

Developing a Value Table for Alaska's Public School Performance Incentive Program

Richard Hill
Revised August 1, 2006

Background

Alaska has shown considerable interest in describing the effectiveness of schools through the monitoring of changes in student performance over time. The Center for Assessment has proposed a system of "value tables" to measure the progress of students from one year to the next. In this system, a school is awarded a certain number of points depending on each student's performance level on any given year, when compared to that student's performance level the previous year. Alaska considered this a reasonable system, and during the 2004-05 school year, staff from the Department of Education and Early Development (EED), working with staff from the Center, developed a value table that might be used for Alaska. Details on this work are included in an earlier publication, "Establishing a Value Table for Alaska." (Hill, May, 2005)

In 2006, the Alaska legislature passed an omnibus education bill that contained a section providing for a public school performance incentive program. Under this law, up to 850 certificated employees of schools would be eligible for bonuses of up to \$5,500 (and an indeterminate number of noncertificated employees would be eligible for bonuses of up to \$2,500). The money was intended "to serve as an incentive for public school personnel to create a learning environment in which the students at that school demonstrate improved academic achievement more rapidly than would usually be expected, or, if already at an advanced level of achievement, continue to perform at this advanced level."

While EED felt that a value table system would be an appropriate method to identify the schools that should receive these rewards, the specific value table that had been developed in 2005 should not be used. The original value table had been developed for possible use with Alaska's No Child Left Behind accountability system, and therefore the categories of "Proficient" and "Advanced" had been combined into one (since NCLB requires that all outcomes of Proficient or higher to be equivalent). However, the Alaska legislation, by specific reference to students "at an advanced level," required that those two levels be considered distinct from each other. Also, the original value table was established by a group of educators serving on an on-going advisory committee. While it was appropriate to use this group to establish values for an NCLB accountability system, EED felt that a select committee should be used to determine the values that would be used for a system in response to legislation (and with considerable money attached).

In addition, as EED staff had presented the original value table to educators and legislators around the state, they learned that most people thought the category of "Proficient" should be divided into two categories so that students could show progress within that level. The lower two levels, "Far Below Proficient" and "Below Proficient," already had been divided into two categories in the original value table for that same reason. For all these reasons, EED felt that it should discard the original value table and develop a new one specifically designed to respond to the new legislation.

The Value Table Meeting

In order to receive recommendations on what the value table should be, EED hosted a meeting in Anchorage on June 7, 2006. The following people were invited, and all attended:

Carol Comeau, Superintendent, Anchorage Schools
Les Daenzer, Assistant Superintendent, Lower Kuskokwim Schools
Virginia Degnan, District Test Coordinator, Bering Strait Schools
Cydney Duffin, Principal, Colony High School, Mat-Su Borough Schools
Rhonda Gardner, Assistant Superintendent, Anchorage Schools
Jeanette Hayden, Teacher Mentor, Anchorage Schools
Michael Perkins, Principal, Mears Middle School, Anchorage Schools
Carl Rose, Executive Director, Alaska Association of School Boards
Roger Sampson, Commissioner, Department of Education & Early Development
Ann Shortt, Superintendent, Fairbanks North Star Borough Schools
Betty Walters, Superintendent, Kodiak Island Borough Schools
Denise Wilkinson, Principal, Polaris K-12 School, Anchorage Schools

Commissioner Sampson started the day by welcoming the attendees, introducing them, and providing some background on the legislation and the purpose of the day. Les Morse, Director of Assessments for EED, provided more specific background, including what the old value table looked like and the rationale for redoing it. Richard Hill, a consultant to EED from the Center for Assessment, then led the meeting from that point on.

We started by first explaining the overall assignment to the committee: ranking every cell within the value table, from the outcome that was most desirable (having a student go from Far Below Proficient Minus to Advanced) to least valued (having a student go from Advanced to Far Below Proficient Plus). In addition, they were to indicate whether the progression was a fairly steady one from cell to cell, or whether there were particular cells (or sets of cells) that were closer to each other, or further apart from each other, than the typical amount for cells.

