INDIVIDUAL MISCLASSIFICATION in SSC REPORT

Cronbach, L.J., Bradburn, N.M & Horvitz, D.G. Sampling and Statistical Procedures used in the California Learning Assessment System. Report of the Superintendent's Select Committee July 1994 . Sacramento, CA: California State Department of Education.

Table 9. Risk of misclassification associated with various SEs

SE	0.8	0.7	0.6	0.5
Probability of misclassification (at least one-step)	0.56	0.51	0.45	0.39
Probability of two-step misclassification	0.08	0.05	0.02	0.01

SSC CALCULATIONS

Exact numerical calculations from SSC report are shown in Part B of Table 7. These (SSC/Cronbach) calculations can be thought of as a very special and limited case of the type of misclassification calculations Tables 1-3. The key point is that SSC/Cronbach's results are limited to ONLY the following (rough) proportions of kids in each of the 6 Performance Levels-- `(.05, .15, .3, .3, .15, .05)`
and that's not a distribution of student scores we see in many schools (i.e. 80% at level 3 or above rather rare).

Details on SSC (Cronbach) calculations. These calculations treat the 1 through 6 category labels for the six Performance Levels as a score range--actually we can say scores range from .5 to 6.5 with category divisions at (1.5, 2.5, 3.5, 4.5, 5.5). Now, let X be true score and Y be observed score; the SSC calculations use a bivariate Normal distribution for X and Y. Work with the conditional distribution of Y|X=c, where for example c = 3.3. Then the conditional distribution Y|X=c has mean c and variance = (standard error)^2. So for the calculations which are based on true performance level of 3 (range 2.5 to 3.5) a value of X between 2.5 and 3.5 is chosen. Ten to compute the Probability off-by-two-or-more from the conditional probability distribution (set to be normal), the probability under the curve exceeding 4.5 (i.e. PL5 and above) and that less than 1.5 (PL1) is computed (gullwing procedure; see G-theory materials). And, specifically, a standard that this error-- **Probability off-by-two-or-more-- not exceed .05** is pursued.

Aside: For the specific calculation for the range of category 3 of 2.5 to 3.5, choose X-values at every .05 and then average these numbers (a very poor man's integral). These numbers (also with finer increments .01 or .001) are given in Part B of Table 5. (My understanding is that LJC used the very very crude increment .2 with no weighting by the shape of the distribution).

Now what do these SSC computations tell you? It turns out these calculations tell you what my misclassification calculations tell you if you only look at results for category 3 (or equivalently 4) and use only the proportion in the 6 categories to be:
`{0.0512352, 0.155873, 0.292892, 0.292892, 0.155873, 0.0512352}`,
which is a distribution more peaked in the middle than used in Tables 1 and 4. (Of course, some small variation of these category proportions will also yield close results.) Misclassification matrix for this distribution given below. Parts A and B of Table 7 match perfectly for s.e. values of .6, .7, and .8 (esp. if you use the finer and more accurate .01 increment in Part B). (Probability off-by-one-or more numbers also match; see misclassification matrix).

How to make the equivalence between the two computations? A little detail. Start with a finite score range of .5 to 6.5. Take as the approximation to the Normal assumption, the triangular distribution. A symmetric triangular distribution over (.5, 6.5) has mean 3.5 and variance 1.5. So for my calculations take the X distribution as N(3.5, 1.5) (and thus Y as N(3.5, 1.5 + (s.e.)^2). (This N(3.5, 1.5) distribution leaves .007 probability outside each end of (.5, 6.5).) This X-distribution coupled with the set category boundaries of (1.5, 2.5, 3.5, 4.5, 5.5) creates the proportions in the 6 categories of {0.0512352, 0.155873, 0.292892, 0.292892, 0.155873, 0.0512352}, which is the approx. situation for which the SSC calculations are relevant. Score reliability is given through SSC standard error value and the equal width categories determine true proportions in each of the six categories .

TABLE 7 LJC Table

A. Rogosa Computations to obtain SSC numbers
X Normal with mean 3.5, variance 1.5 (without loss of generality)
True Performance Level Probabilities:
{0.0512352, 0.155873, 0.292892, 0.292892, 0.155873, 0.0512352}

SSC s.e.	Equivalent Reliability	Probability off-by-two-or-more						
		Overall	Conditional					
			PL1	PL2	PL3	PL4	PL5	PL6
.6	.806452	0.01998	.015	.0151	.02344	.02344	.0151	.015
.7	.753769	0.0395542	.02846	.02916	.04703	.04703	.02916	.02846
.8	.700935	0.0644374	.04479	.0465	.07742	.07742	.0465	.0448

B. LJC SSC Report-Style Computations for Category 3 (X = 2.5 to 3.5)
(Approximating an integral)

SSC s.e.	Probability off-by-two-or-more		
	Average over .05 increments	Average over .01 increments	Average over .001 increments
.6	.0224362	.023412	.0236328
.7	.045745	.0469975	
.8	.0760267	.0773911	

==

Some Standard Error--Reliability Translations

A translation between Reliability and the SSC Standard Errors.
Table 6 below provides such using the equivalences above. Remember the Cronbach probability numbers are based on the category proportions:
{0.0512352, 0.155873, 0.292892, 0.292892, 0.155873, 0.0512352}.

TABLE 8 Reliability--Standard Error Equivalences

Reliability	Equivalent SSC Standard Error (on .5, 6.5 scale) (using N(3.5, 1.5)
0.65	sig -> 0.898717
0.675	sig -> 0.849837
0.7	sig -> 0.801784
0.725	sig -> 0.754298
0.75	sig -> 0.707107
0.775	sig -> 0.659912
0.8	sig -> 0.612372
0.825	sig -> 0.564076
0.85	sig -> 0.514496
0.875	sig -> 0.46291
0.9	sig -> 0.408248
0.925	sig -> 0.348743
0.95	sig -> 0.280976

==

Misclassification Matrix for s.e. = .7 -> reliability = .7538

true cat	obs cat 1	2	3	4	5	6
1	.7265	.245	.0279	.0006	0	0
2	.2208	.4893	.2607	.0286	.0006	0
3	.0217	.2215	.4903	.2412	.0249	.0005
4	.0005	.0249	.2412	.4903	.2215	.0217
5	0	.0006	.0286	.2607	.4893	.2208
6	0	0	.0006	.0279	.245	.7265