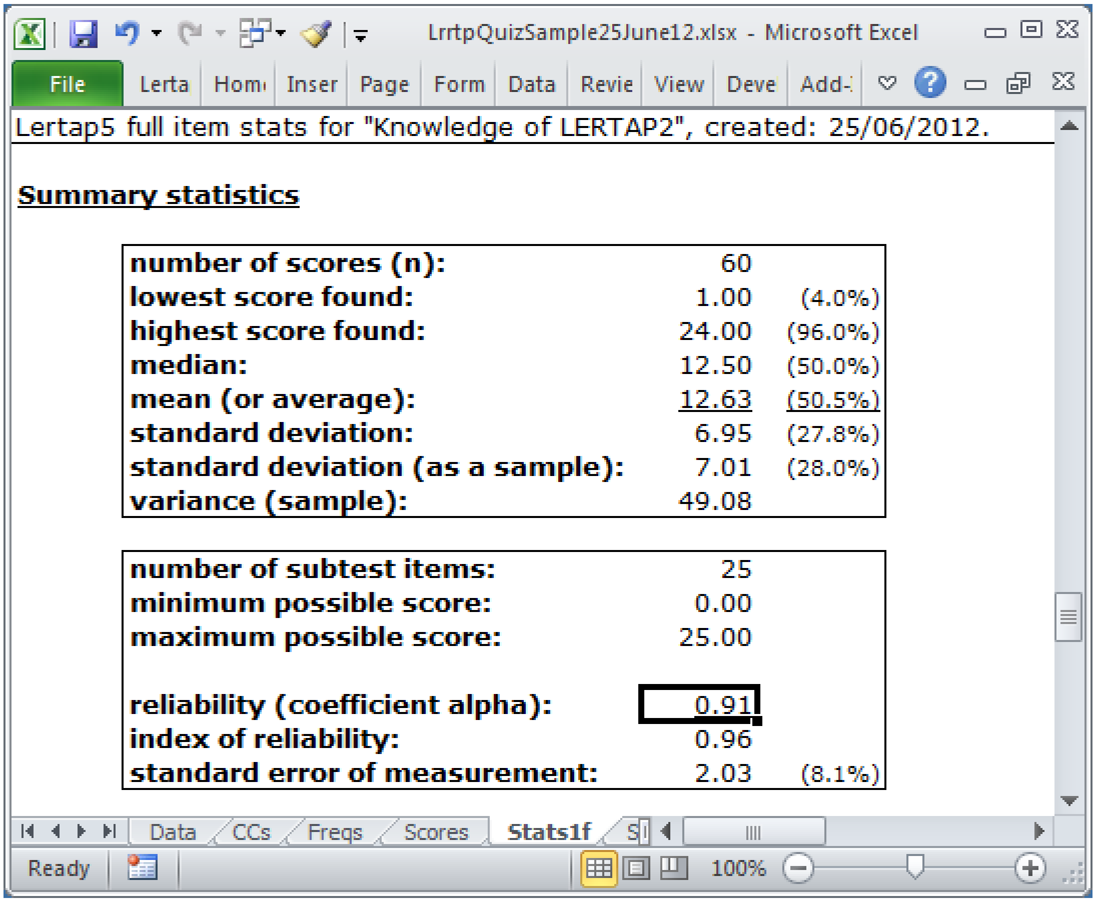
Q2 (c4)

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	4	0.07	-0.01	-0.01	12.50	-0.02
B	0.00	12	0.20	-0.41	-0.58	7.00	-0.81
C	0.00	7	0.12	-0.29	-0.48	7.00	-0.81
D	0.00	8	0.13	-0.27	-0.42	7.88	-0.68
E	1.00	29	0.48	0.66	0.83	17.66	0.72

Q3 (c5)

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	2	0.03	-0.08	-0.20	9.50	-0.45
B	0.00	1	0.02	-0.05	-0.15	10.00	-0.38

Click here to bone up on this section of a Statsf report (the one showing above).



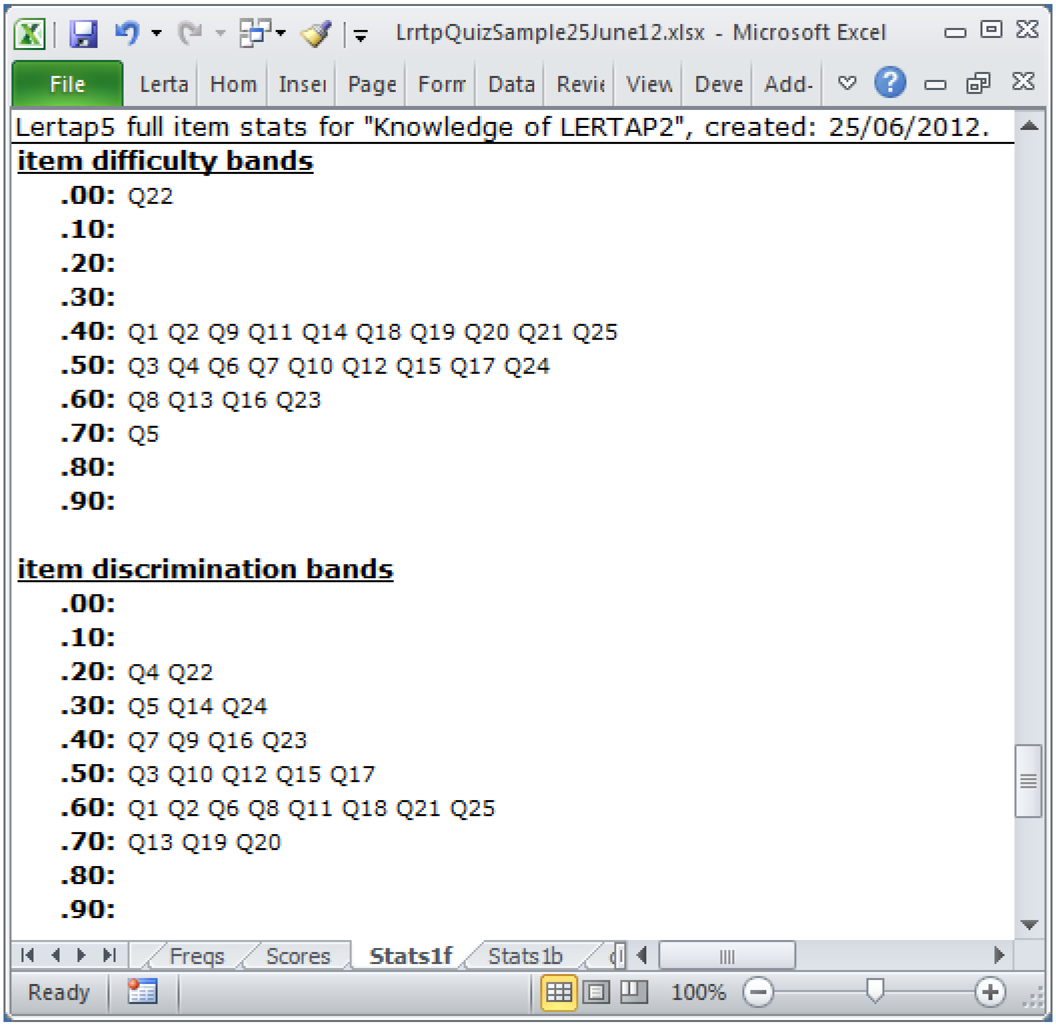
Lertap5 full item stats for "Knowledge of LERTAP2", created: 25/06/2012.

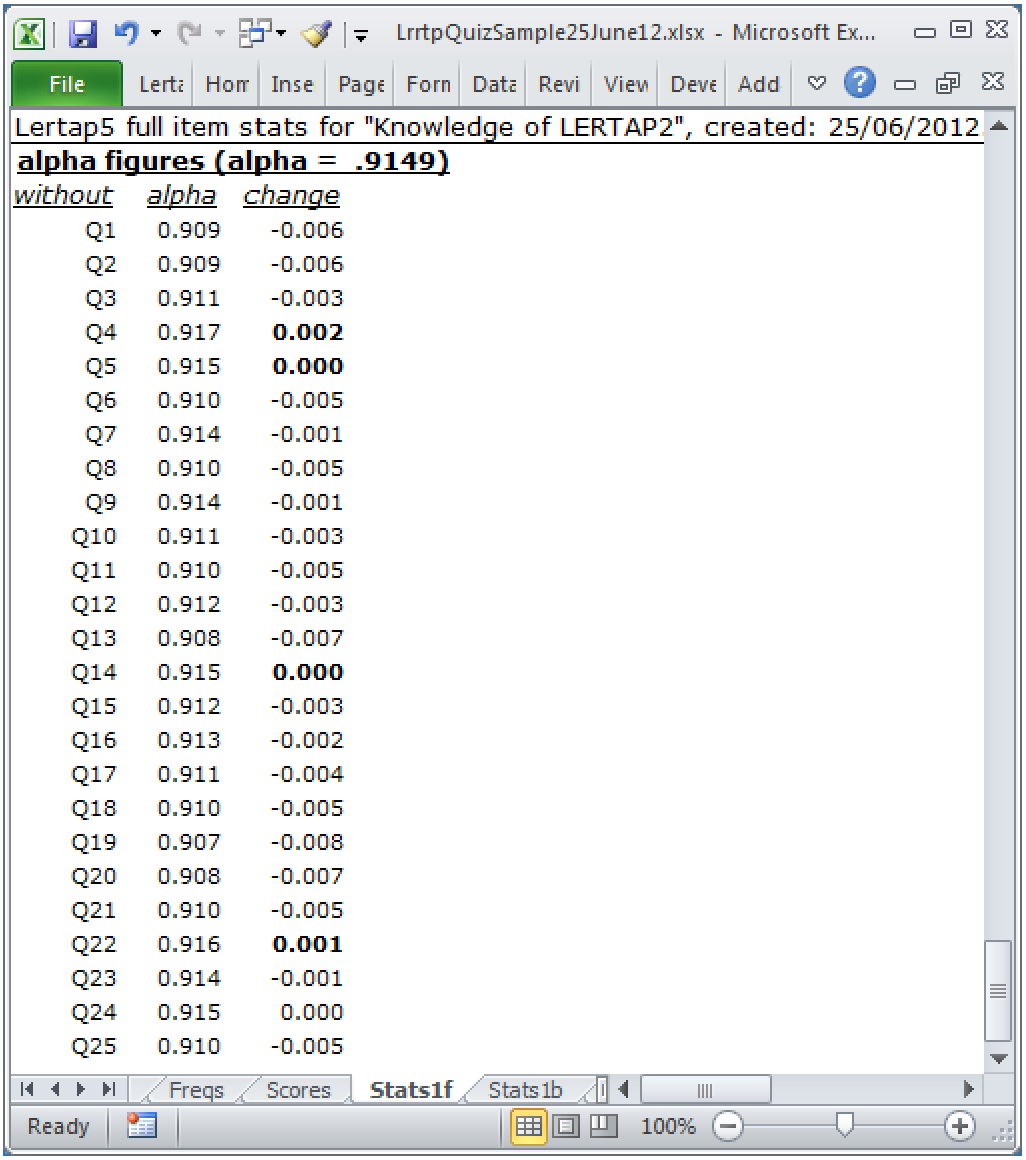
Summary statistics

number of scores (n):	60	
lowest score found:	1.00	(4.0%)
highest score found:	24.00	(96.0%)
median:	12.50	(50.0%)
mean (or average):	12.63	(50.5%)
standard deviation:	6.95	(27.8%)
standard deviation (as a sample):	7.01	(28.0%)
variance (sample):	49.08	

number of subtest items:	25	
minimum possible score:	0.00	
maximum possible score:	25.00	
reliability (coefficient alpha):	0.91	
index of reliability:	0.96	
standard error of measurement:	2.03	(8.1%)

Click here to read about the "Summary statistics" section of a Statsf report, as shown above.


LrrtpQuizSample25June12.xlsx - Microsoft Excel
File
Lerta
Hom
Inse
Page
Form
Data
Revie
View
Deve
Add-
Lertap5 full item stats for "Knowledge of LERTAP2", created: 25/06/2012.
item difficulty bands
.00: Q22
.10:
.20:
.30:
.40: Q1 Q2 Q9 Q11 Q14 Q18 Q19 Q20 Q21 Q25
.50: Q3 Q4 Q6 Q7 Q10 Q12 Q15 Q17 Q24
.60: Q8 Q13 Q16 Q23
.70: Q5
.80:
.90:
item discrimination bands
.00:
.10:
.20: Q4 Q22
.30: Q5 Q14 Q24
.40: Q7 Q9 Q16 Q23
.50: Q3 Q10 Q12 Q15 Q17
.60: Q1 Q2 Q6 Q8 Q11 Q18 Q21 Q25
.70: Q13 Q19 Q20
.80:
.90:
Freqs
Scores
Stats1f
Stats1b
Ready
100%



Lertap5 full item stats for "Knowledge of LERTAP2", created: 25/06/2012

alpha figures (alpha = .9149)

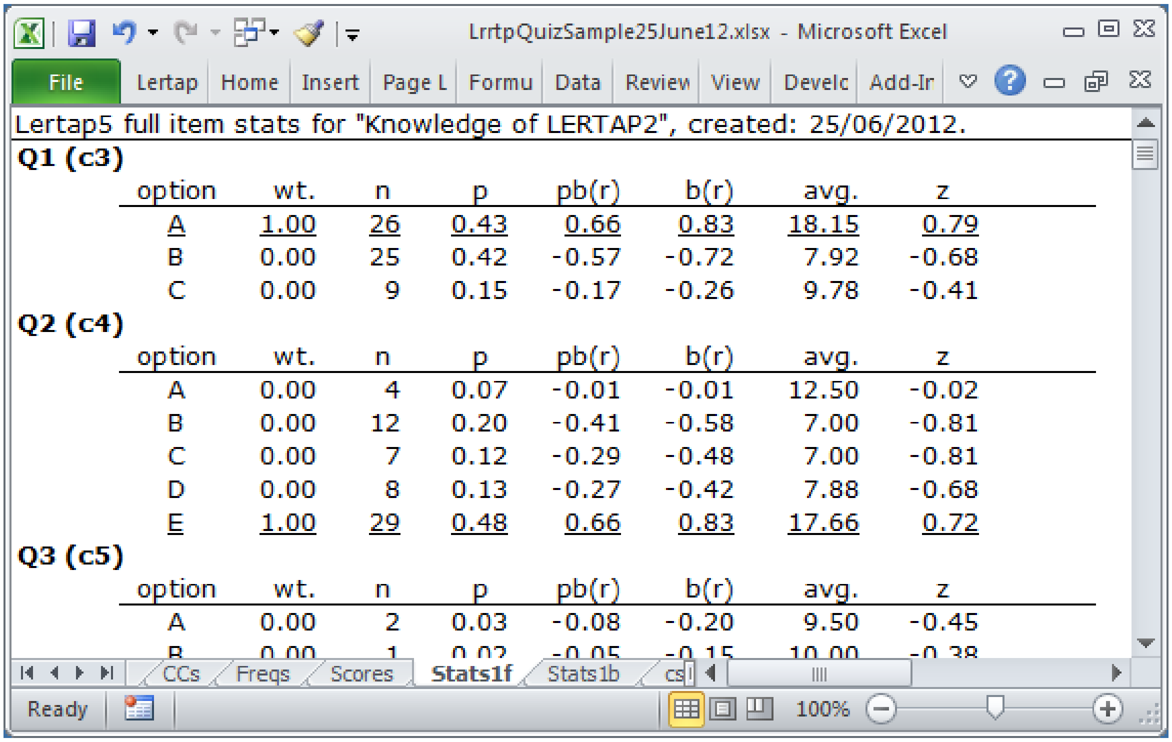
without	*alpha*	*change*
Q1	0.909	-0.006
Q2	0.909	-0.006
Q3	0.911	-0.003
Q4	0.917	**0.002**
Q5	0.915	**0.000**
Q6	0.910	-0.005
Q7	0.914	-0.001
Q8	0.910	-0.005
Q9	0.914	-0.001
Q10	0.911	-0.003
Q11	0.910	-0.005
Q12	0.912	-0.003
Q13	0.908	-0.007
Q14	0.915	**0.000**
Q15	0.912	-0.003
Q16	0.913	-0.002
Q17	0.911	-0.004
Q18	0.910	-0.005
Q19	0.907	-0.008
Q20	0.908	-0.007
Q21	0.910	-0.005
Q22	0.916	**0.001**
Q23	0.914	-0.001
Q24	0.915	0.000
Q25	0.910	-0.005

Freqs | Scores | Stats1f | Stats1b

Click here to read about the Stats1f output seen immediately above.

6.4.3.1.1 Full item statistics

By far the longest section in any Statsf report is that which has the detailed item statistics. Here you will find a table of results for each test (or subtest) item.



Lertap5 full item stats for "Knowledge of LERTAP2", created: 25/06/2012.

Q1 (c3)

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	1.00	26	0.43	0.66	0.83	18.15	0.79
B	0.00	25	0.42	-0.57	-0.72	7.92	-0.68
C	0.00	9	0.15	-0.17	-0.26	9.78	-0.41

Q2 (c4)

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	4	0.07	-0.01	-0.01	12.50	-0.02
B	0.00	12	0.20	-0.41	-0.58	7.00	-0.81
C	0.00	7	0.12	-0.29	-0.48	7.00	-0.81
D	0.00	8	0.13	-0.27	-0.42	7.88	-0.68
E	1.00	29	0.48	0.66	0.83	17.66	0.72

Q3 (c5)

option	wt.	n	p	pb(r)	b(r)	avg.	z
A	0.00	2	0.03	-0.08	-0.20	9.50	-0.45
B	0.00	1	0.02	-0.05	-0.15	10.00	-0.38

The screen snapshot above shows results for items from the 25-item cognitive subtest found in the "Lertap Quiz".

The items in this subtest used upper-case letters as response options. Sixty (60) students took the test.

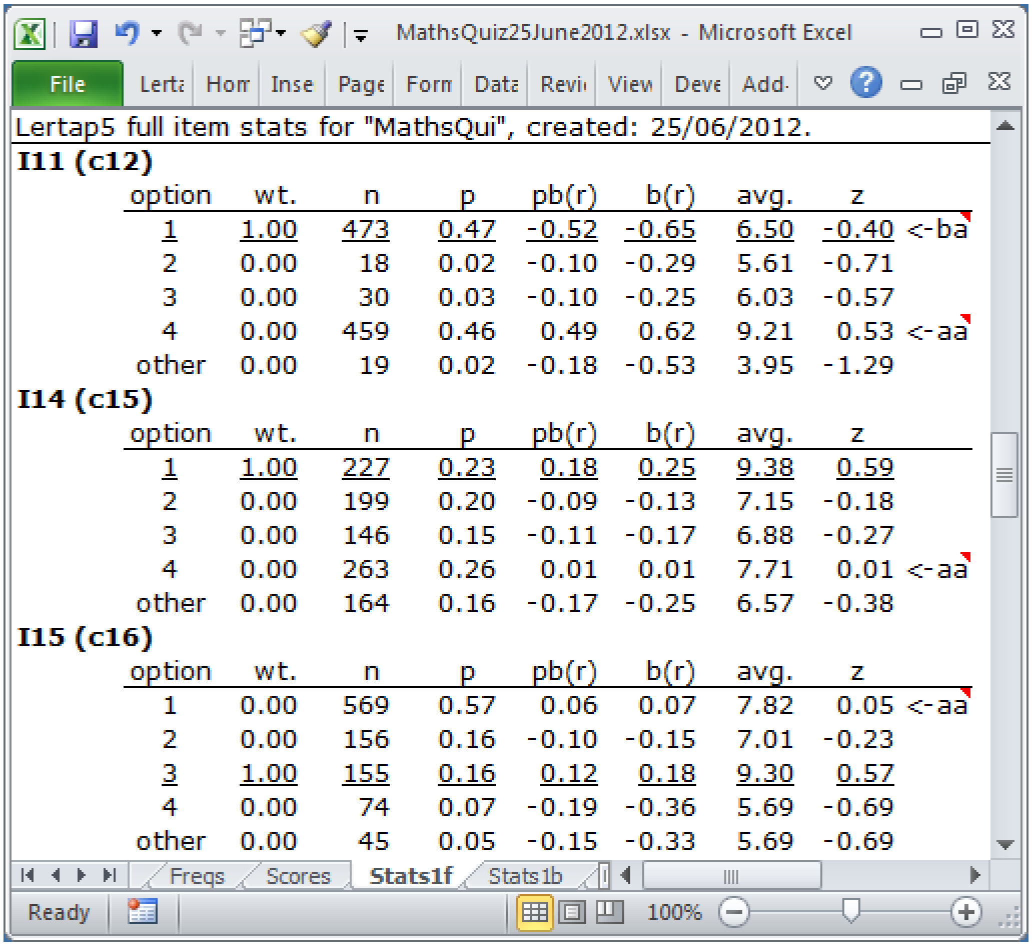
The table below provides an interpretation of the column headings seen in the report.

option	The response option selected by the student. Response options may be letters (upper or lower case) or digits (from 0 to 9). If the option corresponds to the keyed-correct answer, then the option is underlined, as are all the statistics for the option.
wt.	The number of scoring points given for selecting the option.
n	The number of people who selected the option.
p	The proportion of people who selected the option. Multiply this by 100 to get the corresponding percent. In the example above, 43% of the students selected option A on

	Q1. If the option is the keyed-correct answer, then p is usually referred to as "item difficulty". Thus, the difficulty of Q1 is 0.43. In the Stats1b report, p is "diff". Note that an item with a difficulty of 0.80 would be referred to as an "easy" item as 80% of the students got it right. Conversely, a difficulty of 0.20 would be a "hard" item as only 20% got it right.
pb(r)	The point-biserial correlation with the criterion score. To compute this, all those who selected the option are given a "score" of 1 while those who did not are given a score of 0. These scores are then correlated with the criterion score. If the option is the keyed-correct answer, then pb(r) is usually referred to as "item discrimination". In the Stats1b report, pb(r) for the keyed-correct answer is denoted as "disc". The pb(r) values will range from -1.00 to +1.00. Items with a discrimination of at least 0.30 are often referred to as having good discrimination. Above, both Q1 and Q2 would be said to have good discrimination as their pb(r) values for the correct option are above 0.30. (Lertap corrects pb(r) values for "part-whole inflation" whenever the item is part of the criterion measure.)
b(r)	The biserial correlation with the criterion score. Sometimes the option selection "scores" of 0 and 1 used to compute pb(r) are regarded as a dichotomous code for an underlying, or "latent", variable having a range of scores. If we assume the latent variable to have a "normal" (or Gaussian) distribution, then a well-known equation may be applied to pb(r) to derive b(r), which is then taken as an estimate of the correlation between the latent variable and the criterion score. Note that it is possible for b(r) to be greater than 1.00.
avg.	The average criterion score for those who selected the option. On Q2 (above), the 29 students who selected option E, the keyed-correct option, had an average criterion score of 17.66. In this case, 17.66 is the mean test score for these 29 students -- the "criterion" is the test score. The test had 25 items, and all items were scored on a right-wrong basis, with 1 point for right, 0 otherwise. The maximum possible criterion score is thus 25. When the "criterion" is the test score itself, it's called an "internal criterion". At times an "external" criterion measure is on hand, and item performance may be assessed by correlating item results with the external criterion score. Please refer to this topic for more information.

z	avg. expressed as a "z-score". Sixty (60) students took our 25-item test. Their average test score, that is, the mean of the 60 test scores, was 12.63. The standard deviation of the 60 scores was 6.95. With these two figures on hand, the "z score" for the 29 students who selected option E on Q1 is (17.66 minus 12.63) divided by 6.95, or 0.72. z-scores are "standard scores". No matter what the test mean and standard deviation may be, when z-scores are formed, the mean of the z-scores will always be zero (0.00), and the standard deviation will always be one (1.00). A positive z-score occurs when the original score is greater than the overall average criterion score. z-scores are frequently interpreted by making reference to the "normal curve". A z-score of 1.00 in a normal distribution is at the 84th percentile; only 16% of all test takers will have a z-score higher than 1.00. A z-score of -1.00 in a normal distribution is at the 16th percentile; only 16% of all test takers will have a z-score less than -1.00. (More about z here.)

Consider now another test, a 15-item mathematics quiz given to 999 high school students.



Lertap5 full item stats for "MathsQui", created: 25/06/2012.