Given our experience working with a committee in establishing the original value table, we knew that there were two issues that we needed to clarify for this committee before they started their deliberations. First, they were instructed to rank the cells in the table according to their desirability, not according to how difficult they might be to attain. For example, it might be more educational challenging to move a student from Far Below Minus to Far Below Plus, but most people (including the legislators who specifically cited the advanced level in the bill) would consider it more desirable to move a student from Proficient to Advanced—and therefore, that latter combination should be ranked higher than the former one. Second, the judgments about desirability might vary from grade to grade and from content area to content area, so we instructed the committee that when they felt that was the case, they should consider that they were constructing the table for students moving from grade 6 to grade 7 (the middle grades in the system) in reading (the content area that avoids the extremes of mathematics, where many more students score at the lowest levels, and writing, where few students score Advanced, especially at the higher grades). We also provided the committee with two givens that EED had already decided, based on conversations with policy makers around the state: first, maintaining status at higher levels was more desirable than maintaining status at lower levels (thus, for example, it is more desirable for a Proficient student to remain at Proficient than it is for a student who is Below Proficient to maintain that same status), and second, that whenever a student scored Far Below Proficient Minus in the second year, that cell would receive a value of zero,

regardless of what the student's performance had been the previous year. The committee expressed no concerns with either of those two givens.

Finally, before proceeding, we knew (again, from previous experience) that we need to address the issue of how much gain a student could reasonably make from one year to the next. That is, if a student progressed from Far Below Basic Minus to Advanced in one year, could that be a legitimate result, or would it just be a sign that something likely was misleading in the student's level one of the two years. If it were the latter, then we would need to establish some sort of maximum on the amount of improvement a student could make in one year. However, one committee member immediately responded that it would be quite possible for a student to legitimately go from the lowest level to the highest level in one year, and cited the progress she had observed in students who were just learning English. The committee expressed general agreement with that position, and as a result, we set no maximums on the amount of improvement we would expect a student to make legitimately.

The task presented to the committee was a substantial one—to rank the 42 combinations of performance levels (seven levels the first year by six levels the second) in order from most to least desirable. We suggested to the committee that the most efficient way to approach the problem was to first rank the cells in each of the diagonals and then to decide where the diagonals interweaved. So, for example, it would be easier to first rank all the cells in which students' performance was unchanged from one year to the next (indeed, EED had already done that for the committee—remaining at Advanced was more desirable than remaining at Proficient Plus, remaining at Proficient Plus was more desirable than remaining at Proficient, and so on), and then rank all the cells in which students' performance the second year was one level higher than it had been the previous year. Then, all one would need to do at that point would be to decide whether there was any point of intersection between the two diagonals; for example, decide whether it is more desirable for a student to remain at Advanced than it is to progress from Far Below Proficient Minus to Far Below Proficient Plus.

To assist the committee in making their decisions, we provided them with two sets of information before they started their assignment. We showed them the percentages of students performing at each performance level for each grade and content area. We also showed them cross-tabulations that provided the number and percentage of students at each performance level in 2006, given their performance level in 2005. We provided two such tables for each content area: one was the total result across all grades, and the second was the results for students going from grade 6 to grade 7.

At this point, we divided the committee of 12 into four groups of three people each and provided them with decks of cards representing the 42 cells in the value table. After the groups made their initial decisions about how to order the cards, each presented to its rationale to the others. After that, the groups reconsidered their decisions and reordered the cards as they saw appropriate, then recorded their judgments about the ordering of the cells on a form we provided. Once that was completed, the Commissioner made appropriate closing comments and the meeting was adjourned.

The Committee Judgments

The four groups all completed the assignment. The data from one group provided rankings that were inconsistent with reasonable values (for example, the group rated students moving from Proficient to Advanced as its most desirable outcome, while students moving from Below Proficient to Advanced was considerably less desirable), so their data were not included in the final averages. One group indicated there were two significant breaks in their rankings, and another indicated there was a small

range in which they felt all the cells were of equal value. Two of the groups indicated there having students move across the line to Proficient was a significant accomplishment, so that, for example, a student moving up one level to become Proficient should be more valued than moving up one level anywhere else. To summarize the data, it seemed simplest to compute initially the average rank the groups gave each of the cells. Those averages are reported in Table 1.

