An Alternative Method of Measuring Exceptional Growth

(and whatever we call the opposite of growth)

The current value added measure of teacher effectiveness (VAMOTE) being refined by CMS measures growth in terms of points above or below a predicted score. This system is seen as an improvement over previous measures (percent of students passing) in that it considers the effects of social factors and a student's previous performance to predict a score for the student's expected growth. The VAMOTE gives a measure of the difference from the expected growth by evaluating the student's actual score with the predicted score.

The acknowledged shortcoming of the VAMOTE is that it puts teachers with classes of students with high expected scores at a disadvantage compared to those teachers with classes of students with low expected scores. As an example take the standardized test scores of two students: Student A whose predicted to actual score is 40 to 55 has shown a 15 point growth. Whereas, student B whose predicted to actual score is 90 to 100 has shown only a 10 point growth. Note that in this situation, for teachers of high performing students, it is impossible for student B to exceed the growth of student A. Advantage goes to student A, not because of any teacher influence but simply because Student A can't score higher than 100 on the test.

To compensate for this disadvantage another approach would measure improvement based on possible growth. In this method student A whose score improved 15 points would be assigned an improvement of 15 points divided by the 60 points - the total points he could have improved – for a score of .25, or 25%. Student B however would be assigned an improvement of 10 points divided by the 10 – the total points he could have possibly improved for a score of 1.0 or 100%. Advantage goes to student B.

Now that we have 2 conflicting measures, the issue remains. Which is the more highly effective teacher?

Here is the formula written in pseudo-code I used to create the relative % increase or decrease column.

```
if predicted score is 0, then 0
   else if actual - pred  > 0, then (actual - pred)/(max score - actual)*100
      else if actual - pred = 0, then 0
         else (pred - actual)/(actual - min score)*100.
```

Here is a description of the same formula written in something closer to English (but not much)

If there is no predicted score make the improvement 0
If the actual is better than predicted, then divide the difference by the total number of possible improvement points (%).
If the actual = the predicted, then put 0.
If the actual is less than predicted (negative), then divide the difference by the total number of points it could have gone down (%).

Further Discussion

The extreme case above doesn't show the full picture. In the cases presented below the maximum score is 178 and the minimum score is 128.

In the first case each of the actual scores is 2 points less than the predicted score.

pred	actual	p-a	rel % incr or decr
174	172	-2	-4.34783
172	170	-2	-4.54545
170	168	-2	-4.7619
168	166	-2	-5
166	164	-2	-5.26316
164	162	-2	-5.55556
162	160	-2	-5.88235
160	158	-2	-6.25
158	156	-2	-6.66667
156	154	-2	-7.14286
154	152	-2	-7.69231
152	150	-2	-8.33333
150	148	-2	-9.09091
148	146	-2	-10
146	144	-2	-11.1111
144	142	-2	-12.5
142	140	-2	-14.2857
140	138	-2	-16.6667
138	136	-2	-20
136	134	-2	-25
134	132	-2	-33.3333
		0	0
		0	0
	avg	-2	-10.6395

max	min	number of students
178	128	21

In the second case, again the maximum score is 178 and the minimum score is 128, but now each of the actual scores is 2 points more than the predicted score.

pred	actual	p-a	rel % incr or decr
174	176	2	50
172	174	2	33.33333
170	172	2	25
168	170	2	20
166	168	2	16.66667
164	166	2	14.28571
162	164	2	12.5
160	162	2	11.11111
158	160	2	10
156	158	2	9.090909
154	156	2	8.333333
152	154	2	7.692308
150	152	2	7.142857
148	150	2	6.666667
146	148	2	6.25
144	146	2	5.882353
142	144	2	5.555556
140	142	2	5.263158
138	140	2	5
136	138	2	4.761905
134	136	2	4.545455
		0	0
		0	0
	avg	2	12.8134

max	min	number of students
178	128	21

In the third case, the relative increase is set at 50 %. Note that as the predicted scores are lower, the points needed to make a 50% gain increase.

pred	actual	p-a	rel % incr or decr
174	176	2	50
172	175	3	50
170	174	4	50
168	173	5	50
166	172	6	50
164	171	7	50
162	170	8	50
160	169	9	50
158	168	10	50
156	167	11	50
154	166	12	50
152	165	13	50
150	164	14	50
148	163	15	50
146	162	16	50
144	161	17	50
142	160	18	50
140	159	19	50
138	158	20	50
136	157	21	50
134	156	22	50
		0	0
		0	0
	avg	12	50

max	min	number of students
178	128	21

In the last case, the relative increase is set at -50 %, i.e. a 50% decrease. Here the predicted scores are lower, the points needed to make a 50% loss decrease.

pred	actual	p-a	rel % incr or decr
174	151	-23	-50
172	150	-22	-50
170	149	-21	-50
168	148	-20	-50
166	147	-19	-50
164	146	-18	-50
162	145	-17	-50
160	144	-16	-50
158	143	-15	-50
156	142	-14	-50
154	141	-13	-50
152	140	-12	-50
150	139	-11	-50
148	138	-10	-50
146	137	-9	-50
144	136	-8	-50
142	135	-7	-50
140	134	-6	-50
138	133	-5	-50
136	132	-4	-50
134	131	-3	-50
		0	0
		0	0
	avg	-13	-50

max	min	number of students
178	128	21

The case of the unwilling student

I have a co-worker who was complaining about some of his students who attended on exam day, but who refused to make a serious effort at taking the exam. I entered the all of his students' predicted scores and actual scores into my spreadsheet using the same format and calculations from the alternative method of measuring document. In the second spreadsheet I deleted the predicted scores for the 2 students who would not make an effort on the exam and the student who was taking the class for the third time (and was still predicted to pass by the EVAS system.) That causes their scores not to be counted in the teacher average at the bottom. I also marked some other details that the teacher thought were significant like the 2 LEP students who did not have predicted scores, but still passed the exam.