I11 (c12)

option	wt.	n	p	pb(r)	b(r)	avg.	z	
1	1.00	473	0.47	-0.52	-0.65	6.50	-0.40	<-ba
2	0.00	18	0.02	-0.10	-0.29	5.61	-0.71	
3	0.00	30	0.03	-0.10	-0.25	6.03	-0.57	
4	0.00	459	0.46	0.49	0.62	9.21	0.53	<-aa
other	0.00	19	0.02	-0.18	-0.53	3.95	-1.29	

I14 (c15)

option	wt.	n	p	pb(r)	b(r)	avg.	z	
1	1.00	227	0.23	0.18	0.25	9.38	0.59	
2	0.00	199	0.20	-0.09	-0.13	7.15	-0.18	
3	0.00	146	0.15	-0.11	-0.17	6.88	-0.27	
4	0.00	263	0.26	0.01	0.01	7.71	0.01	<-aa
other	0.00	164	0.16	-0.17	-0.25	6.57	-0.38	

I15 (c16)

option	wt.	n	p	pb(r)	b(r)	avg.	z	
1	0.00	569	0.57	0.06	0.07	7.82	0.05	<-aa
2	0.00	156	0.16	-0.10	-0.15	7.01	-0.23	
3	1.00	155	0.16	0.12	0.18	9.30	0.57	
4	0.00	74	0.07	-0.19	-0.36	5.69	-0.69	
other	0.00	45	0.05	-0.15	-0.33	5.69	-0.69	

Three items from the maths quiz are on display in the screen snapshot above. They have been selected in order to show how Lertap "flags" item options which may have problems.

There were 15 items on the quiz, all scored on a right-wrong basis, with one point given for each correct answer. The overall mean test score was 7.67 with a standard deviation of 2.88.

Now, if an item is meant to be one which helps us identify the strong students, we expect that those who get the item right will be the best students. Those who get it wrong should be the weakest students

This did not happen on I11. The keyed-correct response was option 1, selected by 473 students. Their avg. test score was 6.50, corresponding to a *negative* z-score of 0.40. The last option, 4, was selected by 459 students whose avg. test score was 9.21, with a *positive* z-score of 0.53. Weak students are getting I11 right, while strong students are getting it wrong.

To draw our attention to this, Lertap has "flagged" I11 twice. We've got a "ba" flag for the first option, and an "aa" flag for the last option.

The "ba" flag appears whenever an option with a wt. above 0.00 has a negative z value. "ba" means "below average". This flag will wave whenever the students who select the keyed-correct option have a below-average test score (we expect such students to have an above-average test score).

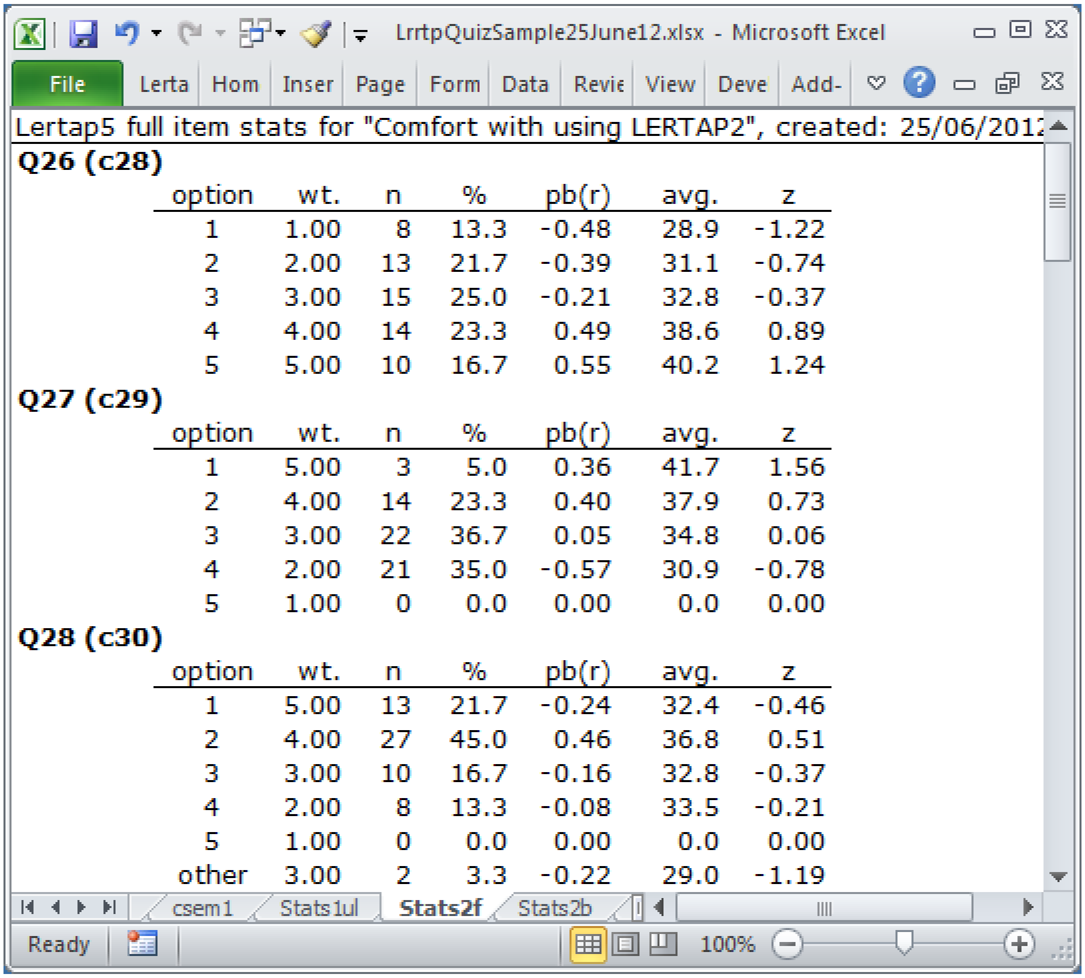
The "aa" flag will show whenever an option with wt. at or below 0.00 has a positive z value. Here, "aa" stands for "above-average" -- it'll come out to flap in the breeze whenever a "distractor" (an incorrect option) is selected by above-average students. This is an unwanted outcome as we expect the above-average students to get the item right.

What has happened is that I11 has been mis-keyed. The correct option on I11 was 4, not 1. An error was made when the *key line was typed into the CCs worksheet. Once this error is corrected we can expect Stats1f to display no flags at all for I11.

Why has Lertap flagged I14 and I15? Because both of these items have a distractor selected by students with above-average test scores. The flagged distractors for these two items have z values just above zero, meaning that the students who selected them were slightly above average. This is unexpected. It probably indicates that the items need to be reviewed -- what is there about these distractors which has made good students see them as plausible? (Asking the students is likely to uncover the underlying reasons.)

There is one more flag which may appear in a Stats1f report. It's "<-no". It will appear whenever an item option has not been selected by anyone.

It is possible to stop Lertap from using these flags in its Stats1f reports. There is a setting in Row 58 of Lertap's System worksheet which controls it.



Lertap5 full item stats for "Comfort with using LERTAP2", created: 25/06/2012

Q26 (c28)

option	wt.	n	%	pb(r)	avg.	z
1	1.00	8	13.3	-0.48	28.9	-1.22
2	2.00	13	21.7	-0.39	31.1	-0.74
3	3.00	15	25.0	-0.21	32.8	-0.37
4	4.00	14	23.3	0.49	38.6	0.89
5	5.00	10	16.7	0.55	40.2	1.24

Q27 (c29)

option	wt.	n	%	pb(r)	avg.	z
1	5.00	3	5.0	0.36	41.7	1.56
2	4.00	14	23.3	0.40	37.9	0.73
3	3.00	22	36.7	0.05	34.8	0.06
4	2.00	21	35.0	-0.57	30.9	-0.78
5	1.00	0	0.0	0.00	0.0	0.00

Q28 (c30)

option	wt.	n	%	pb(r)	avg.	z
1	5.00	13	21.7	-0.24	32.4	-0.46
2	4.00	27	45.0	0.46	36.8	0.51
3	3.00	10	16.7	-0.16	32.8	-0.37
4	2.00	8	13.3	-0.08	33.5	-0.21
5	1.00	0	0.0	0.00	0.0	0.00
other	3.00	2	3.3	-0.22	29.0	-1.19

The Stats2f report seen above shows how the detailed item statistics look for an affective subtest, in this case the 10-item survey included in the Lertap Quiz dataset.

In contrast to cognitive items, where most often only one of an item's options will be scored (that is, have a wt. which is above 0.00), each option on an affective item is usually scored. The three items pictured above, Q26, Q27, and Q28 all use five options: {1,2,3,4,5}. A student who selects the first option on Q26 will get one point. Selecting the first option on Q27 and Q28, on the other hand, will "award" the student with five points.

These items are example of polytomous scoring. To read more about how Lertap deals with items of this sort, please branch to this topic.

What happens when students don't answer questions, or are not given the opportunity to answer all questions? How does Lertap go about scoring in these cases? See the missing data topics to find out.

Finally: see the manual! There is much more about item statistics in the manual, and it's not a difficult read at all.

Related tidbits:

Flags are also waved in Stats1b reports, where they appear in the ? column. Read more.

The use of flags in Stats1f reports is controlled by an "Annotate Stats-f reports" setting in the System worksheet (it is possible to turn them off).

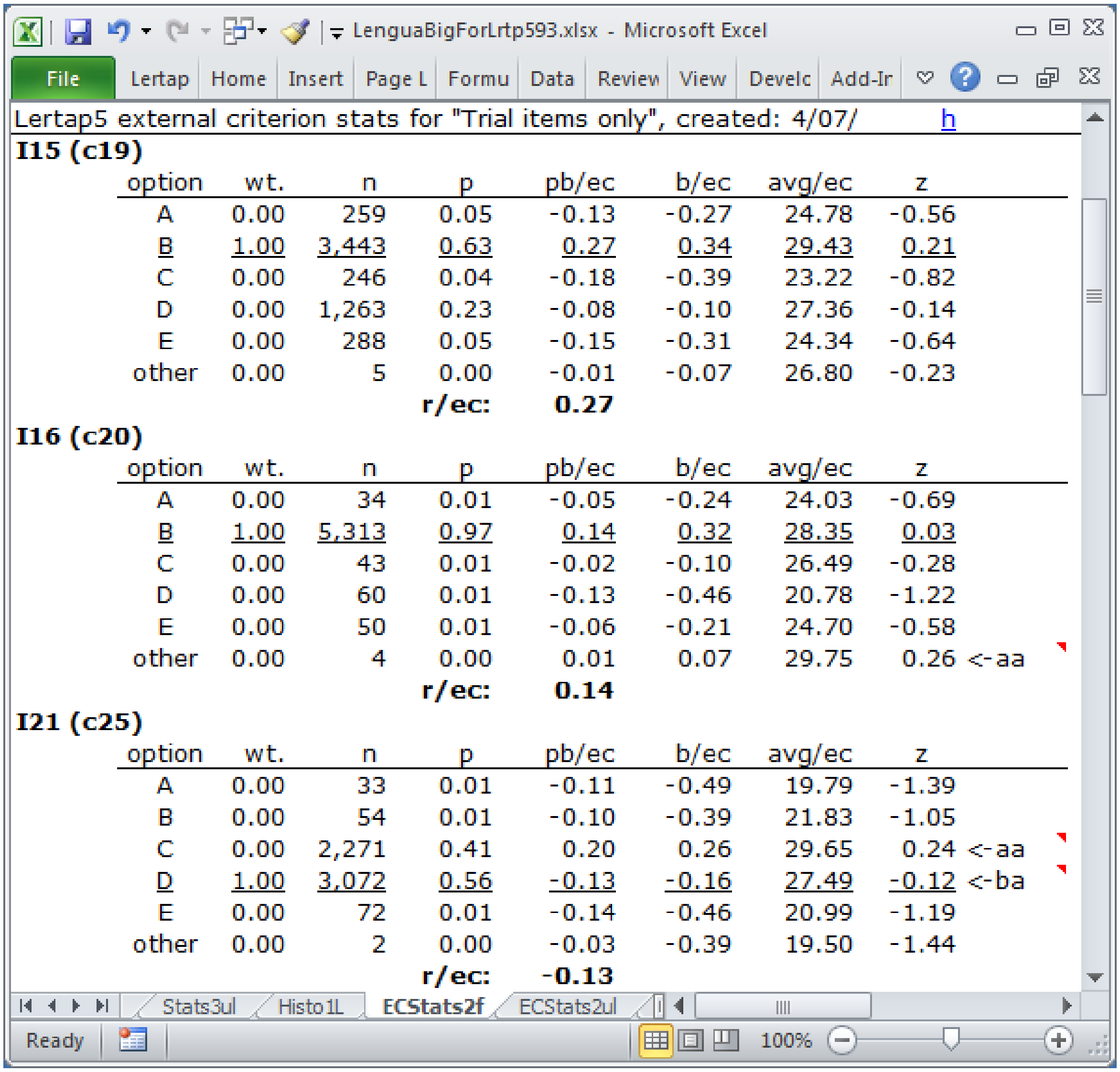
Chapter 7 of the manual has a more in-depth discussion, with references to relevant literature. It may be the most popular chapter in the manual, sometimes used in test and measurement classes. Get a pdf copy of this chapter with a click here.

An easy-to-read technical paper with more details about how Lertap 5 calculates item correlations is available here.

How to print Lertap's reports? Not hard at all, especially if you take in this topic.

6.4.3.1.1.1 With external criterion

Statsf reports become "ECStatsf" reports when an external criterion analysis has been selected. The discussion found in this topic is based on results from the LenguaBIg dataset. The "Core" score was used as the external criterion, and the second subtest, "Trial items only", was selected for the analysis.



Lertap5 external criterion stats for "Trial items only", created: 4/07/ h

I15 (c19)

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**				

I16 (c20)

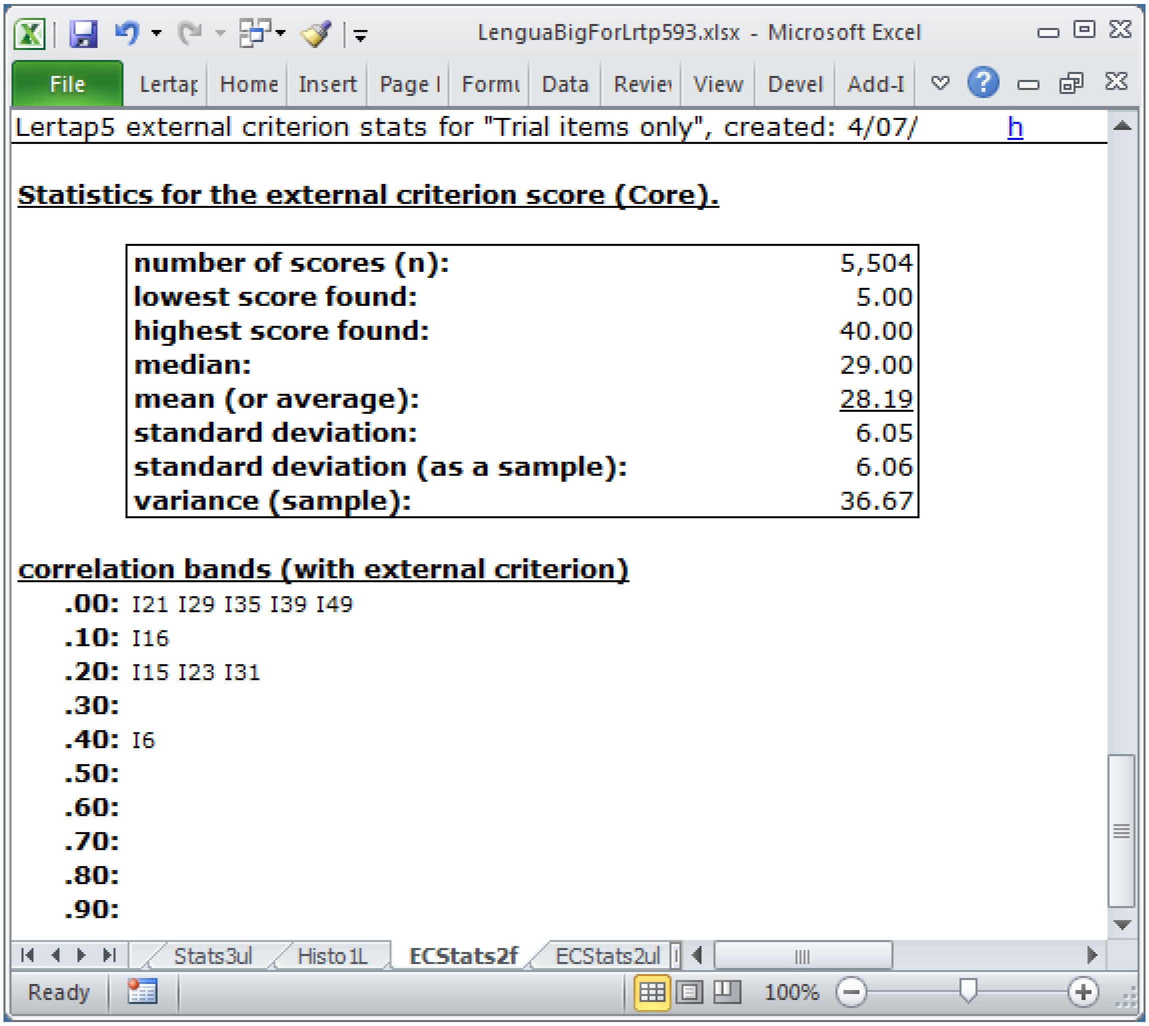
option	wt.	n	p	pb/ec	b/ec	avg/ec	z	
A	0.00	34	0.01	-0.05	-0.24	24.03	-0.69	
B	1.00	5,313	0.97	0.14	0.32	28.35	0.03	
C	0.00	43	0.01	-0.02	-0.10	26.49	-0.28	
D	0.00	60	0.01	-0.13	-0.46	20.78	-1.22	
E	0.00	50	0.01	-0.06	-0.21	24.70	-0.58	
other	0.00	4	0.00	0.01	0.07	29.75	0.26	<-aa
			r/ec:	**0.14**				

I21 (c25)

option	wt.	n	p	pb/ec	b/ec	avg/ec	z	
A	0.00	33	0.01	-0.11	-0.49	19.79	-1.39	
B	0.00	54	0.01	-0.10	-0.39	21.83	-1.05	
C	0.00	2,271	0.41	0.20	0.26	29.65	0.24	<-aa
D	1.00	3,072	0.56	-0.13	-0.16	27.49	-0.12	<-ba
E	0.00	72	0.01	-0.14	-0.46	20.99	-1.19	
other	0.00	2	0.00	-0.03	-0.39	19.50	-1.44	
			r/ec:	**-0.13**				

option	The response option selected by the student. Response options may be letters (upper or lower case) or digits (from 0 to 9). If the option corresponds to the keyed-correct answer, then the option is underlined, as are all the statistics for the option.
wt.	The number of scoring points given for selecting the option.
n	The number of students who selected the option.
p	The proportion of students who selected the option. Multiply this by 100 to get the corresponding percent. In

	the example above, 56% of the students selected option D on I21.
pb/ec	The point-biserial correlation with the external criterion score. To compute this, all those who selected the option are given a "score" of 1 while those who did not are given a score of 0. These scores are then correlated with the external criterion score.
b/ec	The biserial correlation with the external criterion score. Note that it is possible for b/ec to be greater than 1.00.
avg/ec	The average external criterion score for those who selected the option. On I21 (above), the 3,072 students who selected option D, the keyed-correct option, had an average external criterion score of 27.49. The "ba" flag is showing because this avg/ec was below the mean of the external criterion scores, a fact which is also flagged by the negative "z" of -0.12. The "aa" flag on option C, a distractor, indicates that the 2,271 students who selected C, with an average external criterion score of 29.65, were above average students, at least as measured by their external criterion scores. These are generally unwanted outcomes; the flags are there to draw attention to a potential problem with I21.
z	avg/ec expressed as a "z score".
r/ec	The Pearson product-moment correlation between the item and the external criterion score. When an item has only one option with a non-zero "wt." value, r/ec will be the same as pb/ec.



The "Statistics" table for an ECStatsf report is similar to the "Summary statistics" found in an ordinary Statsf report, as seen in the next topic.

The "correlation bands" plot the r/ec values found for the items. In this example, only one item, I6, had a reasonable correlation with the external criterion (generally we'd want these correlations to be at least 0.30). As seen in the upper part of this topic, I21's r/ec was -0.13; items with negative r/ec values are found in the ".00:" correlation band.

6.4.3.1.1.2 Difficulty calculations

The difficulty of a cognitive item is traditionally defined as the proportion of people who answered the item correctly. If, for example, 80% of test takers identified the correct option to Item 1, we'd say Item 1's difficulty was 0.80.

But what if there is more than one right answer to Item 1? What then? What do we do when the scoring of a cognitive item is no longer dichotomous, right/wrong, but instead exhibits polytomous scoring? We might consider a different way of expressing item difficulty under such conditions, selecting one of the following Lertap methods.

---

1) proportional

Under this method, item difficulty is the sum of the people who selected one of the correct answers, divided by the total number of people responding. This method counts any response as being correct if its corresponding weight is greater than zero. This method does not take into account any differences which may exist among response weights.

2) item mean

A second way of assessing the difficulty of a cognitive item is to simply use the item's average, its mean. If an item has just one correct answer, and if the weight for that answer is 1.00, then the item's mean will be identical to the proportional index of difficulty.

3) item mean / max. weight (default)

Item means can be greater than zero. Traditionally, item difficulty has been measured on a scale which goes from 0.00 to 1.00; if we divide the item mean by the greatest response weight, we effectively re-scale the mean so that it falls back to the 0.00 to 1.00 range. This method of indexing item difficulty does exactly that. When there's only one correct answer to an item, it yields the same result as 1) above.

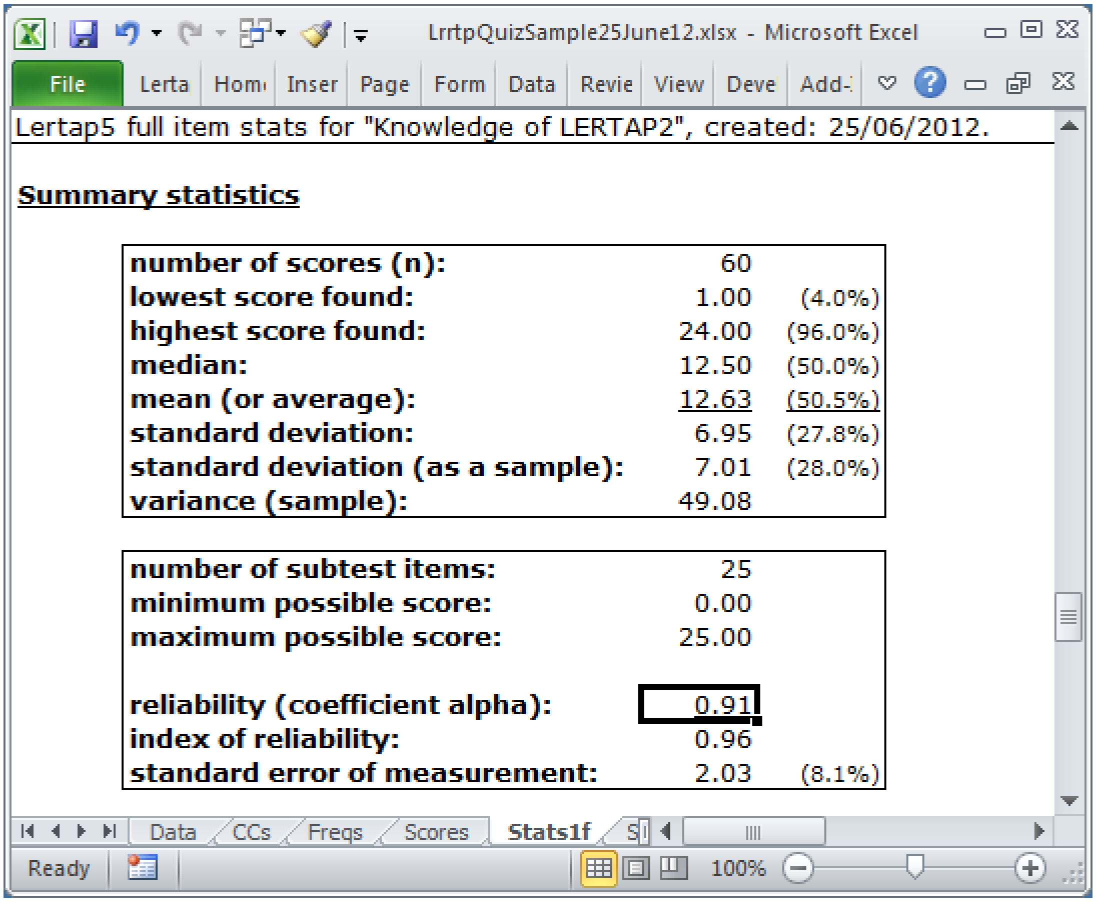
As indicated, Lertap's default method is 3), item mean divided by the maximum response weight. To change it to one of the other methods, do this: (1) make a change in Row 19 of the System worksheet in the Lertap5.xlsm file; (2) save and close the Lertap5.xlsm file.

Finally, we should mention where cognitive item difficulties are displayed. They're shown in the item difficulty bands found towards the bottom of the Stats1f report, and they have their very own column in the Stats1b report.