Table 1
Average Rankings (Rounded to the Nearest Whole Number)
Assigned to the Cells by the Three Groups

Previous Year Level	Current Year Level						
	Far Below Proficient Minus	Far Below Proficient Plus	Below Proficient Minus	Below Proficient Plus	Proficient	Proficient Plus	Advanced
Far Below Proficient Minus	N/A	23	19	13	4	2	1
Far Below Proficient Plus	N/A	30	22	16	6	5	3
Below Proficient Minus	N/A	33	28	20	13	8	7
Below Proficient Plus	N/A	36	31	27	17	12	11
Proficient	N/A	38	35	32	22	17	14
Proficient Plus	N/A	40	37	34	26	19	17
Advanced	N/A	42	39	38	26	23	18

Translating Panel Data into Value Table Entries

First, as noted in Table 1 above, we computed the average ranking for each cell in the matrix, rounding to the nearest whole number. The average ranks ranged from 1 (for the FBP- to Advanced cell) to 42 (for the Advanced to FBP+ cell). Next, to add meaning to the entries in the table, we assigned a score of 100 to the Proficient to Proficient cell. That is, when the question is asked, “What does a score of 100 mean?” the easy answer will be, “Maintaining a Proficient student at that same level from one year to the next.” The average panel ranking for the Proficient to Proficient cell was 22. From that point, it was straightforward to decide to increase Value Table entries by 5 points for every increase of panel ranking of 1 point, with a maximum of 200 and a minimum of 0. The translation generated the Value Table provided in Table 2.

Table 2

Possible Value Table for Alaska, Strictly Following Committee Rankings

Previous Year Level	Current Year Level						
	Far Below Proficient Minus	Far Below Proficient Plus	Below Proficient Minus	Below Proficient Plus	Proficient	Proficient Plus	Advanced
Far Below Proficient Minus	0	95	115	145	190	200	200
Far Below Proficient Plus	0	60	100	130	175	185	195
Below Proficient Minus	0	45	70	110	145	170	175
Below Proficient Plus	0	30	55	75	125	150	160
Proficient	0	20	35	50	100	125	140
Proficient Plus	0	10	25	40	85	115	125
Advanced	0	0	15	20	80	95	120

Suggested Revisions to Committee Recommendations

While the Value Table provides values that generally are internally consistent, a careful look at it suggested some minor revisions.

The first involves those students that drop 1 or 3 levels and end up in Current Year below Proficient. In the original table, students that end up at Below Proficient Minus after dropping 1 or 3 levels get more points than the students in adjacent levels. That is, going from Below Proficient Plus (BP+) to Below Proficient Minus (BP-) (a drop of one level) is assigned a value of 55 points, while going from BP- to FPB+ is 45 points and going from BP+ to BP- is 50. That doesn't happen when students maintain their level (a steady progression from 60 to 120 points) or drop 2 levels (a steady progression from 30 to 80 points).

Similarly, when students go up 3 or 4 levels, that earns a high value if the current year level becomes at least Proficient (which is a good thing, since that's consistent with committee recommendations), but then declining the higher you go above Proficient (again, that is a logical pattern). So, for example, if students go up 3 levels and wind up at BP+, that's worth 145 points, but if they go up 3 levels and wind up at Proficient, that's worth 175. Then the reward declines to 170 and 160 points for moving up 3 levels to Proficient + and Advanced, respectively. That all forms a coherent, justifiable pattern. If a student goes up 2 levels and winds up at BP- or BP+, that's worth 115 and 130 points, respectively, but it's worth 145 points if the student ends up at Proficient. Again, that is all consistent with the diagonal representing going up 3 levels. But now, if the student goes up 2 levels to Proficient +, that's worth 150 points, or 5 more than going up 2 levels to Proficient. To be consistent with the other diagonals, that should be worth no more than (and probably less than) going up 2 levels to Proficient.