Note that the change when the students were removed was not dramatic, but it was significant. It made the average score go from 0.3636364 to 0.968254, or around 0.6 point difference. I am sure the line between the highly effective teachers who are rewarded and those who are not will be less than an average of 1 point on an exam.

I do not have a recommended solution for this issue. My initial reaction was to exclude all students who could be identified as not really "taking" the test. However, my fear is that there would be a teacher who would exploit that loophole by advising certain students who would be likely to score far below the predicted score to just guess "b" on every question.

with students who bubbled only b and 3peat

pred	actual	p-a	rel % incr or decr
149	141	-8	-32
150	151	1	3.5714286
158	161	3	15
144	139	-5	-25
150	146	-4	-15.384615
142	141	-1	-5.5555556
157	154	-3	-9.0909091
156	149	-7	-21.875
142	138	-4	-22.222222
154	134	-20	-66.666667
153	151	-2	-6.8965517
148	150	2	6.6666667
145	142	-3	-14.285714
157	165	8	38.095238
143	144	1	2.8571429
150	156	6	21.428571
150	156	6	21.428571
145	134	-11	-52.380952
146	150	4	12.5
147	152	5	16.129032
150	148	-2	-7.6923077
153	141	-12	-41.37931
150	145	-5	-19.230769
146	146	0	0
147	144	-3	-13.043478
147	145	-2	-8.6956522
156	156	0	0
149	143	-6	-24
144	136	-8	-40
149	149	0	0
144	149	5	14.705882
153	158	5	20
161	157	-4	-10.810811
147	151	4	12.903226
	163	0	0
154	146	-8	-26.666667
150	144	-6	-23.076923
	139	0	0
142	141	-1	-5.5555556
	150	0	0
142	144	2	5.5555556
143	150	7	20
147	147	0	0
	150	0	0
143	145	2	5.7142857
146	157	11	34.375
	154	0	0
156	164	8	36.363636
157	154	-3	-9.0909091
155	157	2	8.6956522
157	161	4	19.047619
155	163	8	34.782609
157	154	-3	-9.0909091
157	163	6	28.571429
153	156	3	12
163	166	3	20
157	162	5	23.809524
158	163	5	25
161	165	4	23.529412
154	152	-2	-6.6666667
156	160	4	18.181818
164	173	9	64.285714
164	178	14	100
159	162	3	15.789474
153	150	-3	-10.344828
156	158	2	9.0909091
164	166	2	14.285714
164	169	5	35.714286
149	150	1	3.4482759
164	158	-6	-15
162	168	6	37.5
	avg	0.3636364	3.6261166

max	min	number of students
178	124	66

without students who bubble only b and 3peat

pred	actual	p-a	rel % incr or decr	
149	141	-8	-32	
150	151	1	3.5714286	
158	161	3	15	
144	139	-5	-25	ld
150	146	-4	-15.384615	
142	141	-1	-5.5555556	
157	154	-3	-9.0909091	
156	149	-7	-21.875	
142	138	-4	-22.222222	
	134	0	0	20+ abs, 10 min
153	151	-2	-6.8965517	
148	150	2	6.6666667	
145	142	-3	-14.285714	
157	165	8	38.095238	ld
143	144	1	2.8571429	
150	156	6	21.428571	
150	156	6	21.428571	
	134	0	0	30+ abs, < 30 min
146	150	4	12.5	
147	152	5	16.129032	
150	148	-2	-7.6923077	
153	141	-12	-41.37931	
150	145	-5	-19.230769	
146	146	0	0	
147	144	-3	-13.043478	
147	145	-2	-8.6956522	
156	156	0	0	
149	143	-6	-24	hi
144	136	-8	-40	
149	149	0	0	
144	149	5	14.705882	
153	158	5	20	
161	157	-4	-10.810811	ag
147	151	4	12.903226	
	163	0	0	
154	146	-8	-26.666667	
	144	0	0	3peat
	139	0	0	lep
142	141	-1	-5.5555556	
	150	0	0	
142	144	2	5.5555556	lep
143	150	7	20	
147	147	0	0	
	150	0	0	lep
143	145	2	5.7142857	lep
146	157	11	34.375	
	154	0	0	lep
156	164	8	36.363636	Honors below
157	154	-3	-9.0909091	
155	157	2	8.6956522	
157	161	4	19.047619	
155	163	8	34.782609	
157	154	-3	-9.0909091	
157	163	6	28.571429	
153	156	3	12	
163	166	3	20	ag
157	162	5	23.809524	
158	163	5	25	
161	165	4	23.529412	
154	152	-2	-6.6666667	
156	160	4	18.181818	
164	173	9	64.285714	ag
164	178	14	100	
159	162	3	15.789474	
153	150	-3	-10.344828	
156	158	2	9.0909091	
164	166	2	14.285714	
164	169	5	35.714286	
149	150	1	3.4482759	
164	158	-6	-15	
162	168	6	37.5	ag
	avg	0.968254	6.054734	

max	min	number of students
178	124	63