When the item difficulty calculation method has been set to 2) above, Lertap's item difficulty bands can come under stress since they use a 0.00 to 1.00 scale. In this case, Lertap momentarily pops into the 3) method, re-scaling the mean so that it will fall into one of the bands. However, the item mean will display correctly in the Stats1b report.

6.4.3.1.2 Summary statistics

The "Summary statistics" section of a Statsf report consists of two small tables.



number of scores (n)	This figure corresponds to the number of students whose item responses are found in the Data worksheet. Lertap creates a test score for each student when the Elmillon option is used. The actual test scores are found in the Scores worksheet.
lowest score found	Of all the scores, this figure is the lowest one found. The figure in parentheses, (4%) in this example, expresses the lowest score as a percentage of the maximum possible score. In this example, the maximum possible score was 25.
highest score found	The highest score found. The figure in parentheses, (96%) in this example, expresses this score as a percentage of the maximum possible score.

median	The score which corresponds to the 50th percentile. In this example, half of the scores are below 12.50 and half are above.
mean (or average)	The average of the test scores, also called the "mean" and the "arithmetic mean". The mean and the median will be equal when the scores are symmetric about the mean.
standard deviation	This is one measure of how spread out the scores are. If all the scores are the same, the standard deviation will be zero. If you are familiar with the equations used to calculate this statistic, Lertap uses the equation with "n" in the denominator to derive this figure.
standard deviation (as a sample)	If the students tested are considered to be a sample drawn from a larger population, an unbiased estimate of the population standard deviation is given by this figure. If you are familiar with the equations used to calculate the standard deviation, Lertap uses the equation with "n-1" in the denominator to derive this value.
variance (sample)	The square of the standard deviation (sample). Another measure of how spread out the scores are.
number of subtest items	This is the number of items in the test (or subtest).
minimum possible score	Corresponds to the lowest possible score which a student could get on the test. For a cognitive test, this is usually zero, meaning no items correct.
maximum possible score	The highest possible test score. On a cognitive test this is usually equal to the number of test items, with a student getting one point for each item answered correctly.
reliability (coefficient alpha)	Derived by applying Cronbach's method to find coefficient alpha. The maximum possible value is 1.00. A "good" cognitive test might be expected to have a reliability of at least 0.80; a "good" affective scale might be expected to have a reliability of at least 0.70. Note (however), that these "good" values are not always appropriate. A "mastery" test, for example, will sometimes be acceptable even when its reliability, as measured by alpha, is low. Affective scales frequently have alpha reliabilities below 0.70 -- when this happens, results are often interpreted on an item by item basis instead of the overall scale score.

	Notes: (1) coefficient alpha may take on negative values; when this happens Lertap reports 0.00 as the reliability figure. (2) KR-20, another reliability estimate, is the same as alpha when cognitive test items are all dichotomous -- see "tidbits" below. (3) McDonald's reliability estimate, "omega", will appear at the end of an IStats report (reference).
index of reliability	In "CTT", classical test theory, this is the correlation between observed scores and true scores. It is computed by finding the square root of the reliability estimate.
standard error of measure ment	Often abbreviated as "SEM", this is a measure of measurement precision, inversely related to reliability. Very reliable tests will have little measurement error. The SEM is often used to form "confidence intervals". Adding and subtracting one SEM from any test score gives what is often called the "68% confidence interval", a range of scores which, given certain assumptions, captures the true test score with a probability of 0.68. In CTT, SEM is an average figure. A more accurate estimate of measurement error is found by computing the "conditional standard error of measurement", a figure which varies by test score. Lertap's "csem" report provides estimates of measurement error by test score.

Related tidbits:

To see how Lertap can be used to calculate split-half reliability estimates, have a go with the "HalfTime" dataset.

Lertap will also compute an estimate of McDonald's omega reliability estimate. See the last section in this document.

An example of equivalent-forms reliability (also known as parallel forms reliability) may be seen here.

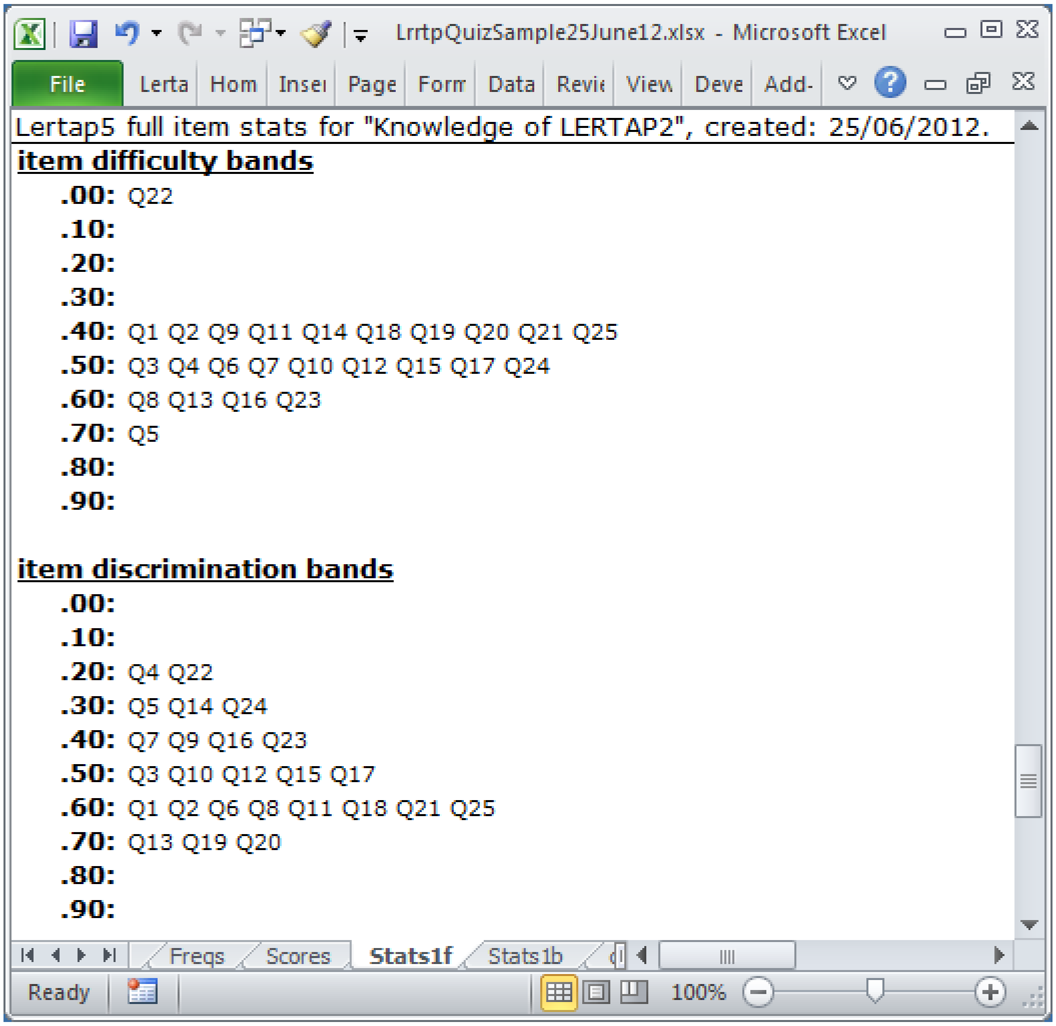
KR-20 and KR-21 are reliability estimates stemming from the work of Kuder and Richardson (1937). KR-20 and coefficient alpha will be equal when all of the test's items are scored on a right-wrong (or "dichotomous") basis, which is common for a cognitive test. When one or more of a test's items have more than one correct answer, or when an affective scale is in use, then "polytomous" scoring is in effect and coefficient alpha is the preferred estimate of reliability.

Some third-party accreditation bodies, such as NCCA (the National Commission for Certifying Agencies), will explicitly request the value of KR-20 for a test. If your test items are all dichotomous, scored on a right-wrong basis, then write in the value of coefficient alpha for KR-20 (they're the same in this case). If one or more of the test items has/have more than one right answer, then write in the value of coefficient alpha and indicate on the form that alpha has been used as KR-20 was not appropriate. (More about linking Lertap's statistics to NCCA reports is covered in this little paper.)

KR-21 is a simplified form of KR-20, used in the very rare case when all items have the same difficulty (in other words, it's not much used at all -- in its day it was a computational shortcut to KR-20).

6.4.3.1.3 The bands

Three sections of data are found at the end of a Statsf report.

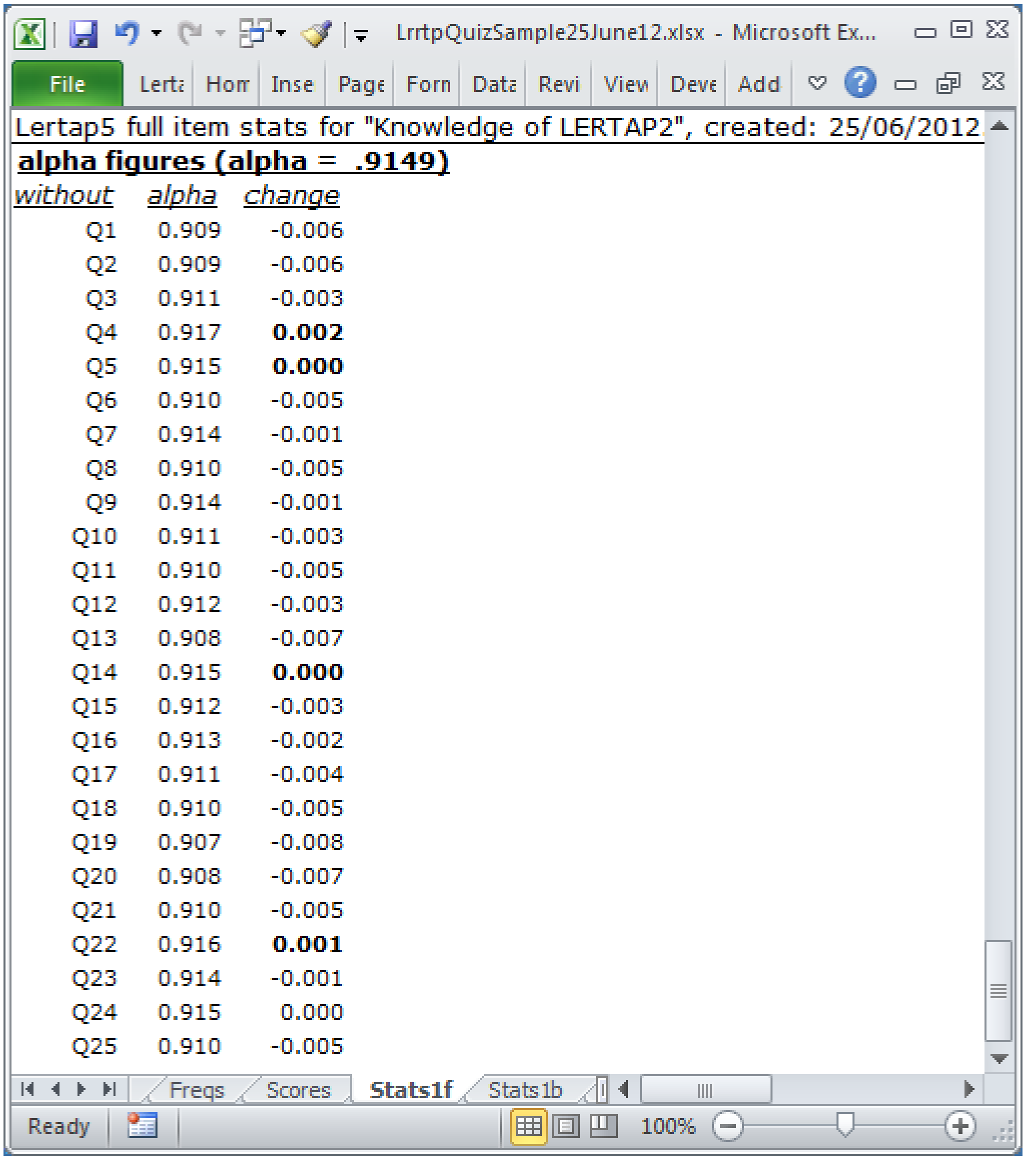


The bands summarize item difficulty and discrimination figures for cognitive tests (there are also bands for affective measures -- please see the next topic).

In this example, most items had difficulty values (proportion correct) lying between 0.40 and 0.60. Four items had difficulties above 0.60 but less than 0.70. One item, Q5, had a difficulty greater than 0.70, while another, Q22, had difficulty less than 0.10.

Item discrimination values are interpreted in the same manner. For this test, only five items (Q4, Q22, Q5, Q14, and Q24) had discrimination values below 0.40.

Item discrimination figures may be negative. When this happens such items will appear in the .00 band.



Lertap5 full item stats for "Knowledge of LERTAP2", created: 25/06/2012

alpha figures (alpha = .9149)

without	*alpha*	*change*
Q1	0.909	-0.006
Q2	0.909	-0.006
Q3	0.911	-0.003
Q4	0.917	**0.002**
Q5	0.915	**0.000**
Q6	0.910	-0.005
Q7	0.914	-0.001
Q8	0.910	-0.005
Q9	0.914	-0.001
Q10	0.911	-0.003
Q11	0.910	-0.005
Q12	0.912	-0.003
Q13	0.908	-0.007
Q14	0.915	**0.000**
Q15	0.912	-0.003
Q16	0.913	-0.002
Q17	0.911	-0.004
Q18	0.910	-0.005
Q19	0.907	-0.008
Q20	0.908	-0.007
Q21	0.910	-0.005
Q22	0.916	**0.001**
Q23	0.914	-0.001
Q24	0.915	0.000
Q25	0.910	-0.005

The last table in a Statsf report indicates how the value of coefficient alpha would change if an item were removed from the test.

In this example, with all 25 test items, the value of alpha is .9149.

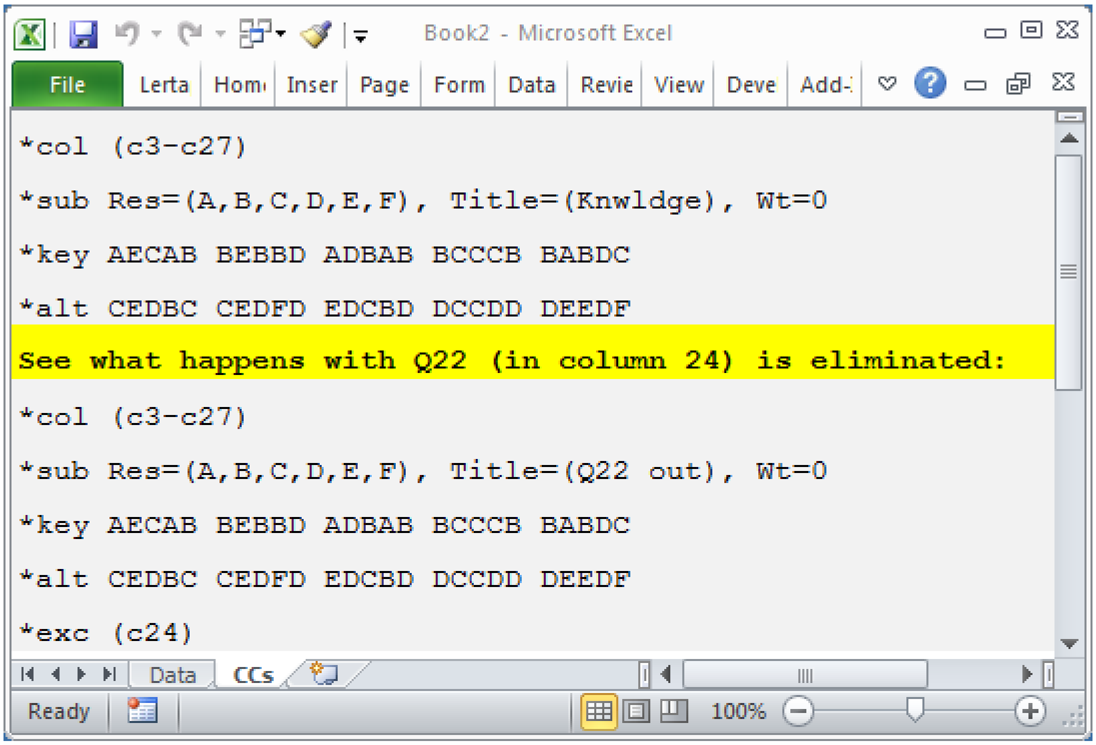
If Q1 were removed, alpha would *decrease* by -0.006, becoming 0.909. Since higher alpha means greater reliability, we would not want to drop Q1 from the test.

However, the situation with Q4 is different. Were we to eliminate it from the test, alpha would *increase* to 0.917. This might be cause to question the inclusion of this item the next time the test is used.

There is a relationship between item discrimination and the value of coefficient alpha. Items with discrimination figures above 0.30 can generally be counted on to boost alpha. Items with lower discrimination values may, on the other hand, serve to bring alpha down.

Related tidbit:

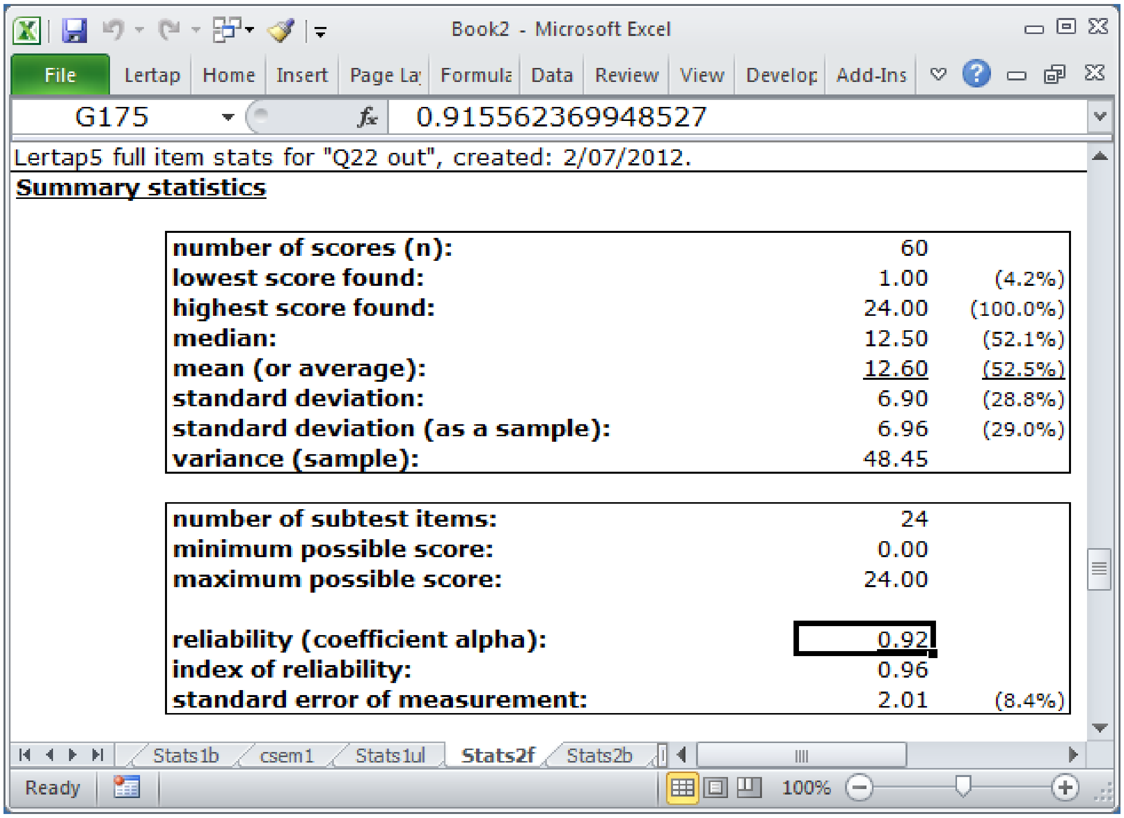
This set of CCs lines could be used to compare the reliability of the whole 25-item test with that from the same test with Q22 excluded (or eliminated):



There are two subtests defined by these lines (remember that the *col lines mark the start of subtest definitions).

A *exc line has been used to exclude the item responses found in column 24 (c24) of the Data worksheet. This was Q22 in this example.

The Stats2f report corresponding to the second subtest looked like this after Elmillon was run:



Lertap5 full item stats for "Q22 out", created: 2/07/2012.

Summary statistics

number of scores (n):	60	
lowest score found:	1.00	(4.2%)
highest score found:	24.00	(100.0%)
median:	12.50	(52.1%)
mean (or average):	12.60	(52.5%)
standard deviation:	6.90	(28.8%)
standard deviation (as a sample):	6.96	(29.0%)
variance (sample):	48.45	

number of subtest items:	24	
minimum possible score:	0.00	
maximum possible score:	24.00	
reliability (coefficient alpha):	0.92	
index of reliability:	0.96	
standard error of measurement:	2.01	(8.4%)

With 24 subtest items, coefficient alpha has the unrounded value seen in the Formula Bar (0.91556 ....).

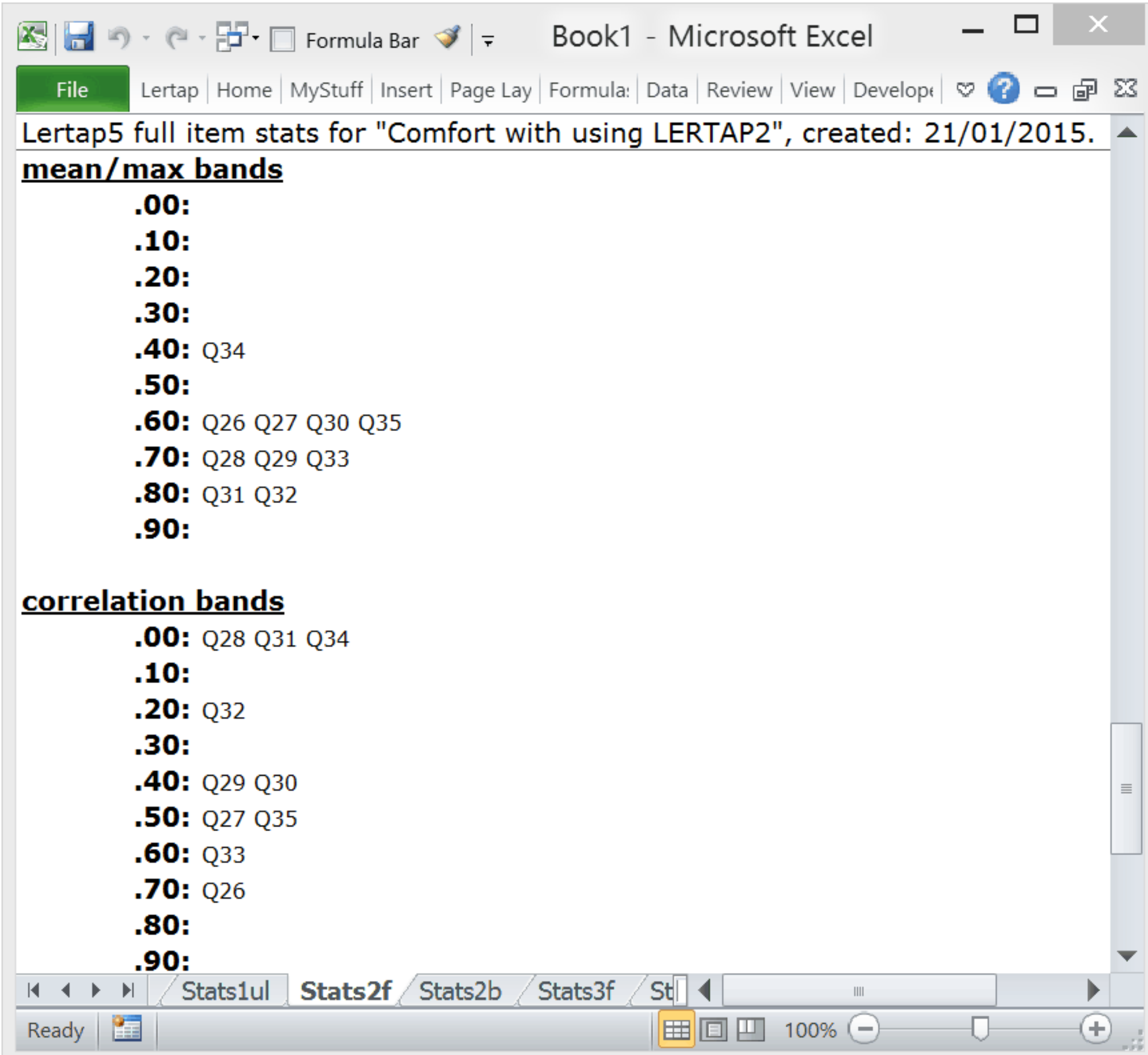
We could have also excluded Q4, found in c6 of the Data sheet pertaining to this example, by using *exc (c6, c24).

