

NCCA Report Details and Lertap Output

Larry Nelson

Curtin University, Western Australia

Document date: 4 November 2013

website: www.larrynelsonstuff.com/lertap/

NCCA, the National Commission for Certifying Agencies, has an annual report form which, among numerous other items, requests detailed "Examination Data" for tests and exams used in accredited certification programs. As of the date above, a sample annual report was available [here](#) on the internet.

This document shows how to get the detailed exam data requested by NCCA from standard Lertap reports.

(The Lertap screen snapshots seen below were made by using the version of Lertap for Excel 2010 and Excel 2013.)

The following table and notes indicate the information requested in the Sample 2011 Annual Report Form used by NCCA.

1	2	3	4	5	6	7	8	9	10
Form name or number	Total # of candidates tested on this exam form in 2011	% of Candidates Passing in 2011	Passing Point ¹	Average Score ¹	Standard Deviation ¹	Standard Error of Measurement ¹	Decision Consistency Estimate ² (of P/F decisions)	Reliability Estimate ³ (of test scores)	Total Number of Items on Exam ⁴

¹Please indicate metric used e.g., scaled score, percent correct, raw scores. If scaled score is used, please provide definition. All four values should be reported on the same scale.

²For the decision consistency estimate, please provide the numeric value in the table and identify whether it is a decision consistency reliability estimate or the standard error of measurement at the cut-score.

³For the reliability estimate, please provide the numeric value in the table and identify the method

⁴If variable length, please provide range and maximum number of items.

The information for the six columns shaded in blue can be found on a Lertap Stats1f report, while data for the three yellow columns may be collected from a Stats1ul report.

Examples from Lertap's Stats1f and Stats1ul reports are shown on the following page.

Lertap5 full item stats for "M.Nur Licensing E390v6.3", created: 27-May-13.

Summary statistics

NCCA Col 2->	number of scores (n):	1,769
	lowest score found:	2.00 (3.3%)
	highest score found:	60.00 (100.0%)
	median:	41.00 (68.3%)
NCCA Col 5->	mean (or average):	40.61 (67.7%)
NCCA Col 6->	standard deviation:	7.71 (12.8%)
	standard deviation (as a sample):	7.71 (12.8%)
	variance (sample):	59.44
	number of subtest items:	60
	minimum possible score:	0.00
	maximum possible score:	60.00
NCCA Col 8->	reliability (coefficient alpha):	0.82
	index of reliability:	0.91
NCCA Col 7->	standard error of measurement:	3.23 (5.4%)

Stats1f

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

Stats1ul

NCCA Col 3 requests the percentage of candidates who passed. In the Stats1ul report, this may be calculated by using n (masters) divided by n (everyone) and expressing the result as a percentage figure. In the example above, this would be 808 divided by 1,769, or 46%. Note that Lertap will display this result if you let your mouse hover above the 808 figure – all the cells with a small red triangle in their upper corner have additional information about the cell's contents.

NCCA Col 4 wants to know what the passing score was. In the Stats1ul report, this is the “cut score”. In the example above, the cut score is 42.

Finally, NCCA Col 8 is for the “decision consistency estimate”. Lertap has no less than four of these: “Livingston's coefficient”, the “Index of dependability” (a Brennan-Kane estimate), the “Prop. Consistent placings” (the p_0 index in the literature), and the “Prop. beyond chance” (known in the literature as “kappa”).

Which of these four to use? There's a strong case to make for the p_0 index (that is, the “Prop. Consistent placings”), with the “Index of dependability” perhaps coming in as the second-best choice. For more information, you might refer to [this webpage](#) and to this [technical paper](#).

Note that NCCA Col 8 could contain “the standard error of measurement at the cut-score” instead (the option is yours, apparently). In the Stats1ul report, this is “CSEM at the cut score”.

Correction added 4 November 2013: I mentioned only five of the six blue columns. The answer for the 10th column way up above, “Total Number of Items on Exam”, is in the Stats1f report where it's shown as “number of subtest items”. In the Stats1f report sample above, the “number of subtest items” is 60.