

Preliminary Report to

Wyoming State Board of Education

***Results of the State Professional  
Judgment Panel Meetings &  
Recommendations on the State School  
Accountability Program***

by

Michael D. Beck

Casper, Wyoming

2-4 October 2012

# ***Results of the State Professional Judgment Panel Meetings & Recommendations on the State School Accountability Program***

## **Abstract**

The Wyoming Professional Judgment Panel (PJP) met for three days in early October in Casper, Wyoming. The meeting had four objectives:

1. To recommend school standards for four Performance Indicators identified as critical components of school accountability;
2. To use these established Performance Indicator standards in a “body of work” approach to set recommended standards for all Wyoming school buildings;
3. To discuss and, as possible, suggest “business rules” to be used in analyzing the component data for the Performance Indicators and the overall school accountability system; and,
4. To review draft plans for an appeals process to be implemented in conjunction with the school accountability system.

At the completion of the meetings, the PJP had met each of the above objectives. They carried out the process of setting school performance standards for each of the Performance Indicators (PIs), used these results to recommend school performance levels, and reviewed and commented on both system business rules and a draft of the system appeals process.

Recommendations of the PJP for both the PIs and the School Performance Levels have been submitted to the Wyoming Department of Education (WDE) so that WDE can implement these cutscores in their statewide database of school-level PIs. Upon WDE’s completion of these analyses and related quality-control efforts, the State Board of Education will have a school-by-school listing of the accountability levels recommended by the PJP for this pilot year.

## **Data Preparation**

All school-based data for the sessions was prepared by the Wyoming Department of Education (WDE) staff based on requests by Michael Beck, the consultant who was charged with facilitating the PJP session. Due to the tight time constraints for planning and carrying out the data analyses and delays in the processing of the required data, several generally minor inconsistencies and irregularities remained in the data at the time of the PJP sessions. These will be clarified and corrected, as necessary, during the process of generating the pilot accountability ratings based on the PJP’s work. In the case of the Equity PI, the facilitator made adjustments in the data provided by WDE to fine-tune these calculations; these adjustments will be conveyed to WDE prior to their generation of the pilot school rankings.

## Meeting Activities

The PJP sessions were held on 2-4 October 2012 at the University of Wyoming Outreach Center in Casper. Sessions were facilitated by Michael Beck and Sheila Potter, consultants engaged by the Wyoming State Board of Education to plan and conduct the PJP activities. A total of 26 PJP members participated in the sessions (24 were present for all three days, one for only Days 1 and 2, and 1 for only Day 3). Several observers attended the sessions, although they did not participate actively in either the discussions or committee judgment processes.

Appendix A presents an outline of the agenda for the three-day session; this agenda was followed closely, although several specific discussions were conducted at somewhat different points in the proceedings. Most particular among the variations from the original agenda was the inclusion, at PJP request, of an opportunity for a third round of committee judgments for the cutscores for the PIs.

The facilitators have conducted similar committee sessions in over a score of states for comparable purposes. They considered the PJP participation to be frank, open, and highly engaged. An effort was made to solicit views from all PJP members, to encourage interaction among the panelists, and to welcome disparate viewpoints. Opportunities were provided throughout for PJP members to seek clarification of unclear information or concepts. It is the opinion of the facilitators that panelists understood their task clearly and were very engaged in the task.

## General Methodology & Procedures

PJP procedures were essentially a “standard-setting” process in which panelists attempt to determine a reasonable minimum “cutscore” that describes achievement of pre-established level of performance. In the case of the PJP, the process was a two-stage activity. First, panelists determined levels of accomplishment for each of the PIs. Then, these PIs were considered in a “body of work” matrix in which panelists determined appropriate School Performance Levels for each possible combination of PI accomplishment.