Related tidbits:

You will no doubt have noticed the Wt=0 settings on the *sub lines seen above? Whenever the CCs lines reference more than one subtest, Lertap is inclined to make a total score by adding up the subtest scores. This "Total" score will then appear in the Scores worksheet. To keep a subtest out of the total score, Wt=0 is used. If all subtests have Wt=0, then the total score will not be made.

6.4.3.1.3.1 Affective bands

These bands will appear when an affective scale has been used:



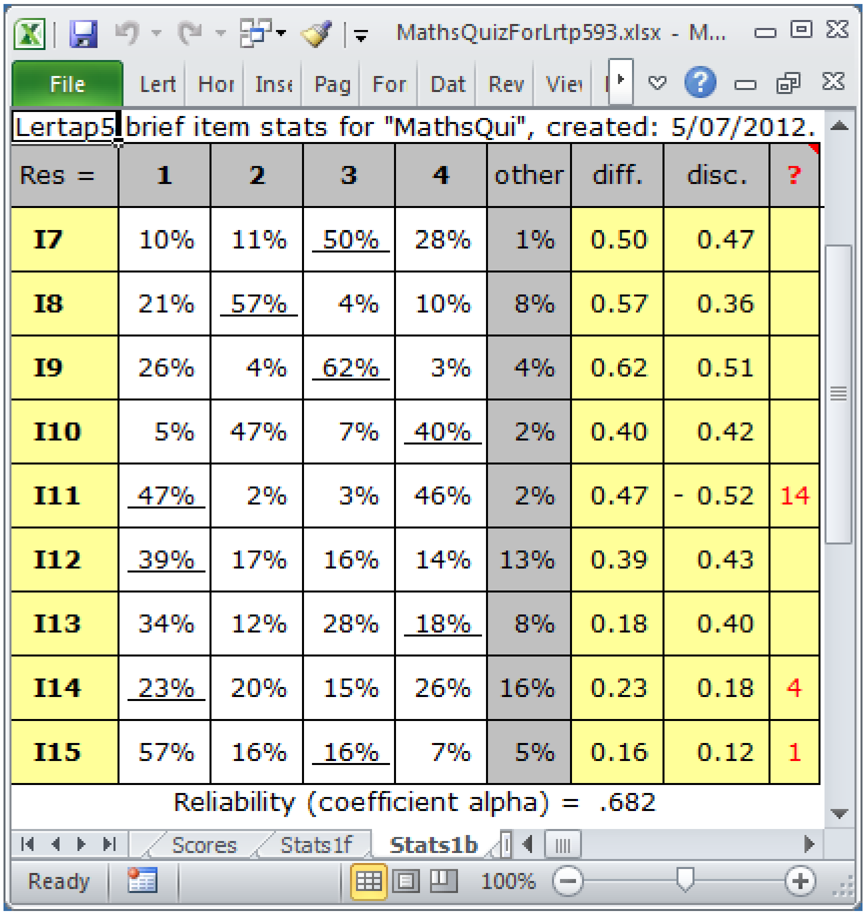
The "Mean/Max" figure is used as an indicator of how polarized an item's mean is. Q34 has the lowest mean/max figure above -- it was a five-option Likert question, going from 1 for strongly disagree to 5 for strongly agree. Its mean turned out to be 2.00. The maximum the mean could have been (had everyone selected 5) was 5.00. Dividing 2.00 by 5.00 gives the mean/max value of 0.40.

The correlation bands summarize the correlation between each item and the criterion score. The criterion score is most often simply the scale score (generally referred to as a "subtest score" in Lertap) but, at times, it may also be an "external criterion" score.

The bands, being just bands, are not too precise. The exact item mean and correlation values are found in the corresponding Stats-b summary (in the example above it would be Stats2b as the display is focused on a worksheet called Stats2f).

6.4.3.2 Brief item statistics

Statsb reports are brief versions of the information found in Statsf reports.



Part of a typical Statsb report for a cognitive test is shown above.

(Cognitive tests have a second part in Statsb, a plot of difficulty and discrimination: see examples in the next topic.)

Results for each item are presented in a single row. The percentage figures reflect the popularity of each item option, that is, the percentage of students who selected the option. This is the same as "p" in the corresponding Statsf report.

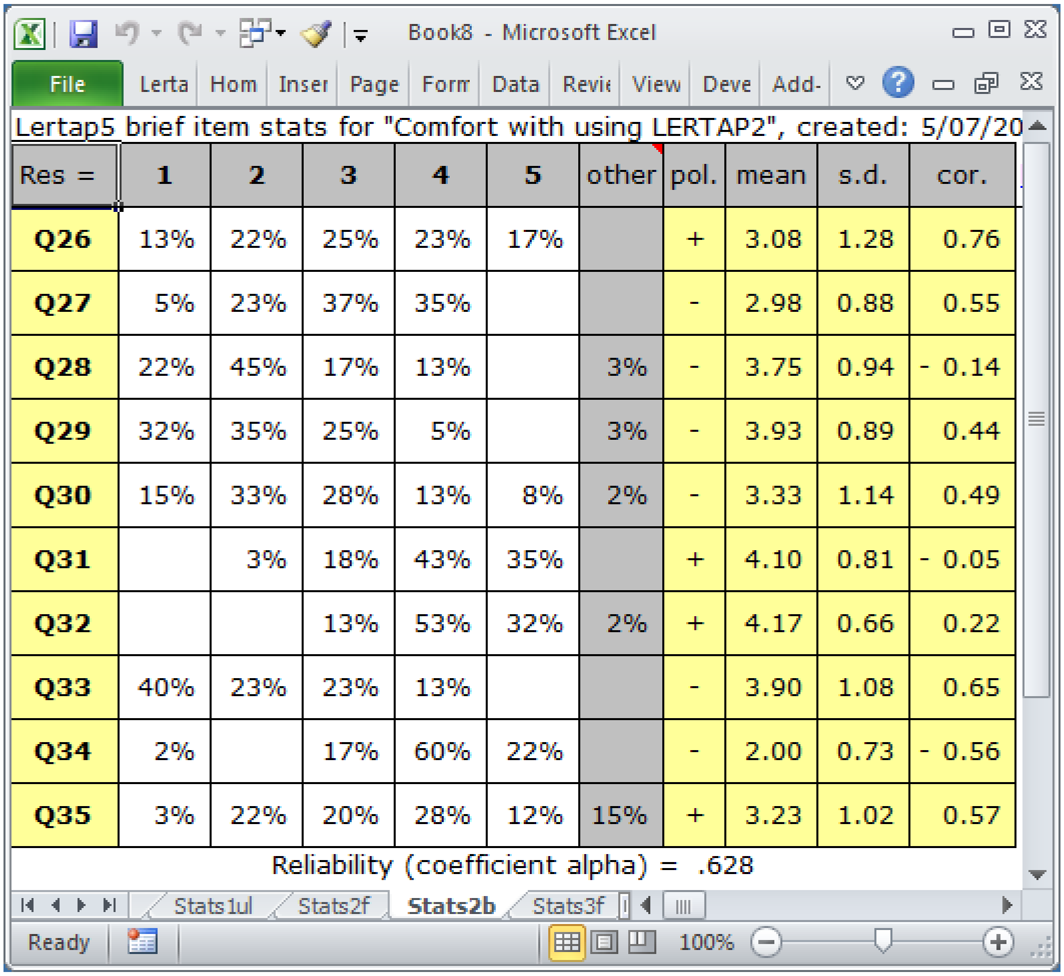
The keyed-correct option's results are underlined. (If an item has more than one keyed-correct option, results for each option are underlined.)

The "other" column gives the percentage of students who did not answer the item, or had an invalid response. Invalid responses are, for example, responses which have been incorrectly entered in the Data worksheet, resulting, perhaps, from a typing mistake. Invalid responses may also result from the use of an optical scanner. "Bubble" answer sheets are used with scanners; students will sometimes shade in more than one bubble for an item, resulting in the scanner using an invalid response code of some sort (such as an asterisk, or a "9").

Item difficulty appears in the "diff." column. If an item has only one keyed-correct option, then diff. will be the same as "p" for the correct answer in the corresponding Statsf report. If an item has more than one keyed-correct option, the percentages for each of these options are summed to get an overall figure.

The "disc." column is the item's discrimination figure. It will be the same as the pb(r) value seen in the corresponding Statsf report if the item has only one correct answer. When an item has more than one correct answer, then disc. is the Pearson product-moment correlation between the item and the criterion score, corrected for part-whole inflation.

The ? mark column is used to flag potential problems. Click here to read more about these flags.



Lertap5 brief item stats for "Comfort with using LERTAP2", created: 5/07/20

Res =	1	2	3	4	5	other	pol.	mean	s.d.	cor.
Q26	13%	22%	25%	23%	17%		+	3.08	1.28	0.76
Q27	5%	23%	37%	35%			-	2.98	0.88	0.55
Q28	22%	45%	17%	13%		3%	-	3.75	0.94	- 0.14
Q29	32%	35%	25%	5%		3%	-	3.93	0.89	0.44
Q30	15%	33%	28%	13%	8%	2%	-	3.33	1.14	0.49
Q31		3%	18%	43%	35%		+	4.10	0.81	- 0.05
Q32			13%	53%	32%	2%	+	4.17	0.66	0.22
Q33	40%	23%	23%	13%			-	3.90	1.08	0.65
Q34	2%		17%	60%	22%		-	2.00	0.73	- 0.56
Q35	3%	22%	20%	28%	12%	15%	+	3.23	1.02	0.57

Reliability (coefficient alpha) = .628

A typical Statsb report for an affective test is shown above.

The "pol." column indicates the "polarity" of the item; if the item has been reverse-scored, then the polarity will be negative.

The item's average score is found in the "mean" column. Its standard deviation is seen under "s.d.", and the correlation of the item with the criterion score, corrected for part-whole inflation, is under the "cor." column. As is the case in other Lertap output, the standard deviation is that for the population (the equation used has "n" in the denominator, not "n-1"). The correlation is a Pearson product-moment coefficient.

A plot of item "cor." values will be found at the bottom of Statsb reports (see examples here).

Missing data can have an effect on the statistics found in many of Lertap's reports, including Statsb. Read more about it here.

Related tidbits:

How to print Lertap's reports? Not hard at all, especially if you take in this topic.

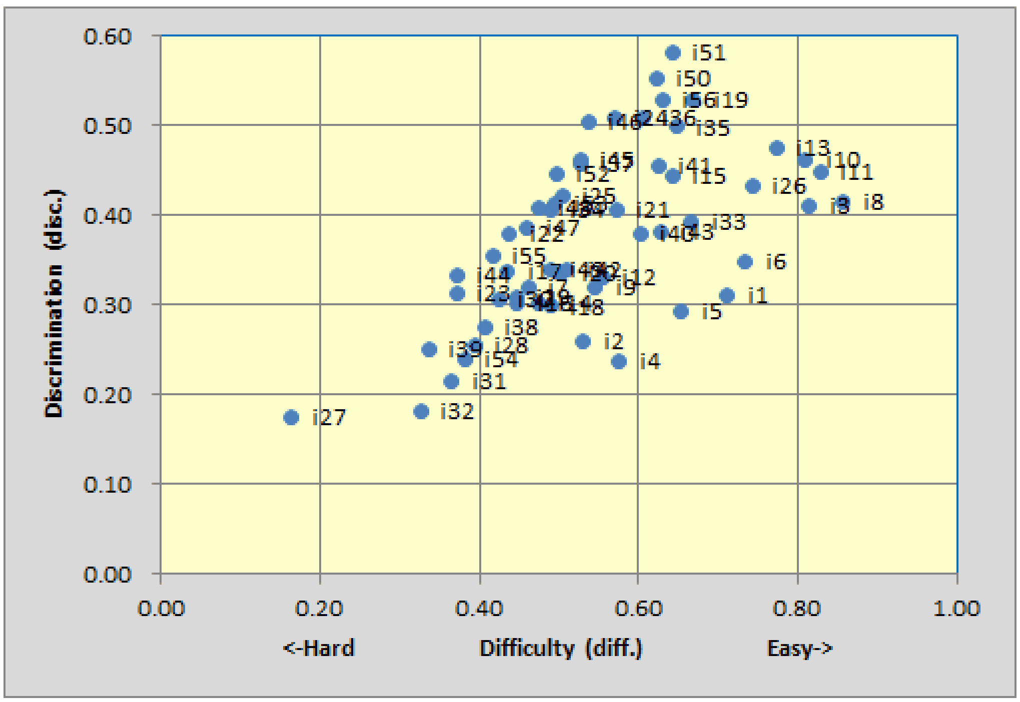
The StatsbSortAZ special macro may be used to sort a Stats_b report, such as Stats1b, on the basis of item difficulty or item discrimination.

Summary statistics for item difficulty and discrimination are now found in Stats_b reports, providing users have Lertap version 5.10.7.2 or later. The Stats1b and Stats2b screen snapshots at the bottom of this topic display these statistics, mean and standard deviation, at the base of the difficulty and discrimination columns.

6.4.3.2.1 Stats-b plot (cog.)

A plot of item difficulty by discrimination is given at the bottom of every Statsb report for cognitive tests.

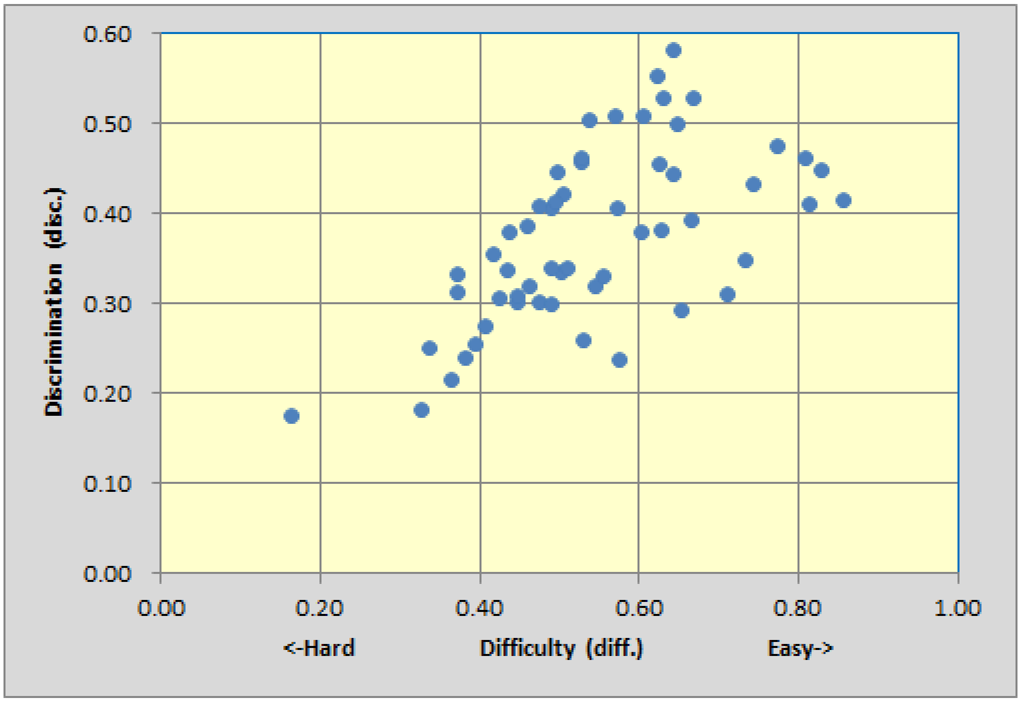
The screen shots found here were taken from the Excel 2010 /2013 version of Lertap 5. In this case a 56-item cognitive test had been processed.



These plots of item difficulty by discrimination are made using the "diff." and "disc." figures displayed in a Statsb report. As these things go (Joe), this plot suggests a test with high reliability -- almost all of the items have discrimination figures above 0.20, with many going over 0.30. Coefficient alpha for this administration of the 56 items was 0.91.

The item labels in the case above were i's followed by an item number, i1 through i56; Lertap uses whatever labels are found in the first column of the Statsb report -- when these labels are longer than four characters, Lertap will use only the last four. For example, if an item label is Ques107, the plot will show s107.

Displaying labels will often make the plot crowded, and a bit hard to read. Fortunately, the labels may be easily hidden. To do so, use the "labels toggle" option. This is what the plot looks like after the option has been used to hide the labels:

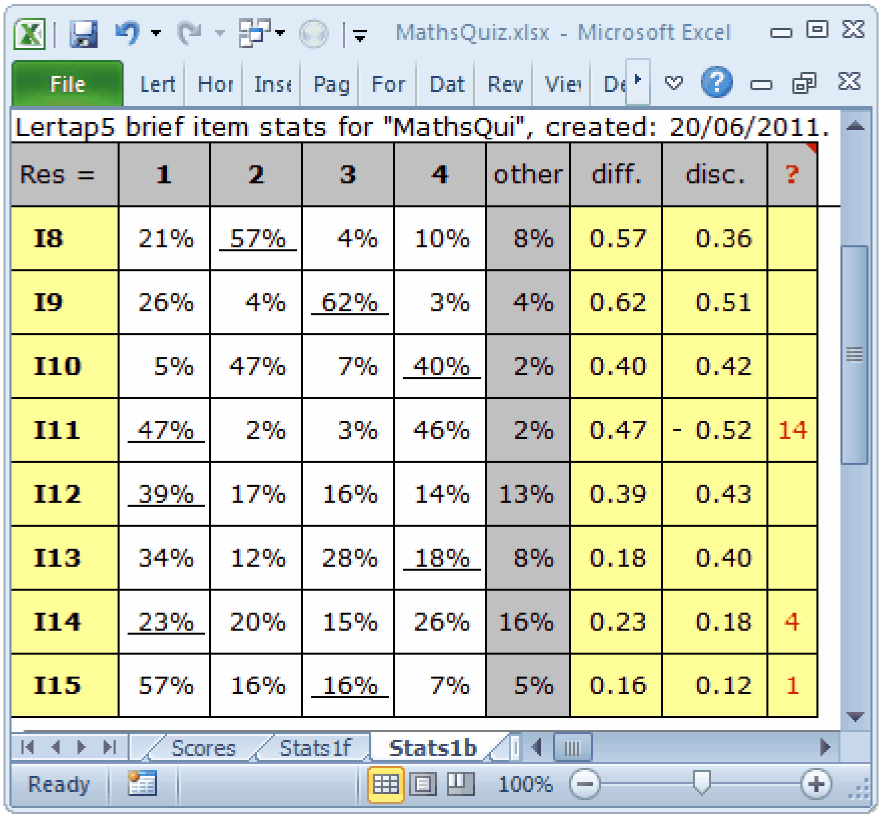


These plots are simply Excel charts; you can reformat them in many ways -- right-click or double left-click at various spots, and options will appear.

6.4.3.2.2 Stats1b ? column

The ? column is used to flag items which may have a problem.

The screen snapshots seen below were taken while using the Excel 2010 version of Lertap.



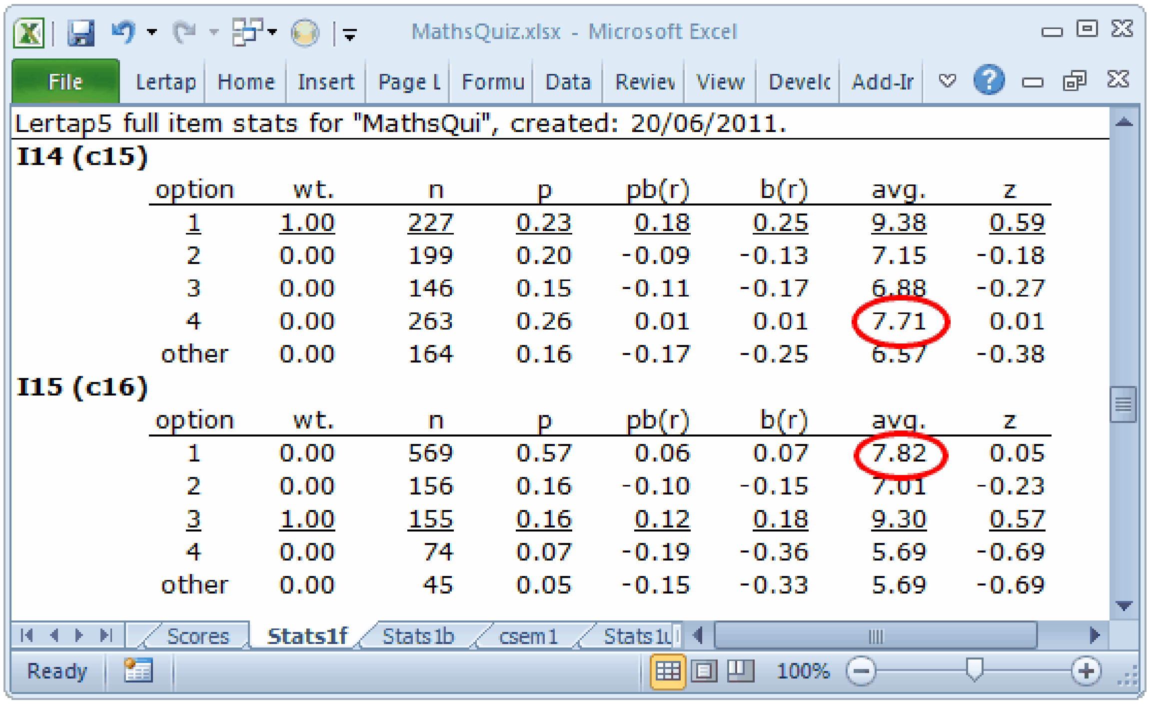
Lertap5 brief item stats for "MathsQui", created: 20/06/2011.

Res =	1	2	3	4	other	diff.	disc.	?
I8	21%	57%	4%	10%	8%	0.57	0.36	
I9	26%	4%	62%	3%	4%	0.62	0.51	
I10	5%	47%	7%	40%	2%	0.40	0.42	
I11	47%	2%	3%	46%	2%	0.47	- 0.52	14
I12	39%	17%	16%	14%	13%	0.39	0.43	
I13	34%	12%	28%	18%	8%	0.18	0.40	
I14	23%	20%	15%	26%	16%	0.23	0.18	4
I15	57%	16%	16%	7%	5%	0.16	0.12	1

An item's distractors will enter in the ? column when the distractor is not selected by anyone, or when it is selected by students with above average proficiency.

"Above average proficiency" means that the students selecting the distractor had an average test or criterion score which was above the mean of all the students who sat the test. (Note that the criterion score may be an external one.)

Let's look at items I14 and I15 in the Stats1f report:



Lertap5 full item stats for "MathsQui", created: 20/06/2011.

I14 (c15)

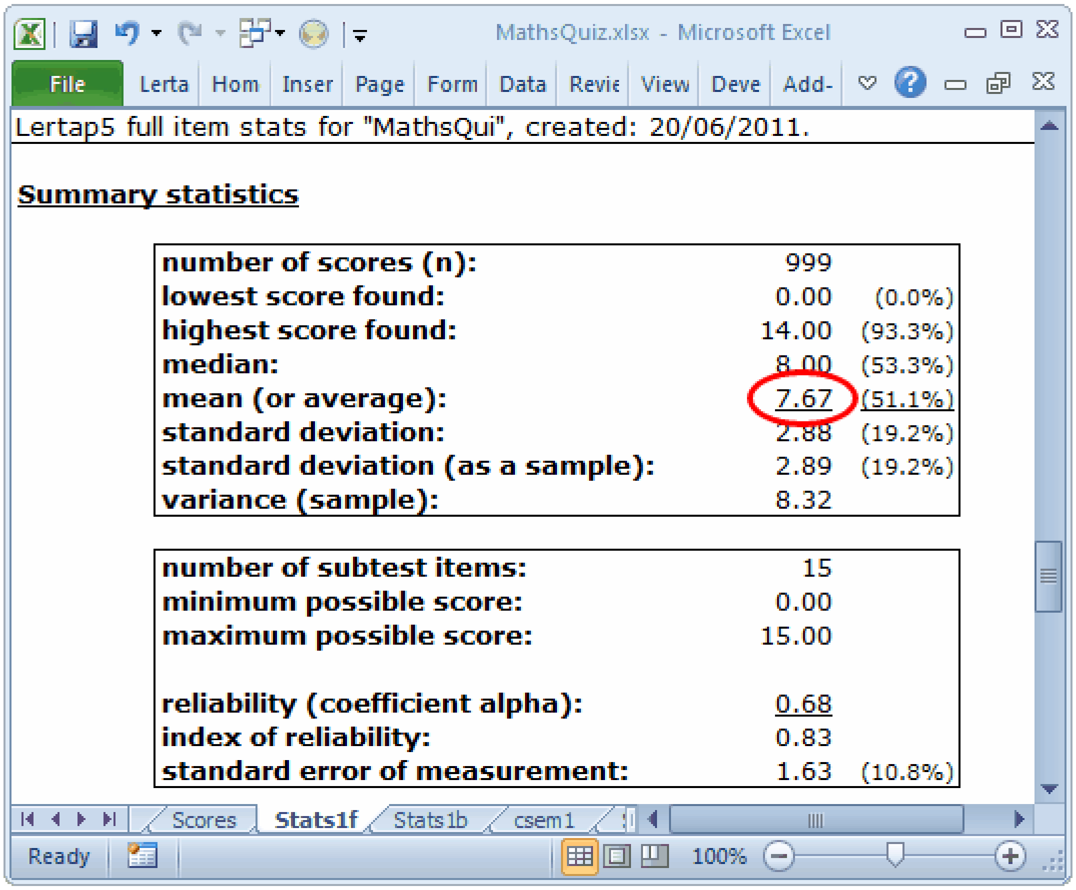
option	wt.	n	p	pb(r)	b(r)	avg.	z
1	1.00	227	0.23	0.18	0.25	9.38	0.59
2	0.00	199	0.20	-0.09	-0.13	7.15	-0.18
3	0.00	146	0.15	-0.11	-0.17	6.88	-0.27
4	0.00	263	0.26	0.01	0.01	7.71	0.01
other	0.00	164	0.16	-0.17	-0.25	6.57	-0.38

I15 (c16)

option	wt.	n	p	pb(r)	b(r)	avg.	z
1	0.00	569	0.57	0.06	0.07	7.82	0.05
2	0.00	156	0.16	-0.10	-0.15	7.01	-0.23
3	1.00	155	0.16	0.12	0.18	9.30	0.57
4	0.00	74	0.07	-0.19	-0.36	5.69	-0.69
other	0.00	45	0.05	-0.15	-0.33	5.69	-0.69

The last option for I14, 4, was a distractor, an incorrect answer (the right answer or answers to an item are always underlined; options with no underlining are the distractors).