In addition to these issues, there is some question about the values assigned when students remain at the same level from one year to the next and are below Proficient. Consistent with the committee’s judgments, students should receive fewer points for remaining at level when they are below Proficient than when they are Proficient or above. In addition, there should be some incremental value for moving closer to Proficient. Thus, for example, look at the row for students who were Below Proficient Plus the previous year. The increment between Below Proficient Minus and Below Proficient Plus should be smaller than the increment between Below Proficient Plus and Proficient—and it is. But the gap in Table 2 is 25 points for the first change, and 50 points for the second. While the second gap should be larger than the first, perhaps they should not be so disparate. The same minor adjustments could also be applied to the Far Below Proficient Plus and Below Proficient Minus categories.

As a result, we offer the following recommended Value Table in Table 3. It is highly similar to the one in Table 2, but eliminates the minor anomalies noted above.

Table 3

Proposed Value Table for Alaska, After Making Minor Alternations

Previous Year Level	Current Year Level						
	Far Below Proficient Minus	Far Below Proficient Plus	Below Proficient Minus	Below Proficient Plus	Proficient	Proficient Plus	Advanced
Far Below Proficient Minus	0	95	115	145	190	200	200
Far Below Proficient Plus	0	70	100	130	175	185	195
Below Proficient Minus	0	45	75	110	150	170	175
Below Proficient Plus	0	30	50	85	125	145	160
Proficient	0	20	35	50	100	125	140
Proficient Plus	0	10	20	40	85	115	125
Advanced	0	0	15	20	80	95	120

Results

Applying this value table to the scores obtained by students in 2005 and 2006, the average growth score attained by schools was 97, with a standard deviation of 12. The correlation of school mean growth scores with 2005 scaled scores was 0.76. That is an important statistic, and reflects a balance of several issues. First, one of the primary reasons that EED wanted to have a system of measuring student change was that it is well established that higher socio-economic schools tend to have higher scoring students, regardless of the quality of instruction they are receiving. As a result, schools that serve students from the wealthier suburban areas almost always outscore those that serve students in the inner city or the remote regions. This leads school personnel to get frustrated with accountability

systems that look at schools' current status only; the schools that serve the wealthiest students have a very high likelihood of being among the highest scoring. Looking at student progress from year to year will give more "have not" schools an opportunity to score well. On the other hand, the most experienced and most qualified teachers tend to work in the schools that serve the wealthiest students, so those schools also should tend to score highest in a system that measures student growth—it's just that the trend should not be as strong for a growth score as it is for a status measure. Also, this value table was intentionally tipped in favor of higher scores for students who are already proficient—and those students are disproportionately in the schools with higher status scores. So it is not surprising that the correlation between starting scores and growth scores is as high as it is; note that it is lower than the typical correlation between socioeconomic status and test scores.

Indeed, the correlation of school mean growth scores with the percentage of students not identified as coming from low-income families was 0.51, and the correlation of growth scores with the percentage of students in the school who are white was 0.64. These modest correlations tell us that there certainly will be a trend for the "have" schools to outscore the "have nots" in this system, but that there will be sufficient number of the latter schools that also score well.

There were 51 schools that had scores of 110 or greater, using the value table recommended in Table 3. That is slightly more than 10 percent of the 483 schools for which we were able to calculate a growth score. Eighteen of those high scoring schools tested 100 or more students; the median was 49. Eleven of them contained less than 50 percent white students; the median was 77 percent. In 11 of the schools, more than 50 percent of the students were from low-income families; the median was 25 percent. In contrast, of the 51 schools that had the highest *status* scores in the state, 26 had 100 or more students, and just 7 had less than 50 percent white students and in only 4 were more than 50 percent of the students from low-income families. So the schools that score highest on growth tend to be smaller and have a higher percentage of non-white and poor students than those schools that score highest on status.