Specifically, four PIs formed the foundation of the PJP deliberations:

***ACHIEVEMENT:*** *defined as the percent of students across grades and assessed content areas whose PAWS and PAWS-ALT scores were at the Proficient level (or higher).* This index was composed of statewide test scores for Grades 3 through 8 and 11 for the content areas of Reading, Mathematics, and Science. Reading and Mathematics tests are administered at each of the above grades; Science tests are administered only at Grades 4, 8, and 11. Grade 11 PAWS and PAWS-ALT scores were included under this PI although this grade will be dropped from the state’s testing program in the 2012-13 school year. Additional discussion of this decision will be provided in the final report of this project; however, fundamentally the decision was

based on the desire to include the Achievement PI for high schools and the facilitators' view that ACT scores are more appropriately considered a "college and career readiness" metric than an achievement metric.

**EQUITY:** *defined as the percent by which the proportion of students in the PAWS/PAWS-ALT performance level of Below Basic was reduced between 2010-11 and 2011-12.* That is, by what percent was the school able to reduce the proportion of students whose PAWS scores were Below Basic? This metric can, of course, be negative – indicating that the school had a higher proportion of such students in the most-recent year compared with the prior year.

**COLLEGE & CAREER READINESS (or READINESS):** This PI, which applied only to schools housing high-school students, was composed of two metrics: (1) *the percent of students who received a high school diploma within four years of entry into the school, and (2) the percent of students whose composite scores on the ACT EXPLORE and ACT Grade 11 assessments met the ACT-determined College Readiness Benchmark.* This latter index was based on the students' scores across the four assessed ACT content areas – English, mathematics, reading, and science. The ACT College Readiness Benchmarks, explained elsewhere, are empirically determined probabilities that a student will be able to do acceptable college-level work.

**GROWTH:** defined as the "median growth percentile" in Reading and Mathematics across all students in the school. The arithmetical calculation of this index is discussed in detail in the report of the WDE consultant for determining this metric (see Domingue, 2012). However, essentially the *Growth PI is determined by calculating each student's change in achievement from year to year and comparing this change/growth to that shown by all other students in the state whose initial achievement was at the same level.* If a student "grew" from year to year at the exact same rate as did others whose pretest score was at the same level, he/she would receive a Growth Percentile of 50. Students who grew at a rate less than that of their pretest peers would receive Percentiles below 50; those whose Reading or Mathematics growth rate exceed those who began at the same level would receive Growth Percentiles above 50. These individual-student percentiles are then arrayed across all students in the school, and the school receives an overall Median Growth Percentile. This is the school's Growth PI.

So, for example, a student may receive a Growth Percentile of 75 in 2011-2012 in Reading, indicating the student is at the 75th percentile of performance when compared with all other Wyoming students in the same comparison group based on prior years' performance. Changes in each student's Growth Percentile over time are one indication of how well the student is performing when compared with students across the state who have a similar achievement profile.

The Median Growth Percentile (MGP) is the indicator of how well each *school* is performing based on how well students in the school are achieving on the PAWS Reading and Mathematics tests.

Note that all PI metrics were defined on a *school level*. That is, the PIs were calculated across all grades served by the school. For example, a school building serving students in Grades 3-5 would have an Achievement PI composed of PAWS and PAWS-ALT data for Reading and Mathematics in Grades 3, 4, and 5 and Science data for Grade 4. The school accountability ratings, based on the School Performance Levels determined by the PJP, were grounded on all of the schools in the state, regardless of grade composition. It is important to note, however, that because of the nature of the PIs, some schools do not have data for each PI. For example, a school housing students at Grades K-6 will not have a CCR/Readiness PI, as data for this PI are available only for high schools. Similarly, a school housing Grades 9-12 will not have a Growth PI, as only one year of PAWS data are available for such schools, making a “growth” determination impossible.

PJP panelists individually made three independent rounds of judgments of the minimum cutscores for Meets the Target and Exceeds the Target for each of the 5 Indicators (4 PIs, with CCR having 2 parts). Following the first and second rounds of judgments, panelists were shown (anonymously) the comparable judgments made by their fellow panelists and extensive interaction among the panelists was encouraged. Panelists were encouraged to offer comments on why they made their judgments, ask questions about the process or the underlying data, and explore the implications of their recommendations. “Impact data,” describing the general statewide implications of the interim judgments were also provided following each round of judgments for each PI.