This option was selected by 263 students. The average criterion score for these students was 7.71, as seen under the avg. column. This was above the average criterion score for all students, which was 7.67, as pictured below.



Lertap5 full item stats for "MathsQui", created: 20/06/2011.

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	

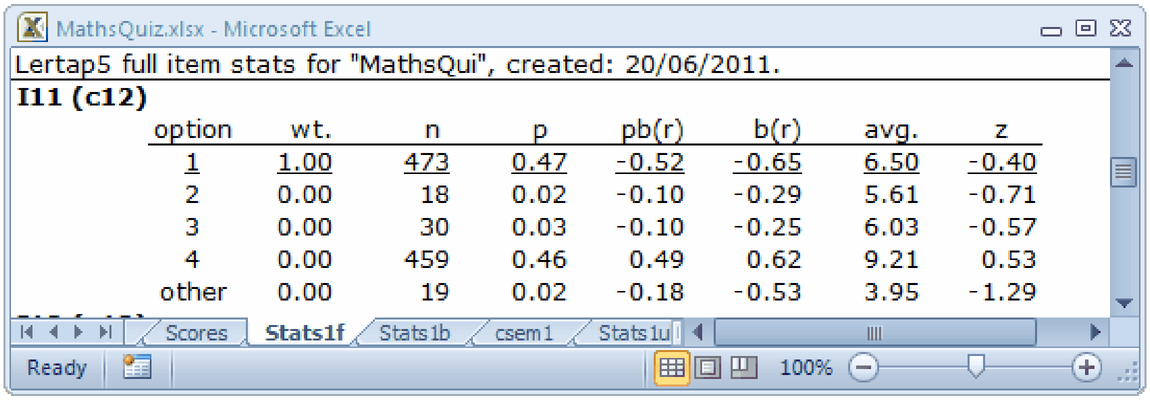
number of subtest items:	15	
minimum possible score:	0.00	
maximum possible score:	15.00	
reliability (coefficient alpha):	0.68	
index of reliability:	0.83	
standard error of measurement:	1.63	(10.8%)

On I15, the Stats1b report flagged the first option. Stats1f shows that this distractor was selected by 569 students, and their avg. score was also above the criterion average.

Keep in mind that these are just flags, notes created by Lertap to suggest that something might be amiss. We usually do not want distractors to be selected by above-average students. When they are we may have some reason to suspect "ambiguity" -- the wording of the distractor may need to be improved. In some cases a decision may be made to "double-key" an item, that is, to score the item in a manner which gives points for more than one answer. In Lertap this is done with a *mws line; an example which uses *mws lines may be found towards the bottom of this topic.

An item's keyed-correct answer (or answers) will enter the ? column when it has been selected by students whose average criterion score is below average.

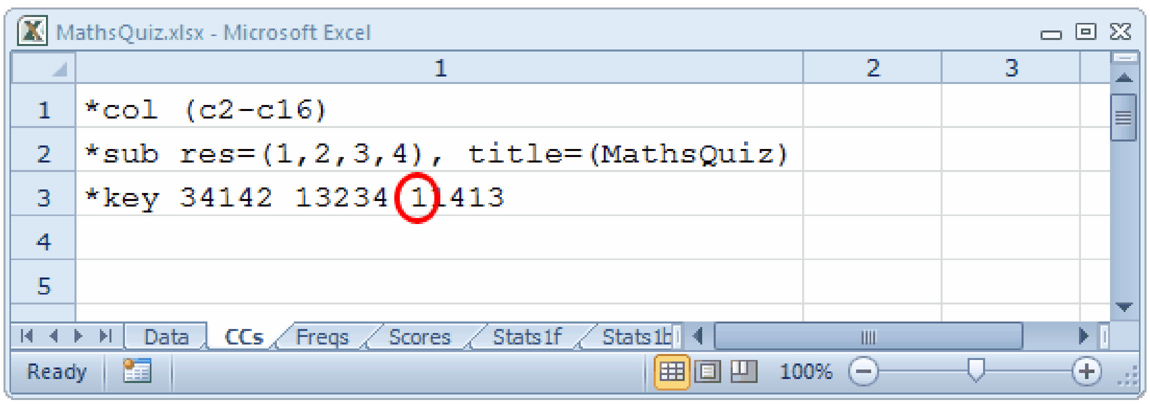
Look now at I11:



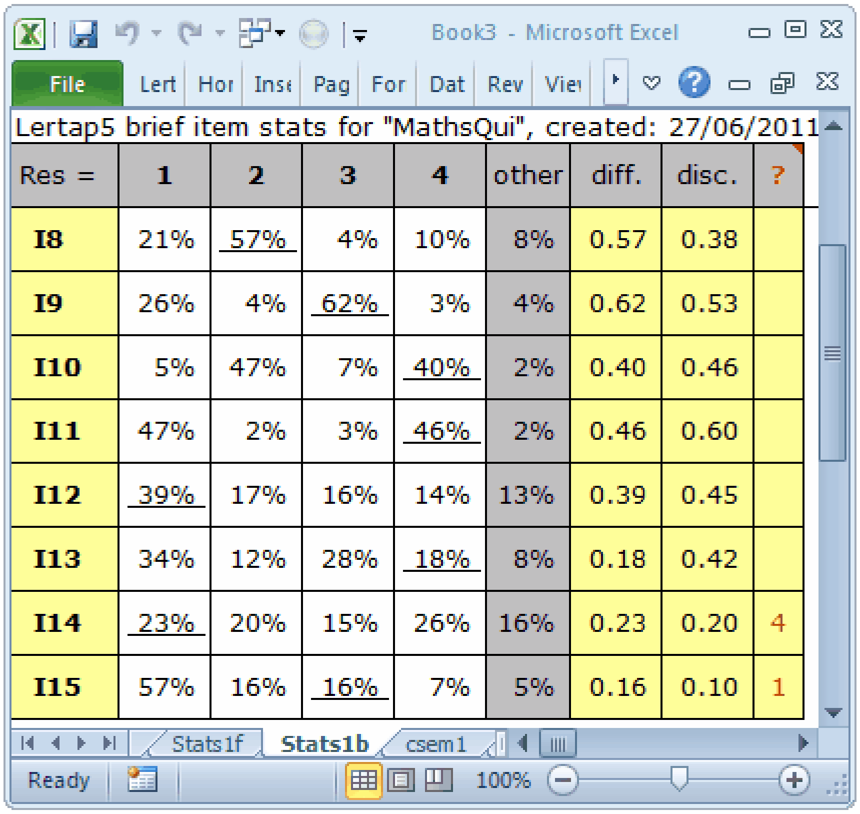
option	wt.	n	p	pb(r)	b(r)	avg.	z
1	1.00	473	0.47	-0.52	-0.65	6.50	-0.40
2	0.00	18	0.02	-0.10	-0.29	5.61	-0.71
3	0.00	30	0.03	-0.10	-0.25	6.03	-0.57
4	0.00	459	0.46	0.49	0.62	9.21	0.53
other	0.00	19	0.02	-0.18	-0.53	3.95	-1.29

The Stats1b report flagged two of this item's options, 1 and 4. The first of these, 1, is the keyed-correct answer, selected by 473 students whose average criterion score (avg.) was 6.50, well *below* the overall criterion average of 7.67. Option 4, a distractor, was selected by 459 students, and their avg. score was well *above* the overall criterion average of 7.67.

I11 has been mis-keyed. An error was made when the *key line for this test was entered in the CCs worksheet:



This sort of error is easy to fix. In this case, we'd change the 1, circled in red, to 4. After doing so, and after once again running the Interpret and Elmillon options, I11's entry in the Stats1b report was clear of flags in the ? column:



Lertap5 brief item stats for "MathsQui", created: 27/06/2011

Res =	1	2	3	4	other	diff.	disc.	?
I8	21%	57%	4%	10%	8%	0.57	0.38	
I9	26%	4%	62%	3%	4%	0.62	0.53	
I10	5%	47%	7%	40%	2%	0.40	0.46	
I11	47%	2%	3%	46%	2%	0.46	0.60	
I12	39%	17%	16%	14%	13%	0.39	0.45	
I13	34%	12%	28%	18%	8%	0.18	0.42	
I14	23%	20%	15%	26%	16%	0.23	0.20	4
I15	57%	16%	16%	7%	5%	0.16	0.10	1

Correcting mis-keyed items should increase the test's reliability estimate, which it did for the test featured here: the reliability went from 0.68 to 0.80.

Related tidbits:

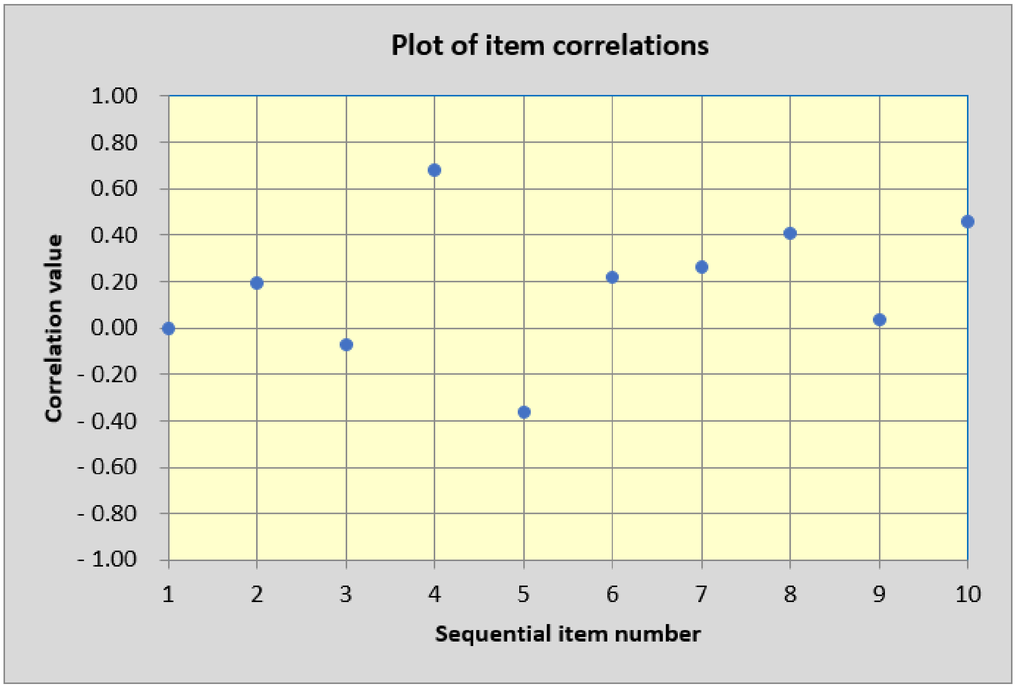
Flags are also waved in Stats1f reports, where they appear in the right margin. Read more.

How to print Lertap's reports? Not hard at all, especially if you take in this topic.

6.4.3.2.3 Stats-b plot (aff.)

A plot of affective item correlations is given at the bottom of every Stats-b report.

The screen shot below relates to the 10 items in the CEQ survey.



Each blue "blip" corresponds to the "cor." value for an item, where "cor." is an abbreviation for "correlation".

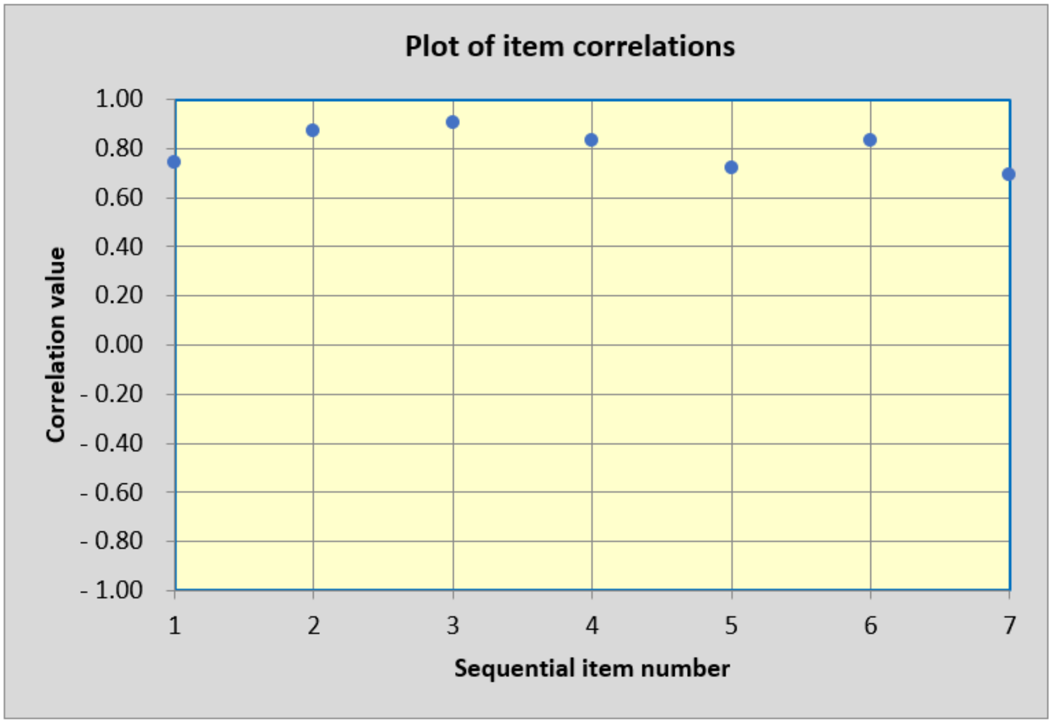
The correlation for the first item, Q1, was 0.00, as shown in the items stats table below. Q5 had the lowest correlation (-0.36).

Lertap5 brief item stats for "CEQ", created: 26-Mar-21.

Res->	1	2	3	4	5	other	pol.	mean	s.d.	cor.
Q1				100%			+	4.00	0.00	0.00
Q2			17%	67%	17%		+	4.00	0.58	0.20
Q3		58%	42%				-	3.58	0.49	- 0.07
Q4	42%	58%					-	4.42	0.49	0.68
Q5	50%	50%					-	4.50	0.50	- 0.36
Q6	25%	50%	25%				-	4.00	0.71	0.22
Q7	50%	50%					-	4.50	0.50	0.26
Q8			17%	58%	25%		+	4.08	0.64	0.41
Q9		8%	25%	58%	8%		+	3.67	0.75	0.03
Q10	25%	33%	42%				-	3.83	0.80	0.46
							Average:	4.06	0.55	0.18

The authors of the CEQ questionnaire never intended for a total score to be made by summing question responses. Of course, Lertap doesn't know this -- it'll go on, pretending that a total score was of interest, and will report scale reliability by computing and displaying the coefficient alpha reliability estimate. It came out to be 0.448. This is low, too low -- these 10 items don't "hang together"; there would be no justification for forming a total score. Of course this is not to say that the survey results were of no value - they certainly were - but the information gathered by the CEQ survey would be interpreted solely on an item by item basis, not a total score.

Here's another example: the plot below is from the seven "technology" items in the DunnSES scale mentioned in this study.



The DunnSES item correlations all tended to be high. No correlation was less than 0.60; there were no negative correlations.

The value of alpha in this case was 0.936 - there would be support for forming and using a total score from these seven items.

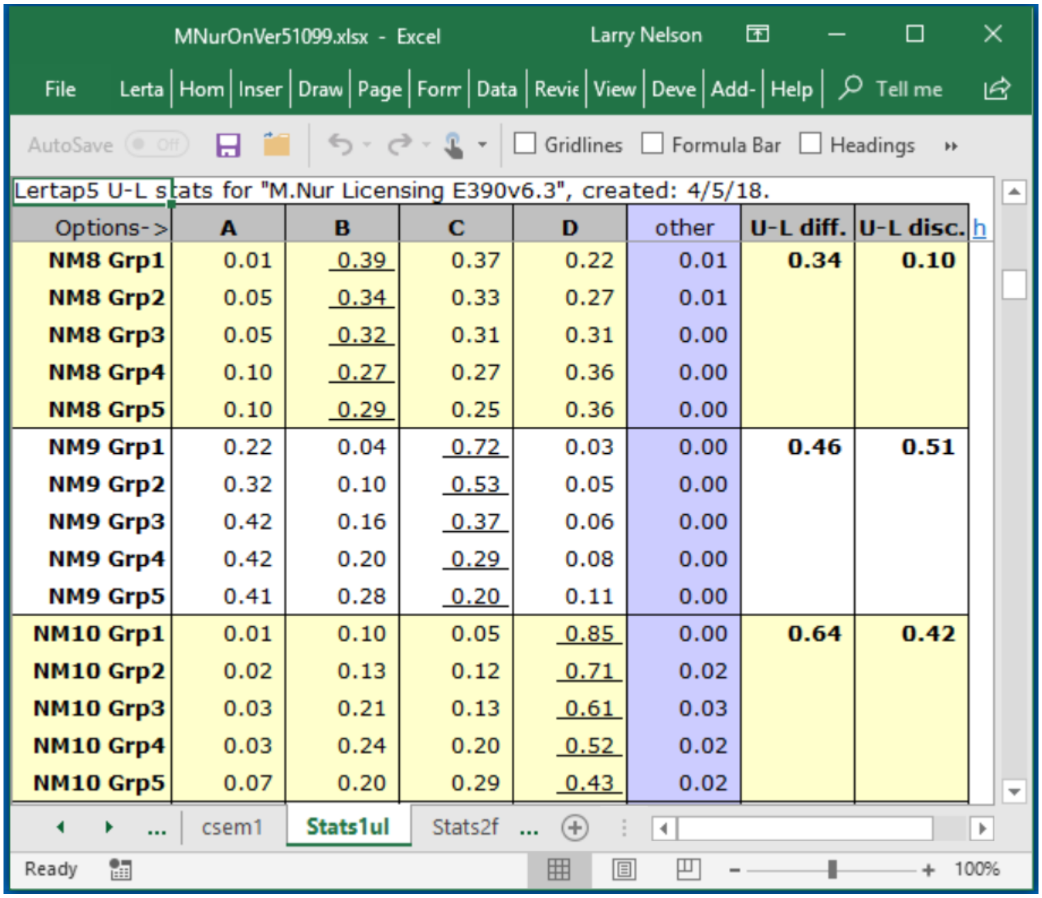
6.4.3.3 Stats1ul

These reports offer another way of looking at how items have performed, one which does not rely at all on the use of correlation coefficients. To some people they're easier to understand. Stats1ul reports provide the gateway to Lertap's very popular "quintile plots".

The "ul" in Stats1ul means "upper-lower". To make a Stats1ul report, Lertap begins by sorting all test scores from highest to lowest. It then uses the sorted scores to form groups of students. The top students go into a group called "Grp1" while the weakest students, the lowest-scoring ones, will go into a group denoted as "GrpX", where X is equal to the total number of groups.

When this method was initially devised, a long time ago, before computers became commonplace, it was often referred to as the "high-low" method, and only two groups were used: the top 27% and the bottom 27% (see Chapter 10 of the manual for more information).

Lertap has a setting in its System worksheet which determines how many groups will be formed. The minimum is two, the maximum is ten, the default is five. When the default value is in effect, the lowest-scoring group will be known as "Grp5".



Lertap5 U-L stats for "M.Nur Licensing E390v6.3", created: 4/5/18.

Options->	A	B	C	D	other	U-L diff.	U-L disc.
NM8 Grp1	0.01	0.39	0.37	0.22	0.01	0.34	0.10
NM8 Grp2	0.05	0.34	0.33	0.27	0.01		
NM8 Grp3	0.05	0.32	0.31	0.31	0.00		
NM8 Grp4	0.10	0.27	0.27	0.36	0.00		
NM8 Grp5	0.10	0.29	0.25	0.36	0.00		
NM9 Grp1	0.22	0.04	0.72	0.03	0.00	0.46	0.51
NM9 Grp2	0.32	0.10	0.53	0.05	0.00		
NM9 Grp3	0.42	0.16	0.37	0.06	0.00		
NM9 Grp4	0.42	0.20	0.29	0.08	0.00		
NM9 Grp5	0.41	0.28	0.20	0.11	0.00		
NM10 Grp1	0.01	0.10	0.05	0.85	0.00	0.64	0.42
NM10 Grp2	0.02	0.13	0.12	0.71	0.02		
NM10 Grp3	0.03	0.21	0.13	0.61	0.03		
NM10 Grp4	0.03	0.24	0.20	0.52	0.02		
NM10 Grp5	0.07	0.20	0.29	0.43	0.02		

The snapshot above depicts a typical Stats1ul report for three items from the M.Nursing dataset, using the default setting of five groups, with 20% of the students in each group.

Results for three items are shown. As you look at them, remember that the figures for an item's correct answer are underlined. The correct answer to the item called "NM8" was B, for example. The figures shown in the main part of the display indicate the proportion of students in each group who selected the item's options.

When we want our items to be ones which are capable of discriminating among the students, picking out the strongest while simultaneously identifying the weakest, then the best students should get the items right, while the weakest falter. This happened on two of the items seen above: item NM9, and item NM10.

On NM9, 72% of the students in the upper group picked out the right answer, dropping steadily to 20% in the lower group. A similar pattern was noted on NM10, although the drop was not quite as marked as 43% of the lower students got the item right.

NM8 was a different case. There's only a small difference between the top and bottom groups, with 39% of Grp1, the top group getting it right, compared to 29% of the low group, Grp5. A closer look reveals that the top group was quite undecided on NM8; many students in this group thought that option C was a good choice, and a fair number also went for option D.

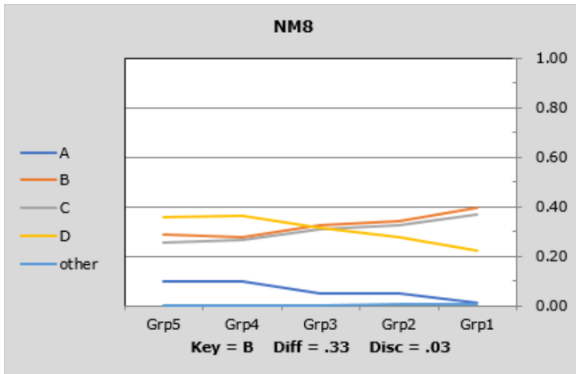
The "other" column gives the proportion of students who did not select one of the item's options, or who had an "invalid response" (which might come from shading in more than one answer of a "bubble" answer sheet).

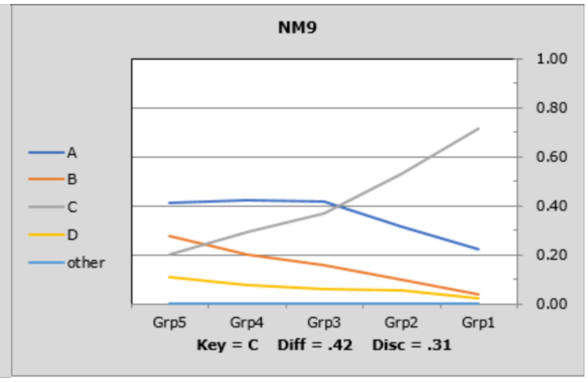
"U-L diff" is a measure of item difficulty, corresponding to the proportion of students who got the item correct (over all groups). It will usually closely correspond to the "diff" value seen in the Stats1b report, and to "p" in Stats1f .

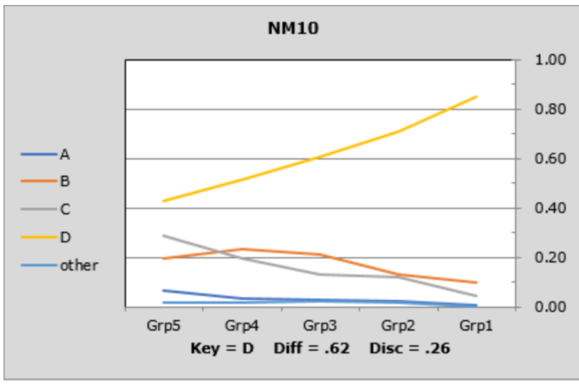
"U-L disc" is an index of item discrimination, designed to be easy to understand: it's simply the difference between the proportions of correct answers in the upper and lower groups. For NM8, the corresponding proportions are 0.39 and 0.29, giving UL-disc=0.10. UL-disc values will frequently *differ* from the "disc" figure found in Stats1b, and from the "pb(r)" in Stats1f; Stats1b and Stats1f use correlation methods to index discrimination, while, as we have seen, Stats1ul uses simple proportions. When an item's "U-L disc" is less than zero, red coloring is used to "flag" this unwanted outcome.

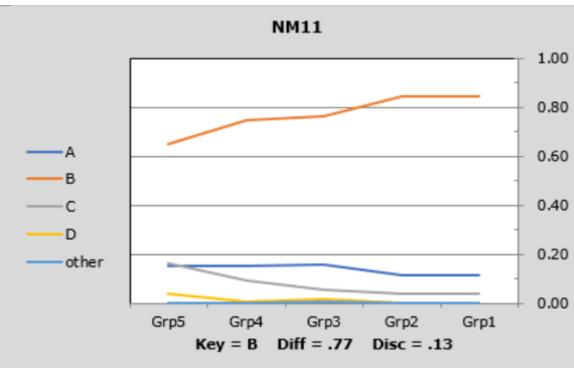
NM9 had the best discrimination of these three items, with UL-disc=0.51 (all of the proportions have been rounded; it would appear that UL-disc for this item should be 0.72-0.20=0.52, but the 0.72 has been rounded up from 0.716713 -- you can always see the unrounded figures by turning on Excel's Formula Bar).

Although the statistics which underlie Stats1ul reports are quite simple ones, based on proportions, nevertheless there's a lot of information to take in. As an alternative to the tables, Lertap is most happy to turn them into charts which are often easier to interpret. Look:









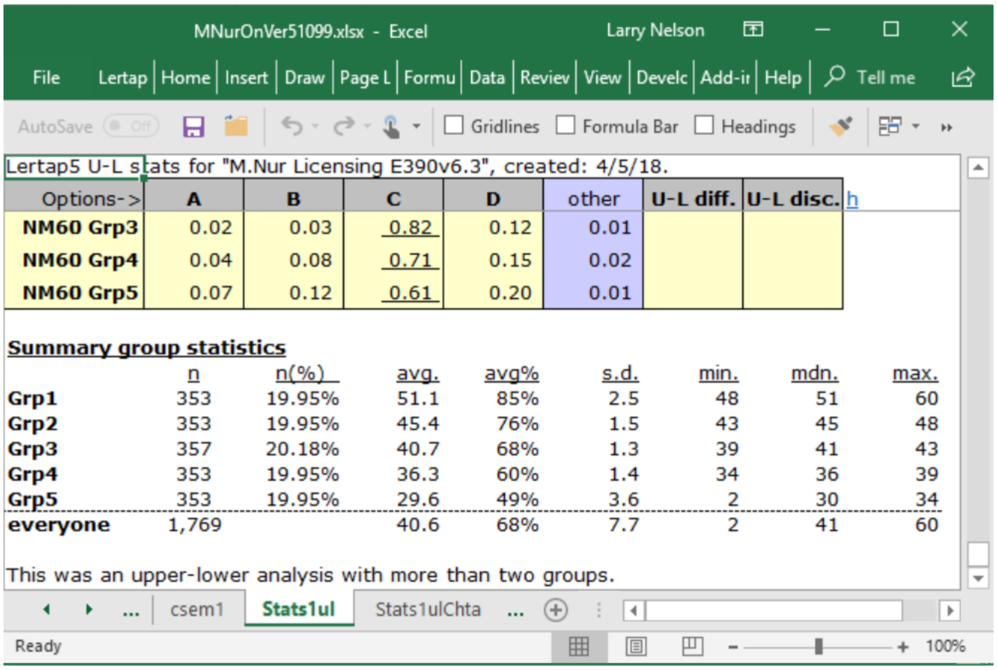
The lines in these graphs are referred to as "trace lines" as they trace how each item's options "perform" in each group. When the Stats1ul results are based on five groups, these graphs are called "quintile plots".

What we frequently want to see is a pattern like that seen for items NM9 and NM10. In the lower group, "Grp5", the proportions for the options should be fairly similar, while in the upper group, "Grp1", most students will (hopefully) be able to identify the correct answer, making the trace lines sort of fan out as they head to the right.

NM8 shows the often-desired pattern in the lower group, but fails to fan on the right.

NM11 was quite easy, even in the lower group. It has no chance of fanning out on the right as it has essentially already fanned on the left.

Graphs such as these are so popular they've spawned numerous supporting documents. Click here for more information.



Stats1ul reports always have a little table at the bottom, such as that shown above. In this case, Lertap has formed five groups, each with about 20% of the students.

n	The number of students in each group.
avg.	The mean (or average) of the test scores for the group.
avg%	The avg. figure expressed as a percentage of the maximum possible score (which was 60 in this case).
s.d.	The standard deviation of the scores for the group. Computed as a "population" value. If you're familiar with the equations used to calculate this statistic, the one used here has "n" in the denominator (not "n-1").
min.	The lowest, or minimum, score found in the group.
mdn.	The median of the group's scores (the 50th percentile).
max.	The highest, or maximum, score found in the group.

Related tidbits:

The screen snapshots seen above were taken from an analysis of the exam responses given by a class of nursing students.

In earlier versions of Lertap 5, no matter how many groups there were, the top group was always called the "Upper" group, while the bottom group was always referred to as the "Lower" group. These labels are also seen in the manual.

Quintile plots have given rise to the term "visual eye-tem analysis", as seen in this paper.

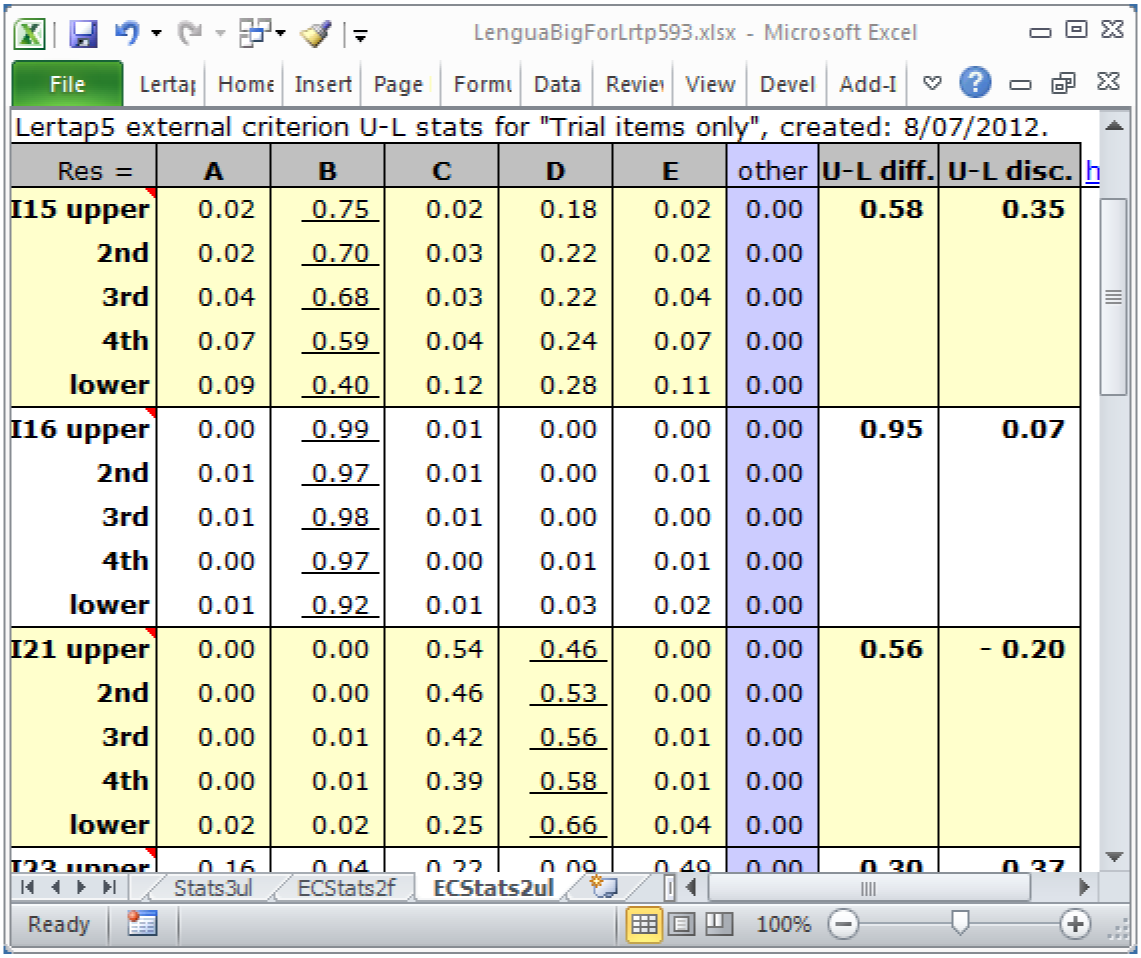
How to print Lertap's reports? See this topic.

6.4.3.3.1 With external criterion

Stats1ul reports become "ECStats1ul" reports when an external criterion analysis has been selected.

The discussion found in this topic is based on results from the LenguaBIg dataset. The "Core" score was used as the external criterion, and the second subtest, "Trial items only", was selected for the analysis. Note that in this example we have a "ECStats2ul" report as we're dealing with the second subtest in the dataset.

Note: the display below uses "upper" and "lower" as group labels. These are now known as groups "Grp1" and "GrpX", as mentioned in the previous topic.

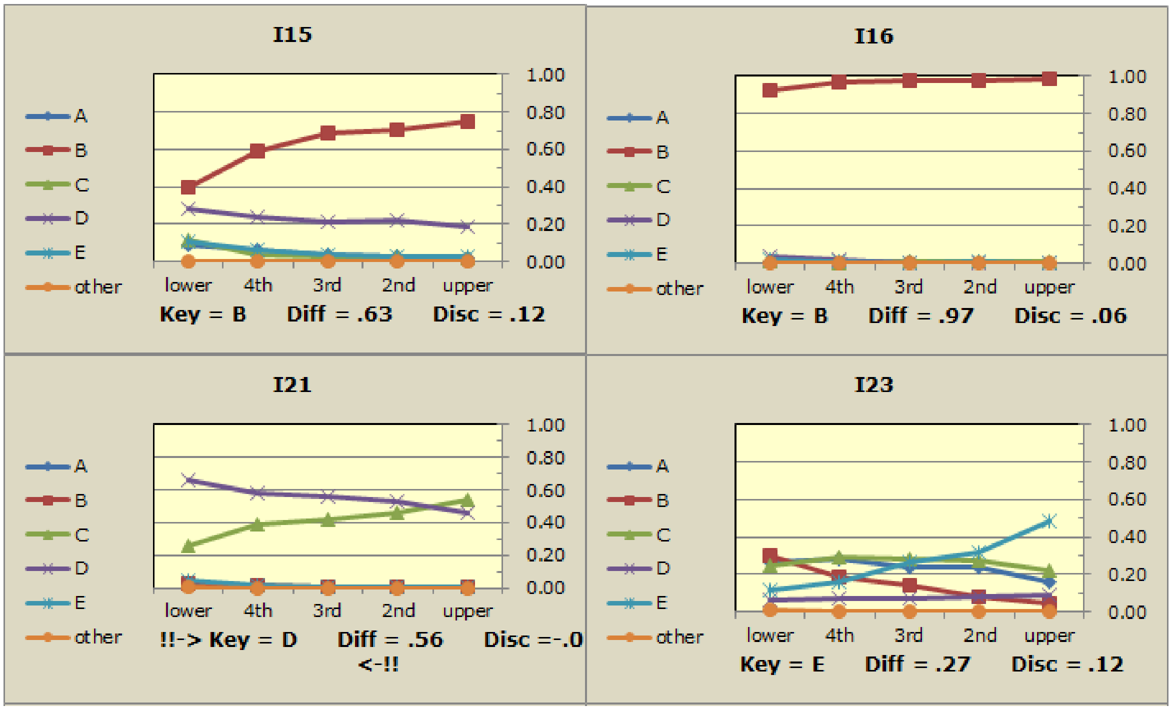


Lertap5 external criterion U-L stats for "Trial items only", created: 8/07/2012.

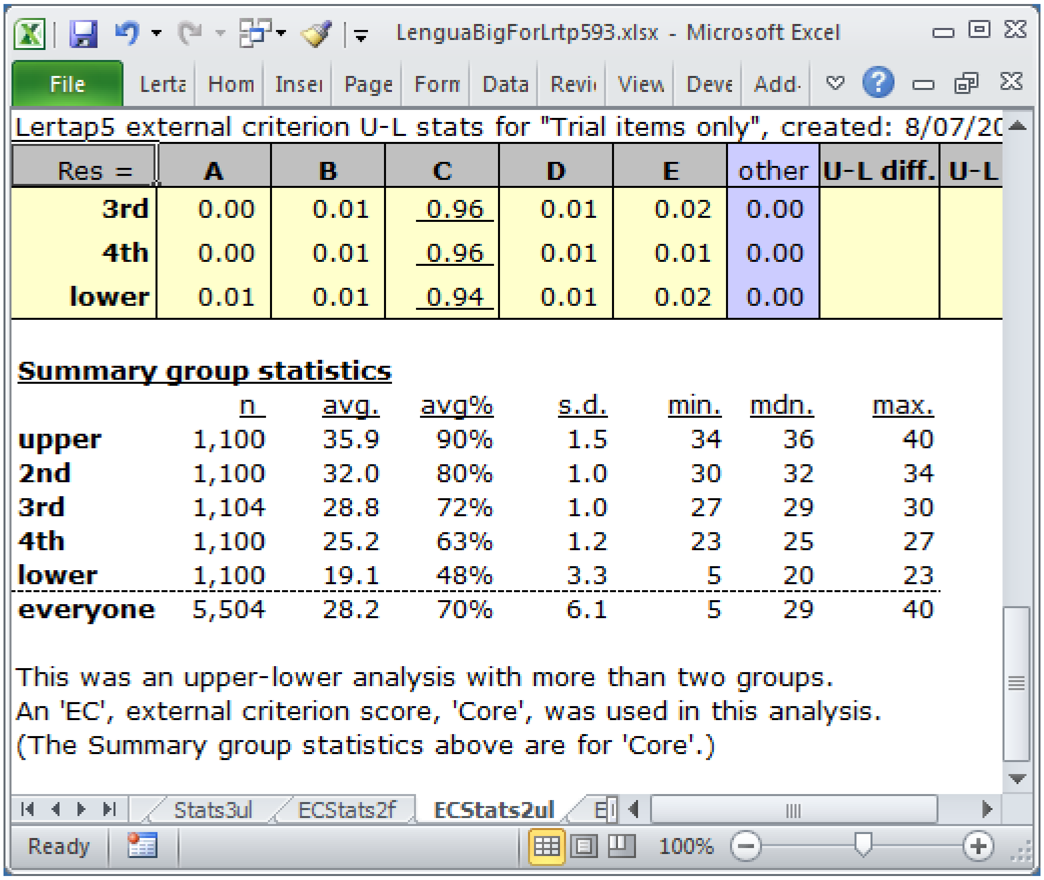
Res =	A	B	C	D	E	other	U-L diff.	U-L disc.
I15 upper	0.02	0.75	0.02	0.18	0.02	0.00	**0.58**	**0.35**
2nd	0.02	0.70	0.03	0.22	0.02	0.00		
3rd	0.04	0.68	0.03	0.22	0.04	0.00		
4th	0.07	0.59	0.04	0.24	0.07	0.00		
lower	0.09	0.40	0.12	0.28	0.11	0.00		
I16 upper	0.00	0.99	0.01	0.00	0.00	0.00	**0.95**	**0.07**
2nd	0.01	0.97	0.01	0.00	0.01	0.00		
3rd	0.01	0.98	0.01	0.00	0.00	0.00		
4th	0.00	0.97	0.00	0.01	0.01	0.00		
lower	0.01	0.92	0.01	0.03	0.02	0.00		
I21 upper	0.00	0.00	0.54	0.46	0.00	0.00	**0.56**	**- 0.20**
2nd	0.00	0.00	0.46	0.53	0.00	0.00		
3rd	0.00	0.01	0.42	0.56	0.01	0.00		
4th	0.00	0.01	0.39	0.58	0.01	0.00		
lower	0.02	0.02	0.25	0.66	0.04	0.00		
I23 upper	0.16	0.04	0.22	0.09	0.49	0.00	**0.30**	**0.37**

The groups in an external criterion analysis are formed by using the external criterion score, which in this case was called "Core". Otherwise, all statistics are formed in the manner described in the previous topic.

Note the negative U-L disc value for I21. More members of the lower group got this item right than in the upper group. You can see this is the corresponding quintile plot:



Of these four items, only I15 shows a tendency to fan out. I16 is very easy. I23 is quite difficult. I21 is goofy: the trace line for the keyed-correct option, D, *decreases* as we move across the plot, the opposite of what would be expected. On the other hand, one of I21's distractors, option C, *increases* from left to right, which is also the opposite of expected. (This pattern is often found when an item has been "incorrectly keyed". It could be that an error has been made, and the keyed-correct option for this item should be changed to C. However, even in this case there would still be many students in the upper group, suggesting that option D may be a plausible correct answer. Indications are that I21 requires revision before being used again.)



Lertap5 external criterion U-L stats for "Trial items only", created: 8/07/20

Res =	A	B	C	D	E	other	U-L diff.	U-L
3rd	0.00	0.01	0.96	0.01	0.02	0.00		
4th	0.00	0.01	0.96	0.01	0.01	0.00		
lower	0.01	0.01	0.94	0.01	0.02	0.00		

Summary group statistics

	n	avg.	avg%	s.d.	min.	mdn.	max.
upper	1,100	35.9	90%	1.5	34	36	40
2nd	1,100	32.0	80%	1.0	30	32	34
3rd	1,104	28.8	72%	1.0	27	29	30
4th	1,100	25.2	63%	1.2	23	25	27
lower	1,100	19.1	48%	3.3	5	20	23
everyone	5,504	28.2	70%	6.1	5	29	40

This was an upper-lower analysis with more than two groups.
An 'EC', external criterion score, 'Core', was used in this analysis.
(The Summary group statistics above are for 'Core'.)

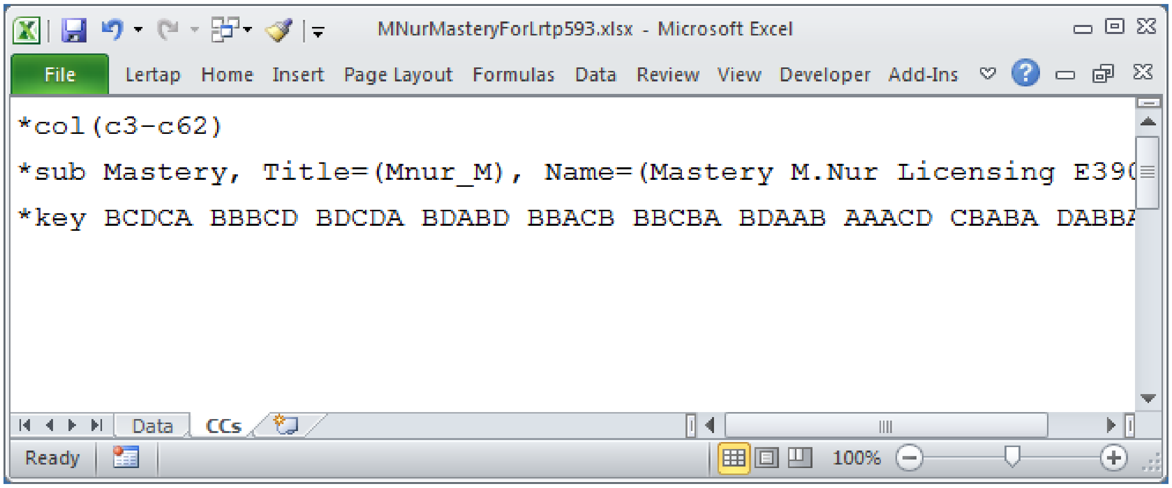
This snapshot shows how the table at the bottom of an ECStatsul report looks when an external criterion has been used to form the groups.

6.4.3.3.2 Mastery mode

"Mastery mode" refers to the situation where some sort of cutoff score has been used to classify test scores into, for example, "masters" and "non-masters", or "pass" and "fail", or "promote" and "keep back".

The cutoff score is at times referred to as a "criterion", as in "CRT", criterion-referenced testing.

Lertap's mastery mode is activated by using the word "Mastery" in a CCs *sub line, as seen here:



Lertap's mastery level is usually set at 70%. Test scores at and above this level make it into Lertap's "masters" group, while the rest go into a group called "others". It is possible to change the mastery level by using, for example, Mastery=80% on the *sub line, or by changing the default setting in row 11 of Lertap's System worksheet.

With the release of version 5.10.5.1 in early 2015, it became possible to use a raw test score for the Mastery setting: Mastery=40, for example, will set the cutoff score at 40 -- in a test with 60 items, this would usually be equivalent to Mastery=67% (assuming each item has just one correct answer worth one point).

Note that prior to version 5.10.5.1, Mastery=X always meant X%. *This is no longer the case*; if the % sign is not used, Lertap will interpret X as a raw test score. But, this does not apply to the setting in row 11 of the System worksheet: a value in row 11 will always be interpreted as percentage figure. In order to use a raw test score as a cutoff, a Mastery= assignment has to be made on the *sub card; this is an example of setting the cutoff to a raw test score of 40:

*sub Mastery=40, Title=(MyTest1), Res=(1,2,3,4)

This will set the mastery level to 40% of the maximum possible test score:

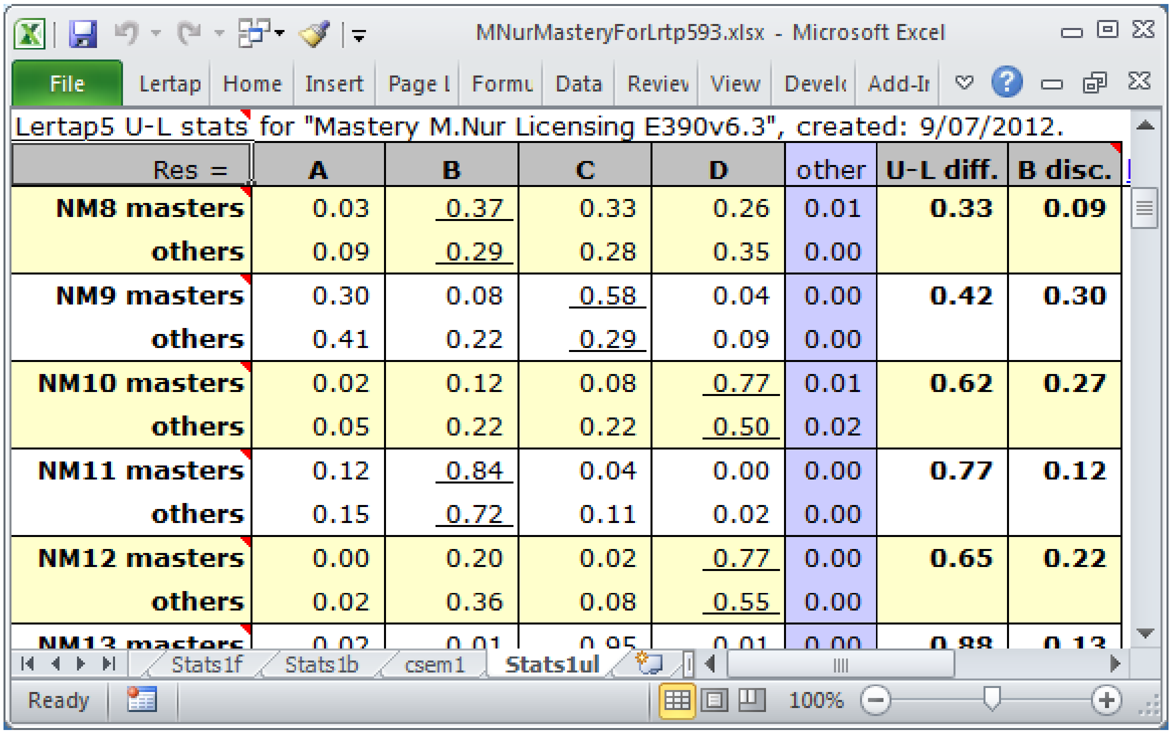
*sub Mastery=40%, Title=(MyTest1), Res=(1,2,3,4)

In the line below, since there is no = sign, Lertap will look to row 11 in the System worksheet to get its mastery cutoff setting, and it will always be a % figure.

*sub Mastery, Title=(MyTest1), Res=(1,2,3,4)

> Critical note: when using a Mastery= setting, as opposed to just the word Mastery by itself, it's advisable to keep the Mastery= setting from being the last entry on the *sub line, especially when Wt= is also used on the line. (Wt= is an assignment which controls the weight given to a subtest when it enters the total score composite, something which is relevant when there's more than one subtest.

Wt= assignments should be the last entry on a *sub line. Read a bit more about this by looking at Example C6 in this topic.)

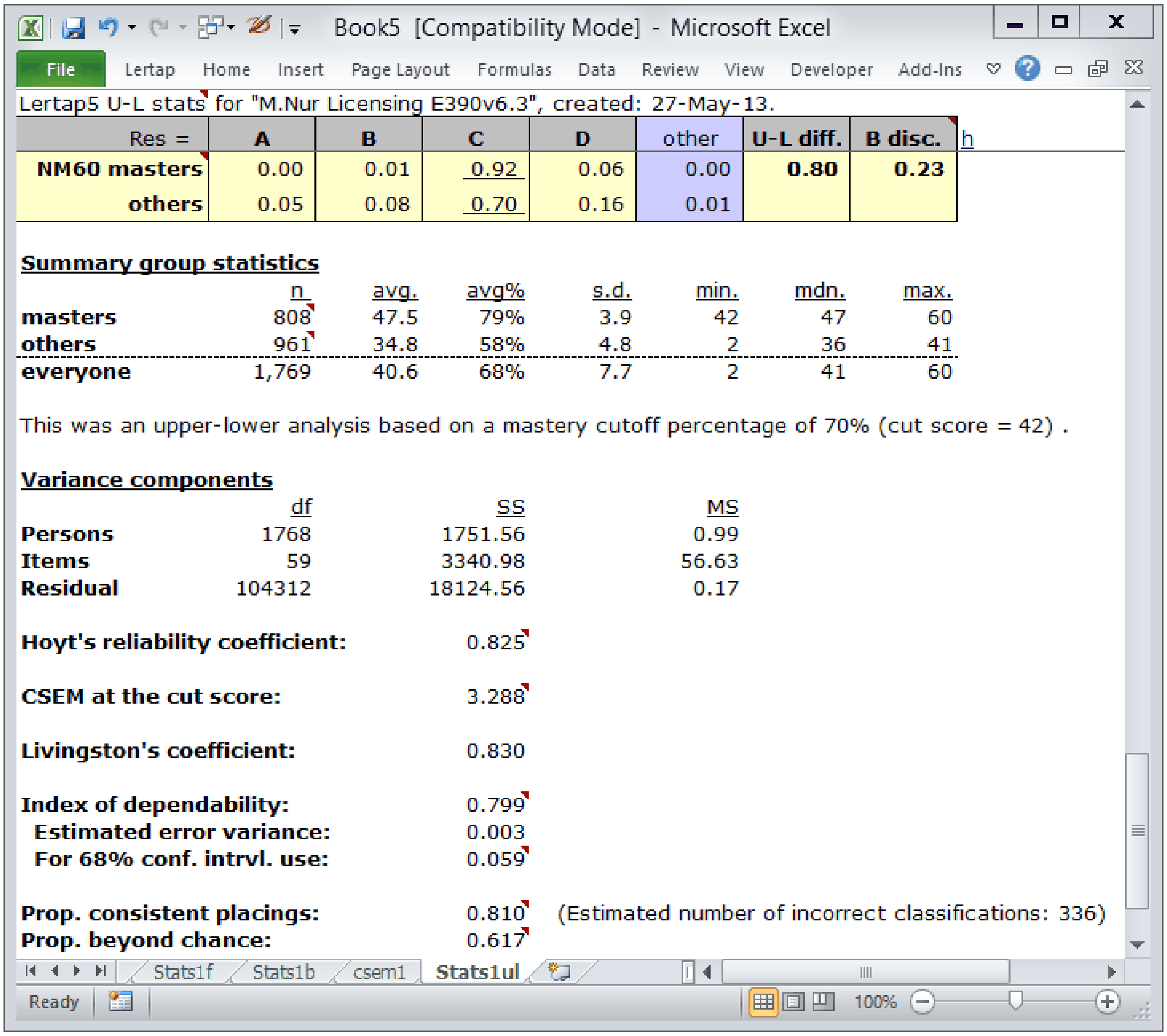


Lertap5 U-L stats for "Mastery M.Nur Licensing E390v6.3", created: 9/07/2012.

Res =	A	B	C	D	other	U-L diff.	B disc.
NM8 masters	0.03	0.37	0.33	0.26	0.01	0.33	0.09
others	0.09	0.29	0.28	0.35	0.00		
NM9 masters	0.30	0.08	0.58	0.04	0.00	0.42	0.30
others	0.41	0.22	0.29	0.09	0.00		
NM10 masters	0.02	0.12	0.08	0.77	0.01	0.62	0.27
others	0.05	0.22	0.22	0.50	0.02		
NM11 masters	0.12	0.84	0.04	0.00	0.00	0.77	0.12
others	0.15	0.72	0.11	0.02	0.00		
NM12 masters	0.00	0.20	0.02	0.77	0.00	0.65	0.22
others	0.02	0.36	0.08	0.55	0.00		
NM13 masters	0.02	0.01	0.95	0.01	0.00	0.88	0.13

The Stats1ul report in mastery mode looks like the snapshot above. With one exception, the statistics which appear are the same as the those found when a mastery level is not used (previous topic). The exception is the last column, "B disc", named after Brennan (1972), who referred to it as "a generalized upper-lower discrimination index". It is simply the difference in the proportions for the correct answer. Unlike the U-L disc value encountered in a regular Stats1ul report, the B disc value for an item depends on the where the cutoff level is -- it would change if, for example, we used 80% as the cutoff instead of 70%.

(Note: there is a slight difference in the way the U-L diff value is calculated when a mastery level is used; students who do not answer an item, or have an invalid response, are excluded from the calculation for that item.)



Book5 [Compatibility Mode] - Microsoft Excel

Lertap5 U-L stats for "M.Nur Licensing E390v6.3", created: 27-May-13.

Res =	A	B	C	D	other	U-L diff.	B disc.
NM60 masters	0.00	0.01	0.92	0.06	0.00	0.80	0.23
others	0.05	0.08	0.70	0.16	0.01		

Summary group statistics

	n	avg.	avg%	s.d.	min.	mdn.	max.
masters	808	47.5	79%	3.9	42	47	60
others	961	34.8	58%	4.8	2	36	41
everyone	1,769	40.6	68%	7.7	2	41	60

This was an upper-lower analysis based on a mastery cutoff percentage of 70% (cut score = 42) .

Variance components

	df	SS	MS
Persons	1768	1751.56	0.99
Items	59	3340.98	56.63
Residual	104312	18124.56	0.17

Hoyt's reliability coefficient: 0.825

CSEM at the cut score: 3.288

Livingston's coefficient: 0.830

Index of dependability: 0.799
Estimated error variance: 0.003
For 68% conf. intrvl. use: 0.059

Prop. consistent placings: 0.810 (Estimated number of incorrect classifications: 336)
Prop. beyond chance: 0.617

Stats1f Stats1b csem1 Stats1ul

Ready 100%

The bottom of a Stats1ul report when a mastery level is in use includes the extra information seen here under the "Variance components" section. This information has to do with assessing how well a mastery test has worked. If a mastery test could be used twice to classify students as "masters" or "others", we would hope that the classification of each student would be the same on each testing, that there would be perfect "agreement" from one testing to the next. However, the measurement error commonly associated with our tests makes this highly unlikely.

Hoyt's reliability coefficient	This reliability estimate comes from the work of Hoyt (1941); he showed how test reliability could be estimated from a variance components table. It is well known that Hoyt's procedure produces the same result at that found by computing Cronbach's coefficient alpha. (Compare Hoyt's estimate with the

	"reliability (coefficient alpha)" figure found in the Stats1f report -- they'll be equal.)
CSEM at the cut score	The standard error of measurement at the cut score. The "C" in CSEM means "conditional" as the value of CSEM depends on, or is "conditioned on", test score. (In this case the cutoff percentage is 70%, equal to a cut score of 42. The value seen here is brought in from the appropriate row in the CSEM 2 column of the csem1 report.)
Livingston's coefficient	An "agreement index" based on a squared-error loss function. Berk (1980) regards this as *"providing meaningful information about the consistency of scores in relation to the cutting score"*. When the cut score is equal to the mean test score, Livingston's coefficient will equal the value of coefficient alpha (and Hoyt's coefficient).
Index of dependabilit y	This agreement index stems from the work of Brennan and Kane (1977). It is also based on a squared-error loss function; unlike Livingston's coefficient, this index takes into account item variance. If all items have the same difficulty it'll be equal to Livingston, otherwise it will be lower than Livingston's coefficient. Because of its incorporation of item variance, this index is sometimes preferred to Livingston's (see Crocker and Algina, 1986).
Estimated error variance	Derived from the work of Brennan and Kane (1977). They referred to this statistic as the "estimated error variance for making decisions about individual persons", specifically noting its relevance to mastery testing.
For 68% conf. intrvl. use	To develop a range which might be expected to capture a student's "universe score" 68% of the time, add and subtract this figure from a student's test score (with the test score expressed as a proportion of the maximum possible test score). "Conf. intrvl" means "confidence interval".
Prop. consistent placings	Berk (2000) regards this statistic, the estimated proportion of students who would be consistently classified, or placed, by using a test as a "Best Buy", referring to it as *"an unbiased estimate of decision consistency which is easy to compute, interpret, and explain".* (Berk implies it is generally much preferred to Livingston's coefficient, and to Brennan and Kane's index of dependability.)

	In the literature, this is referred to as the p_0 index. Lertap estimates p_0 by using a method suggested by Peng and Subkoviak (1980). The "Estimated number of incorrect classifications" output by Lertap, 336 in the example above, is an estimate of the number of students whose classification might change were we to test them again (some of the "masters" would become "others", while some of the "others" would move up to "masters"). In this case 336 = 1,769 multiplied by (1.000 - 0.810).
Prop. beyond chance	This is "kappa", a well-known estimate of the proportion of students who have been correctly classified above and beyond those who would be so classified by chance alone. Berk (2000) suggested that kappa *"is a biased estimate with a long list of limitations and statistical conditions that complicate its interpretation"*.

Note that it is possible for a test to have low reliability, but still have good classification consistency. Some examples of such tests are shown here.

Related tidbits:

NCCA (the National Commission for Certifying Agencies) allows for "decision consistency" estimates to be reported for a test, instead of a conventional reliability estimate (such as coefficient alpha or KR-20). Four of Lertap's statistics could be used as the decision consistency estimate: (1) Livingston's coefficient; (2) the "Index of dependability"; (3) the "Prop. consistent placings" (the p_0 index), and (4) the "Prop. beyond chance" (kappa). Which of these four should be reported for NCCA? The p_0 index might be the best bet, with, perhaps, the "Index of dependability" tagging along. How to use Lertap's output to complete an NCCA report? Got it covered. Where? Here.

Read more about these statistics in Chapters 7 and 10 of the manual, and (especially) in a not-overly-boring paper discussing the use of cutoff scores and many of the statistics seen in this topic, by clicking here.

The citations seen above may be followed here: Get those refs!.

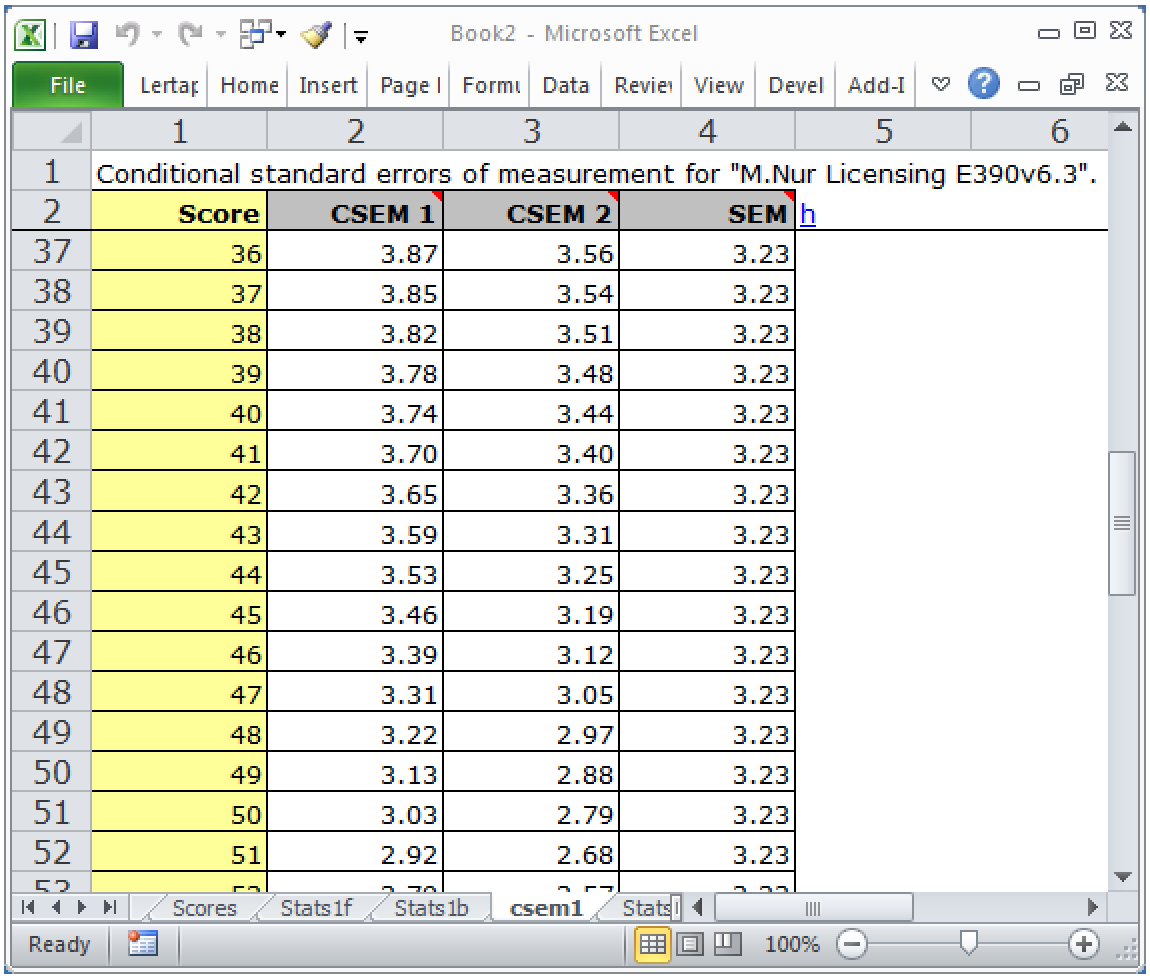
6.4.4 Conditional SEMs

Your read of the Lertap manual will have made you full-bottle on the usual standard error of measurement, SEM, and the invaluable role it has to play in the interpretation of test scores.

The SEM commonly used in CTT, classical test theory, is an average figure, one very frequently applied to each and every student's test score as an estimate of measurement error, no matter the level of the score. It is found in the "Summary statistics" sections of Stats1f reports.

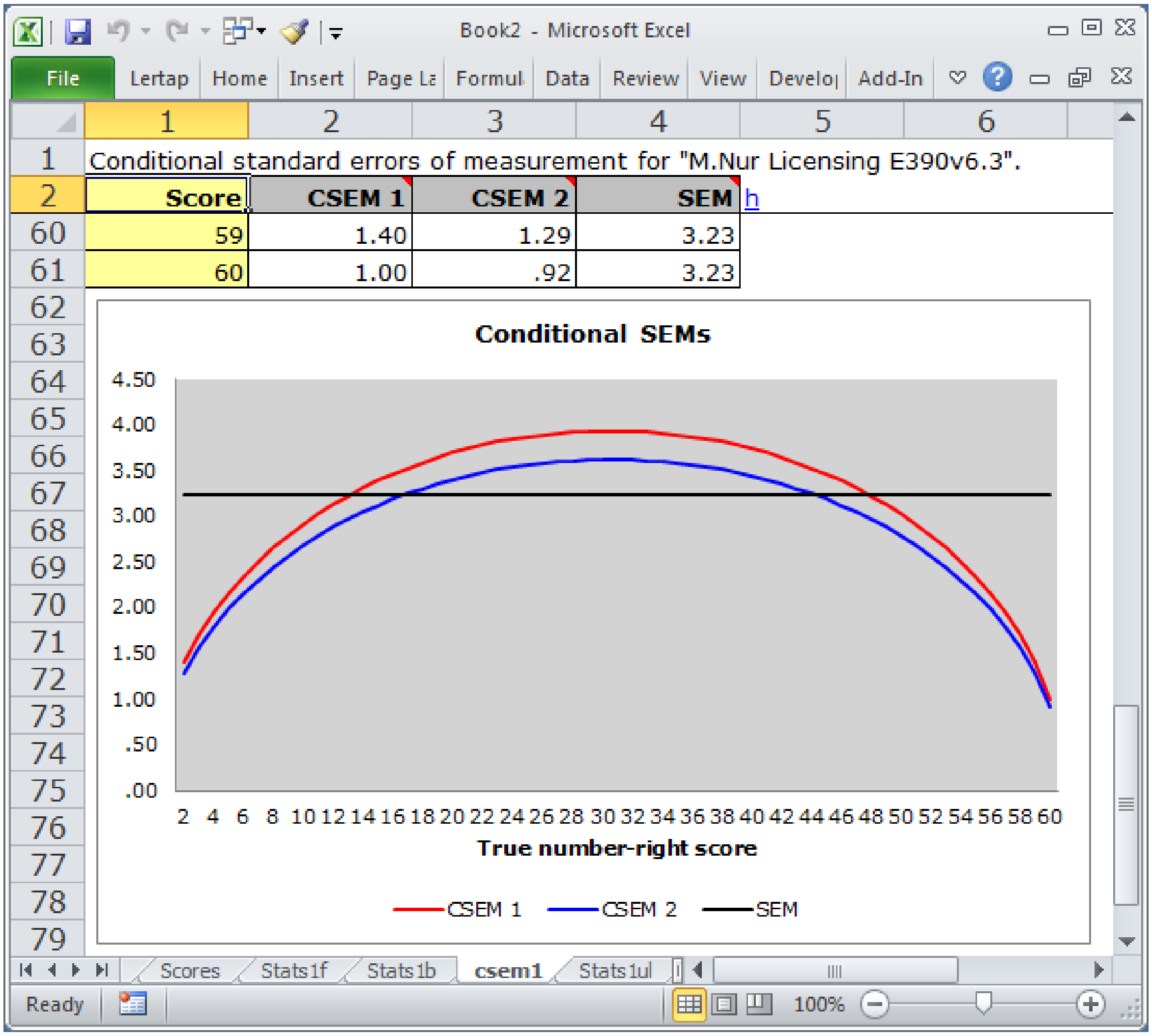
However, it has long been known that standard errors of measurement vary by ability. As we go from test takers with little proficiency, to those with average proficiency, and then on to the strongest students, respective SEMs change. To reflect this, Lertap also computes SEMs for various score levels using methods from a paper by Lord (1984). More exactly, Lertap employs Lord's Method III, the binomial error model, and also the adjustment to Method III estimates, known as Method IV.

A "csem1" Lertap report has two parts: a table followed by a plot. The information seen in the sample output below is from the "M.Nursing" dataset.



Conditional standard errors of measurement for "M.Nur Licensing E390v6.3".

Score	CSEM 1	CSEM 2	SEM
36	3.87	3.56	3.23
37	3.85	3.54	3.23
38	3.82	3.51	3.23
39	3.78	3.48	3.23
40	3.74	3.44	3.23
41	3.70	3.40	3.23
42	3.65	3.36	3.23
43	3.59	3.31	3.23
44	3.53	3.25	3.23
45	3.46	3.19	3.23
46	3.39	3.12	3.23
47	3.31	3.05	3.23
48	3.22	2.97	3.23
49	3.13	2.88	3.23
50	3.03	2.79	3.23
51	2.92	2.68	3.23



There are always three lines in the graphs. The flat one, the horizontal one, corresponds to the usual SEM value, as seen in Lertap's Statsf reports. The SEM value is a constant; it's the same for all test takers, no matter their ability.

The top line, a curve, corresponds to standard errors computed using the binomial error model, labeled as Method III in Lord (1984). The lower curve is what results when Method III estimates are adjusted using Method IV; Method IV estimates will be lower than Method III's whenever the items used in the subtest have different difficulties ("diff" values, to use the parlance of Lertap's Statsb reports).

So, what does the graph tell us?

In this case, results are from a 60-item mastery test with a cut score at 42 (equal to 70% of the maximum possible score of 60).

At the cut score, SEM was 3.23, as it was for all other levels. But the conditional standard errors of measurement at this point were greater: 3.36 using Method IV (CSEM2 in the graph), and 3.65 using Method III (CSEM1 in the graph); these CSEMs are a better reflection of the true state of affairs around the cut score of 42.

Related tidbits:

NCCA (the National Commission for Certifying Agencies) allows for reporting of the "standard error of measurement at the cut-score". Use CSEM2 for this, "Lord's Method IV".

Classical test theory, CTT, has frequently been criticized as failing to acknowledge that errors of measurement vary over score levels. This alleged "failure" is, in turn, often used to highlight an alleged advantage of item response theory, IRT. The criticism is largely misplaced; it is incorrect. CTT and IRT both provide for conditional measurement error estimates. Interestingly, IRT estimates of measurement error are *lowest* in the middle of the score distribution, rising to their highest values at the extremes. CTT estimates are the opposite; as illustrated in the chart above, CTT error estimates are *greatest* in the center. For more information, please refer to this paper.

There's a paper, another best seller, which has lots more about using cut scores, with several examples: www.lertap5.com/Documentation/JERM2007d.pdf.

6.4.5 Printing

There are two main sources of information which discuss getting Excel to print part or all of one of Lertap's summaries.

Click here to read about printing response charts, such as "quintiles" and "quantiles". Refer to this topic for comments on printing the information found in a variety of Lertap's statistical summaries.

6.5 Import & Export

Lertap is an Excel application. Its input and output "files" are standard Excel worksheets nested within a standard Excel workbook.

We have found that users will often have an Excel worksheet with data which they'd like to use with Lertap. If they rename the worksheet to Data, and add another sheet called CCs, will Lertap work?

Yes. And no. Lertap will certainly work, but its output may be poorly formatted and difficult to read. There are often font problems. Lertap has a preferred font: Verdana. If the user's workbook is based on a different font, Lertap's output may be adversely affected.

When such problems arise, we suggest this: use Lertap's New menu to make a new blank Lertap workbook. Then, copy all of the data records from the original workbook, and paste them into Lertap's Data worksheet. This generally works without problem.

We have seen numerous "text" files imported to Lertap without problem. For a fairly thorough example, we recommend a visit to this URL:

http://www.lertap5.com/Documentation/Samples/MondatY/ProcessingMondaty.htm

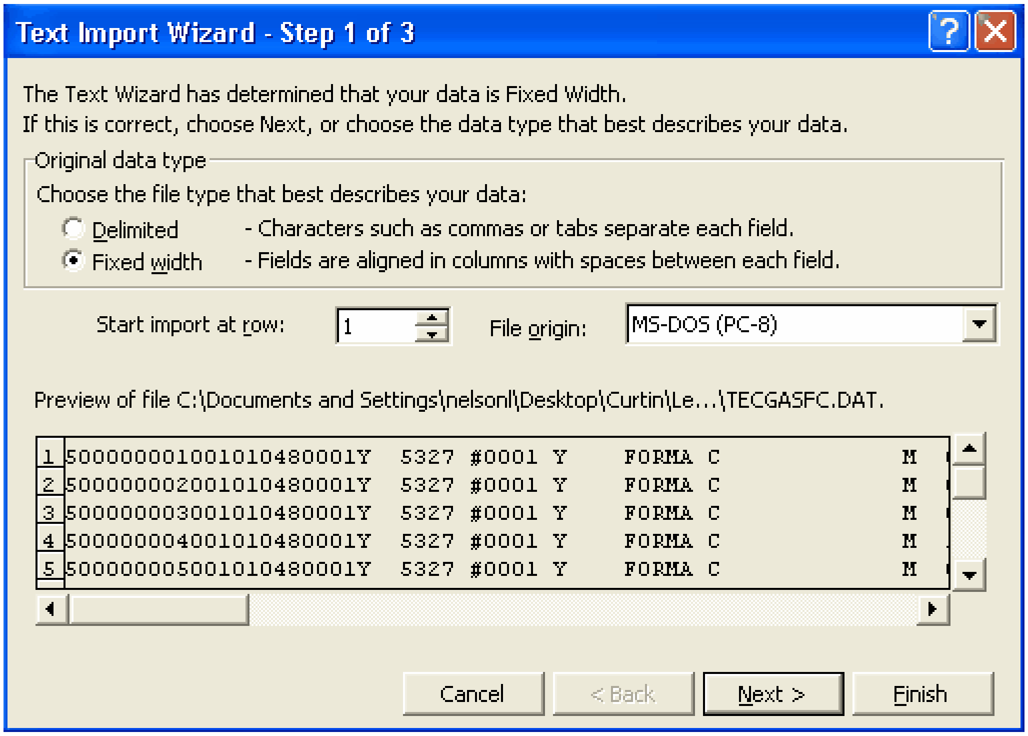
Back a few years, we happened to be in Central Java, Indonesia, where we saw some 20,000 provincial high school test results imported to Lertap from a "dBASE" file set up by a scanner. The dBASE file had three tab-delimited fields: record number, student ID code, and a string of 80 item responses. A straightforward copy-and-paste from the dBASE file to a Lertap DATA worksheet, followed by application of "The Spreader" in the Data worksheet's third column, quickly set up the data for Lertap processing.

Lertap has a special ability to import the text (DAT) files used by the ITEMAN program. This could well serve as a general means of importing data prepared by a scanner. Read about it in the next topic.

Are you aware of Excel's ability to dissect a text file? Excel has a "text import wizard", a useful tool which can be of real help when you've got to take apart a text file (the "DAT" files output from some scanners are usually simple text files).

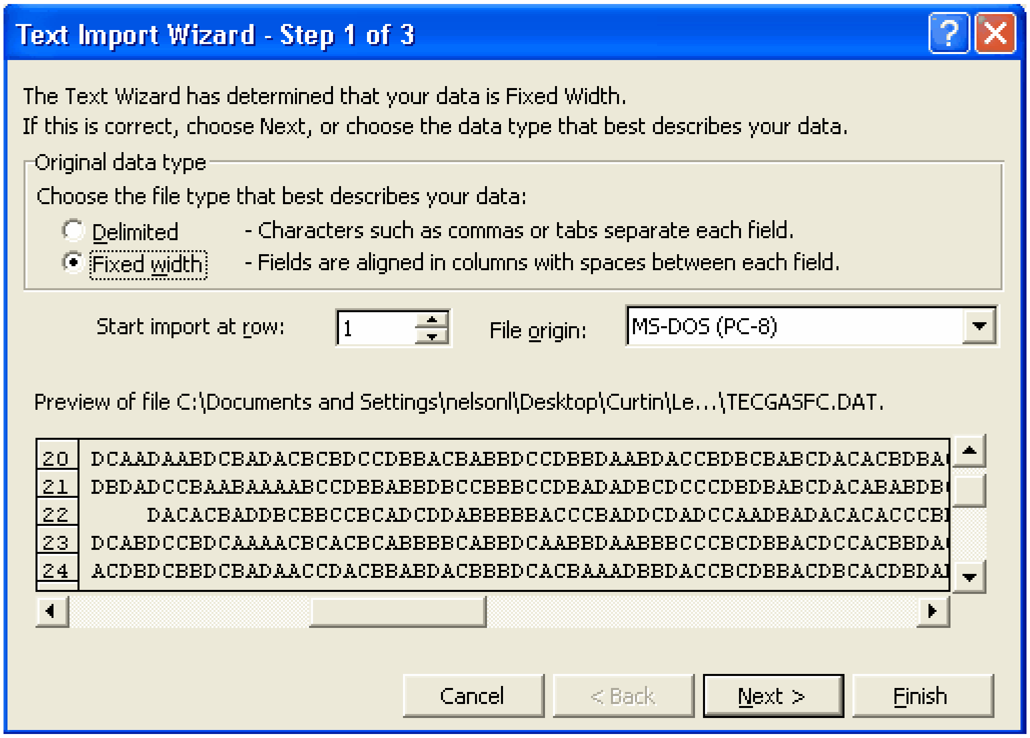
This wizard can be whizzed up in a couple of ways. One way is to use Excel's File / Open options, and under Files of type: ask for "All Files (*.*)", then browse to the file you've got in mind. Or, use Excel's Data / Import External Data, and follow the same steps. If the file you point to is purely text, Excel opens a dialog box which allows the file's contents to be cut out, field by field, and placed in the columns of a new Excel worksheet.

Here's a picture of Excel's text importer in action:



Doesn't the Excel Import Wizard look pretty useful? It is, but beware: it's got a problem. Yes. If there's a string of item responses to be imported, watch out. You have to make sure that the string has no blanks at the start.

Have a look at this screen shot:



Look at row 22 above. It has four blanks at the beginning, probably signifying unanswered questions. When the wizard is asked to import the string of responses, chances are real good it'll simply lop off those four blanks, shifting the string to the left. This is real bad -- this record's test score is going to be wrong.

How to control for this problem? Well, if the file has been created by a scanner, see if the scanner can't be coaxed into saving its data in an Excel-ready format, such as, perhaps, a "csv" file (comma-separated values). Such files come into Excel without having a need to be converted. Another useful format is the trusty old dBASE one mentioned above.

If you want to talk to us about this type of problem, just zip off an email to lertap5@gmail.com. We'll get back to you as soon as we're in from camping.

Note: the topics discussed here are also presented in the "samples" website.

While talking about strings of item responses, don't forget about "The Spreader". It's tailor-made to take strings of responses apart, and it loves to be put to work.

As to exporting Lertap worksheets, making them ready for use in another package: some packages, notably SPSS for Windows, readily work with the 'xls' and 'xlsx' workbook formats used by Excel. In SPSS 17, or later, use File / Open / Data, and select Files of Type: Excel.

Lertap comes ready to export to programs such as XCALIBRE and SCheck: see the discussion related to having Lertap make its own DAT file.

6.5.1 ITEMAN

ITEMAN is another classical item and test analysis system, created by David Weiss of Assessment Systems Corporation (ASC) way back when.

When, exactly? As it happens, about the same time as the first version of Lertap: late 1960s.

Iteman 3 is a system which many people used for years. A major update, inspired in part by Lertap, and in larger part by new staff at ASC, came out in late 2010. It's known as Iteman 4.

How does Iteman 4 compare to Lertap 5? See these papers: an overview, and a discussion of item flags.

Lertap 5 is able to import files made for Iteman 3, and it does so with real ease. This could be a useful and quick way to import data from a scanner. To use the Iteman importer in this manner, you'd first want to download the Iteman user manual from www.assess.com and get an understanding of the four control lines which Iteman wants ahead of the actual data.

Access to Lertap's Iteman importer is via the Macs menu. Will this importer also work with Iteman 4 files? Yes and no. Yes because Iteman 4 allows for two input modes: old (meaning Iteman 3), and new (Iteman 4). Iteman 4 users who have their input in Iteman 3 style can use the importer. Should you have an interest in using the new Iteman 4 input style with Lertap, please use let us know (larry@lertap.com).

Meanwhile, note that it is very easy to have Lertap create files suitable for input to Iteman 4. Should this be of interest to you, read this topic on Lertap and Xcalibre -- Iteman 4 and Xcalibre 4 share the same input files, so activating Lertap's Xcalibre export feature will result in files suitable for use by Iteman 4 as well as Xcalibre 4.

Related titbits:

A paper with more about ITEMAN and Lertap is here: ItemanAndLertap5.pdf (pdf file, about 1.5 MB).

A discussion of the correlation methods commonly found in item analysis programs, with emphasis on why Lertap's results sometimes appear at variance: ItemCriterionCorrelations1.doc (Word file, about 190 KB).

7 R&R&R&R

7.1 Revisions

The manual was printed 1 December 2000. Numerous changes have been made to Lertap since then, some minor, some quite substantial. Read about these changes by using the links below.

With the spawning of the Excel 2007 version of Lertap, released mid-2008, a frequently-revised compendium of the updates which apply specifically to the Excel 2007/2010/2013/2016/2019/365 versions has been developed. It's here:

http://www.lertap5.com/Documentation/UpdatesSummaryLertap57.pdf

A lengthy historical record of changes made before the emergence of Excel 2007 is presented in the following topics.

May 2003: added "Lelp" (version change to 5.2)

June 2003: added an item zapper

July 2003: changed item difficulty calculations

August 2003: what you weighted for (other=) (version change to 5.25)

September 2003: to halve and hold forever; Bilog-MG; tetrachorics

October 2003: eigenvalues & SMCs (version change to 5.3)

November 2003: smiles for quintiles!

February 2004: nothing sword-id about this (XCALIBRE support)

April 2004: consolidation (version change to 5.4)

July 2004: record IDs, & formula scores

September 2004: MDO now means: Missing Data Out!

October 2004: enhanced IStats report (version change to 5.4.5)

November 2004: *exc, a new CCs "card" to exclude items

February 2005: three enhancements (EMQs; MDO; quintile options)

May 2005: production mode added (version change to 5.4.6)

July 2005: beat the cheat? (version change to 5.5)

September 2005: histogram charts anyone?

January 2006: beat the cheat 2 (version change to 5.6)

March 2006: more MDO stuff, plus a did-not-see option

April 2006: let us Mac your day (macros R U!) (version change to 5.6.2)

June 2006: a new recoder, and analysis of variance

October 2006: conditional standard errors of measurement now computed (version change to 5.6.3)

7.1.1 Oct 2006 (Ver. 5.6.3)

Conditional standard errors of measurement are now estimated for cognitive subtests. Read about CSEMs with a click here.

A new supporting document related to the use of CSEMs, mastery tests, and cut scores is available as a Word document (about 300 KB).

A scatterplot of item difficulty by discrimination now appears at the bottom of every Statsb report. With a wee click here you'll see great some examples.

7.1.2 Jun 2006

June 2006: boys will be girls, and at variance?

Recodes, recodes, recodes: a new option on the Move+ Menu will be useful when you need to change boys into girls, cities into countries, and all such. A click here will show how.

The breakouts report has been enhanced; it now features an analysis of variance table at the end. Have a look.

7.1.3 Apr 2006 (Ver. 5.6.2)

April 2006: how about a Big Mac?

No doubt you've overheard people talking about the new macros they've made to customize their copy of Lertap?

It's true, too: you can now get Lertap to link to your own macros.

Read all about it: the Macs Menu.

7.1.4 Mar 2006

March 2006: adjustments to the MDO, and support for "did-not-see" cases

Pairwise exclusions now apply to the calculation of cognitive item correlations when MDO is used on the *sub card: all item-criterion correlations are now corrected for missing data.

A did-not-see option has been added to the System worksheet. This option works in a manner analogous to MDO: if a person has not been presented with the chance to answer an item (for whatever reason), the calculation of item statistics is adjusted accordingly. A gentle click here will let you see more.

Both of these adjustments will be of use when students see different test items. In some current online testing systems (for example), items are sampled from an item bank, with each student getting a subset of items. Not only do students see a sample of test items, but the items included in the sample will vary from student to student, effectively presenting each student with a different test.

7.1.5 Jan 2006 (Ver. 5.6)

January 2006: four enhancements (version change to 5.6)

While most were out celebrating the arrival of yet another new year, there was no rest at Lertap central. Some substantial changes have been made.

Response similarity analysis, RSA

Enhanced work in this area started July last year. Now Lertap produces three reports for looking at the matter of suspect cheating, with an important probability index, "Sigma", added to help put the heat on a cheat. To find out more you'd want your mouse to nibble here.

Lertap breaks down

What's this? A break down? Nope; just had to get your attention. There's now a new option on the Run menu which will break out results by groups.

Say people have sat your test on chemistry at five different campuses of your university. You have coded test venue into one of the columns on the Data sheet. The new option to "Breakout scores by groups" will quickly produce a "Breaks" table with test results organized by group levels, and a spiffy graph to match, "PlotBreaks".
[****]
Read more. *Plus*: have a look at growing whiskers immediately below.

Box your whiskers?

The Shorts menu also sports a new option: "Make box and whiskers from Breaks." Once you've got one of the spiffy new Breaks tables mentioned in the previous paragraph, this option will make a copy of the table, and reformat it so that it will suit one of Excel's built-in options for plotting the performance of your stocks.

Don't have any stocks? No matter. Excel doesn't really realize what it's plotting; the "stock performance" graph which results comes usefully close to being a real boxplot, and it certainly has whiskers.

Burma shave?

The histogrammer now uses improved grammer

The good old line-printer-compatible histogrammer from Lertap 2 days, a favorite of many (well, at least some), is now smarter, being capable of plotting the scores found on three different Lertap reports sheets: Scores, Breaks, and RSAsig. Have a squiz.

Use this with the Shorts menu option to "Make a histogram chart", an option introduced Sep 05, and you'll wow your audience for sure.

7.1.6 Sep 2005

September 2005: a new way to make histograms

Two options were added to the Shorts menu, making it possible to change the number of bars a histogram has, and enabling the creation of histograms without requiring the Analysis ToolPak Add-In. The new histograms, referred to as "histogram charts", are easier to modify.

Note inserted August 2007: the Shorts menu does not exist in the Excel 2007 version of Lertap. The histogram options referred to above are now found under the Histogram topic.

7.1.7 July 2005 (Ver. 5.5)

July 2005: response similarity analysis added (version change to 5.5)

Added support for those interested in investigating whether or not answer copying or sharing may have taken place during an examination.

Two new resources are available under the general rubric of "RSA", response similarity analysis.

As usual, to find out more you'd want to caress your little mouse, and ask it if it wouldn't mind clicking here.

7.1.8 May 2005

A production mode capability was added, making it possible to roll right through, non-stop, from the Run menu's two main options: "Interpret CCs lines", and "Elmillon item analysis".

Settings in this mode also make it possible to roll further, automatically getting histograms, response charts (such as quintile plots), and an item scores matrix (IStats), all without having to wear out your mouse with extra clicks on toolbar options.

Read more about it.

7.1.9 Feb 2005

February 2005: three enhancements

EMQs may now be processed with Lertap. EMQs are extended-matching questions, also known as EMIs (where the I means "item"). A test which makes use of EMQs will typically start out by presenting a series of options, up to 26 of them -- these turn out to be the same as the alternatives, or options, commonly used by MCQs (multiple-choice questions). What makes EMQs different from MCQs? Well, firstly, EMQs use many more options than MCQs. A typical MCQ will have what? Four options? Five? EMQs will have anywhere from 10 to 26.

Then, secondly, EMQ-using tests will have several items which use the same set of options. (This is why the set of options appears before the questions which use them.)

Who uses EMQs? They're popular with the National Board of Medical Examiners in the United States, and they are used in exams created by the United Kingdom's "PLAB", the Professional and Linguistic Assessments Board. We know that EMQs are also frequently used in the Caribbean, particularly in Trinidad and Tobago.

In practical terms, this means that Lertap's Res= declaration may now up to 26 entries; here's a little example, showing an Res= declaration which sets out 20 options:

*Sub Res=(A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q,R,S,T), Title=(EMQs!)

In order to accommodate the use of EMQs, we've changed the way Lertap's *alt card works. It used to be that the entries on the *alt card indicated how many of the Res= characters were used by an item; now the entries actually indicate which Res= character is the last one used by an item. (See example C7 under the Cognitive CCs topic, and remember: *alt is used only when some items use more options than others; if all items use the same number of options, *alt is not needed.)

The second enhancement? New options which control how quintile plots plot. You can have Excel automatically attach data tables to the plots (if you want), and you can tell Lertap to tell Excel that you only want certain items to be quintiled, not all of them. To read about this, you will want to click here.

And finally, the third enhancement concerns adjusting the difficulty index for cognitive items so that unanswered or omitted questions are omitted from the calculation of the index. If you would not like to read about this enhancement, ignore the temptation to click here.

7.1.10 Nov 2004

November 2004: *exc, a new CCs "card" to exclude items

Ways to quickly remove an item from its subtest is a matter discussed in a topic on its own; there's even a paper on the website which deals with the issue in somewhat extensive detail (a link to this paper is found at the end of the remove-an-item topic).

Now removers of items have a new tool: the *exc card, or line, which will probably be the easiest way yet to see that an item, or items, is/are quickly excluded from a subtest. Click here to read about this new method.

7.1.11 Oct 2004 (Ver. 5.4.5)

October 2004: enhanced IStats report (version change to 5.4.5)

More information has been packed into the "IStats" report. Its SMC values are now plotted in a series of ten bands, easing the task of determining the extent to which any single item relates to the others.

The first principal component, or the first principal factor, of the correlation matrix is now extracted, and item-component correlations are displayed in two ways: as a conventional row of values, and in the ten-bands format similar to that used for the SMC coefficients. A technical paper was added to the website to demonstrate how to interpret this new output. Read more about this in the eigenvalues topic.

7.1.12 Sep 2004

September 2004: MDO now means: Missing Data Out!

The power of the MDO option has been increased. Including MDO on an affective subtest's *sub line now gets Lertap to correct its brief stats report (such as Stats1b) for unequal response n's. Items with missing data may now have their statistics adjusted so that they're based only on valid item responses.

This is likely to be a handy revision for survey users. Read more about it by clicking here.

As part of this revision, we modified the CCs lines corresponding to the Lertap Quiz so that they show off the functioning of the new MDO whenever a user takes the cook's tour.

7.1.13 Jul 2004

July 2004: record IDs, & formula scores:

Prior to this revision, data records with ID information had to satisfy two criteria: the ID itself had to reside in either the first or second column of the Data worksheet, and the column header used for the ID field had to begin with the letters ID, or id, or Id, or iD.

Now the first of these requirements is gone. Zapped. The column with ID information may now be any column. This change, prompted by a request from Barbara Foster, University of Texas Southwestern Medical Center, will be welcomed by those who like to put the first item response in the Data worksheet's first column.

But, a warning: there's a potential problem with having the first item's responses recorded in the first column of the Data worksheet. What if someone doesn't answer an item? Some users let a blank, or empty, column represent missing responses. But a blank or empty first column in the Data worksheet has a very special meaning for Lertap: it indicates the end of data. Users who record item responses in column 1 of the Data worksheet should use a special code to cover the case of unanswered items -- for example, perhaps an "x", or maybe a "9".

For more about IDs, click here.

This revision also provides support for users who like to transform test scores, to re-scale them. Any formula may be applied to any score found in the Scores worksheet. Read all about it.

7.1.14 Apr 2004 (Ver. 5.4)

April 2004: consolidation (version change to 5.4)

We upped the version number to 5.4 for non-student users. This was done for two main reasons: we'd made a sufficient number of revisions to warrant a version number increase, and we installed a patch for an execution problem which previously dropped users into "Student mode", inserting an unwelcomed and unexpected new line in their Data worksheets in the process.

7.1.15 Feb 2004

February 2004: nothing sword-id about this:

Support for Bilog-MG users was added in September 2003. Now we've installed similar assistance for XCALIBRE users. XCALIBRE is an IRT program from ASC, makers of the well-known FastTEST item-archiving and test-generating system. (For more comments about FastTEST and Lertap, wiggle your mouse, and click here.)

About the same time, your favourite toolbar was enhanced. Yes. Part of it now sports Shorts. Check it out -- there's handy help for users who like to plot their output.

Note inserted August 2007: the Shorts mentioned above have not disappeared from the Excel 2007 version of Lertap, but they've found themselves dispersed. In particular, the line plotter option referred to is now found under the Basic options section of the Lertap tab.

7.1.16 Nov 2003

November 2003: smiles for quintiles!

This is a BIGGIE, a considerable enhancement to Lertap's processing of cognitive test items. The upper-lower groups analysis module has been expanded so that as many as 5 groups may be processed. Two new types of charts are now available, plotting item results in some remarkably revealing ways. This revision is discussed under the Graphics trio topic.

7.1.17 Oct 2003 (Ver. 5.3)

October 2003: hope you like roots (version change to 5.3)

Added support for latent-root (eigenvalue) and SMC calculations.

7.1.18 Sep 2003

September 2003: to halve and hold forever (?)

An ability to create random samples of data records was added as an option under the Run menu. This will be useful, if not to the whole world, then to those who might use Lertap as a precursor to some subsequent analyses, such as, perhaps, IRT modelling. To ignore this revision, do not click here.

Another option added in September: support for those who love tetrachoric correlations, and Bilog-like data files.

And, late in the month the way The Spreader operates was changed.

7.1.19 Aug 2003 (Ver. 5.2.5)

August 2003: what you weighted for (version change to 5.25)

Two accreditation professionals, one in Puerto Rico, one in Florida, asked for an ability to credit an item even when it was not answered. It was possible to do this before, but now it's easier: the advanced toolbar works better, and a new form of the *mws card has been introduced, one which allows an "other=" weight to be applied.

7.1.20 Jul 2003

July 2003: changed item difficulty calculations

When a journal reviewer suggested alterations to Lertap's procedure for indexing the difficulty of a cognitive item, we allowed our arms to be twisted, and followed his advice. We made it possible for a cognitive item's difficulty to be computed in one of three ways. Read all about it with a click here.

7.1.21 Jun 2003

June 2003: added an item zapper

Sometimes there's a need to quickly remove an item from a subtest without having to re-do a bunch of CCs "cards". Previous versions of Lertap allowed this to happen, and we updated Version 5.2 so that it would, too. A wee click here will explain what we did.

7.1.22 May 2003 (Ver. 5.2)

May 2003: added "Lelp" (version change to 5.2)

Lelp is, of course, Lertap Help. It's what your peepers are feasting on at this very moment. When we took the leap and installed Lelp, we changed the version number to 5.2 (from just "5").

7.2 Resources

There are other Lertap resources.

There's the Lertap manual, which this document has made frequent reference to.

There's the main Lertap website at Fremantle Village:

http://www.lertap5.com/lertap/

The Lertap website has screeds of additional information, including sample data sets useful in measurement classes, or by people just launching their Lertap careers. The website also has a modest series of technical papers highlighting examples of Lertap applications, and discussing current developments (such as the experimental options in Lertap).

A history of Lertap is provided in the manual, and on the website. Lertap's pedigree goes back to the early 1970s. (You may not realize it, but you could have used Lertap in the past when it was masquerading under another title.)

And then there's always our support desk in sunny West Australia. We welcome questions and comments. Write to us at: lertap5@gmail.com.

7.3 References

Please refer to the list of references found at the following URL:

http://www.lertap5.com/lertap/index.html?references.htm

7.4 Rchitect

Larry Nelson is not Lertap's only architect, but he's been the main one by far, having shepherded the system through several versions, and numerous host institutions.

Larry completed a BSc in electrical engineering at Wisconsin (1964); an MSc in Educational Psychology at Wisconsin (1970); and a PhD in Educational Psychology (psychometrics) at Colorado (1973).

He's held a number of academic and non-academic posts in the United States, New Zealand, Venezuela, Indonesia, Thailand, and Australia. All have had something or other to do with applied statistics, data analysis, test development, data banking, and computers.

The software development programs Larry has worked on include data analysis packages for use in Colorado (1972-1973) with NAEP; data banking routines for cross-country test results collected by IEA (1979-1982); item and test analysis software for the Venezuelan Ministry of Education (1971-1972); special-purpose computer-assisted learning programs for use by disadvantaged children and Spanish-language Learners (1973 and later, 1979-1984, New Zealand Education Department); demographic data processing and banking for use in Guayana (1974-1979), la Corporacion Venezolana de Guayana); applied cheating-detection procedures for use in Indonesia's Jawa Tengah province (2011 and 2016); and numerous versions of Lertap.

Lertap has been distributed, over the years, by Antipodes Software Systems (New Zealand and Australia, 1984-2000), Assessment Systems Corporation (2001-2014), Professional Testing Incorporated (2015-2016), and Lertap.com (2016-present). National Computer Systems purchased non-exclusive rights to use Lertap code in their products in 1984. Refer to this link for a more comprehensive history.

At February, 2021, Larry was said to be wearing these hats

Director
Lertap.com
Fremantle Village
South Fremantle, Western Australia

Research Fellow (adjunct)
School of Education
Curtin University
Perth, Western Australia

Professor (adjunct)
College of Research Methodology & Cognitive Science
Burapha University
Bangsaen, Chonburi
Thailand

Driver
Miss Angela's Touring, Camping, & Fishing Excursions
Used to be: justaboutanywhere, Australia
Now, with CV-19, just about no where

Click here for contact information (even has a picture).

Index

- # -

- * -

- ? -

- 2 -

- 3 -

- 8 -

- A -

- B -

- C -

- D -

- E -

- F -

- G -

- H -

- I -

- J -

- K -

- L -

- M -

- N -

- O -

- P -

- Q -

- R -

- S -

- T -

- U -

- V -

- W -

- X -

- Y -

- Z -