The initial plan for the meetings was for the PJP to carry out two rounds of these judgments, with the final round serving as their recommendation (not averaging or combining the two rounds). However, following extensive discussions of the second round of judgments and the statewide implications of these recommendations, several members of the panel requested the opportunity to conduct an additional round of judgments. No panelist disagreed with this request, so a third round was carried out. The majority of PJP members made no changes between their 2<sup>nd</sup> and 3<sup>rd</sup>/final round of recommendations, although several members changed their judgments for one or more of the PIs.

### **PJP Results & Recommendations for the Performance Indicators (PIs)**

Appendix B summarizes the three rounds of judgments in terms of the median recommendation by PI. (Only medians are presented in this Preliminary Report; other summary data, along with anonymous individual panelist recommendations by round, will be presented in the final report.)

As Appendix B illustrates, the median PI recommendations changed across rounds of ratings. However, these changes were generally small. Over rounds, there was a slight tendency for PJP median recommended cutscores to increase, both for Meets the Target and Exceeds the Target.

To summarize the PI judgments of the PJP, they recommended:

- for **Achievement**, that schools must have **75%** of their students at or above Proficient on PAWS or PAWS-ALT in order to be classified as Meeting the Target, and **87%** at or above Proficient to be classified as Exceeds the Target;
- for **Equity**, that schools must have reduced their proportion of Below Basic students between 2011 and 2012 by **6%** to Meet the Target, and by **20%** to be classified as Exceeding the Target;
- for **CCR**, that schools must graduate at least **85%** of their high school students within four years OR must have at least **20%** of their students achieve the ACT Composite Benchmark to Meet the Target, and must attain BOTH of these criteria to be classified as Exceeding the Target;
- for **Growth**, that schools must attain a median growth percentile of **50** to Meet the Target, and a median growth percentile of **60** to Exceed the Target.

### **PJP Results & Recommendations for the School Performance Levels**

Panelists had the opportunity to discuss the final PI judgments prior to making their School Performance Level recommendations. As with the PIs, multiple rounds of judgments were involved in the process. In advance of making the School Performance Level judgments, the PJP discussed the operational meanings they would attach to the ratings – Exceeding Expectations, Meeting Expectations, Partially Meeting Expectations, and Not meeting Expectations.

The activity of establishing standards for the School Performance Levels was essentially a “body of work” standard setting process, in which each panelist independently judged how a school with each possible combination of PI results should be classified. Panelists clearly understood that the process was an iterative one on which any individual panelist may disagree with his/her fellow judges as to the appropriate School Performance Level. Panelists independently made a determination for each possible PI outcome combination, with a total of 135 separate judgments being made. These initial judgments were summarized and presented to the panel, and selected combinations of PI outcomes on which there was significant panel divergence of opinion were discussed in the group. Following discussion, panelists made a revised, final set of recommendations. The median of these judgments is illustrated in Appendix C.

In Appendix C, it should be noted that the significant majority of Wyoming school buildings will have their accountability level determined by 3 of the 4 PIs. These include the 200+ buildings that do not house high school students; thus, no CCR data are available for these schools. Similarly, another 60+ schools will have their performance levels determined by a different 3 of the 4 PIs; in those cases – schools that house only Grades 9-12 – no Growth PI data are available as the Growth metric applies only to Grades 3 through 8. For only a small number (approximately 20) of Wyoming schools

will data for all four of the PI metrics be available. As noted in earlier WDE reports, schools that house only students below Grade 3 will have no School Performance Levels, as no statewide assessment data are available before Grade 3.

Note that for very few of the 135 PI combinations did all members of the PJP reach agreement on the most appropriate School Performance Level. This is both expected from, and a significant advantage of, a judgment-based standard-setting activity. Panelists individually, independently, and anonymously make determinations of the School Performance Level resulting from each combination of PIs. They use for these judgments a shared understanding of the PIs, the standards for those PIs, and the meaning of the School Performance Level labels. Panelists clearly understood that the “PJP Recommendation” would be the *median* of their judgments; such data are presented in Appendix C.

While there was certainly no consensus in the judgments made about most of the possible PI combinations, it is important to note that for just under one-half of the judgments (63 of 135), 20 or more of the 25 panelists agreed on a *single* School Performance Level; in an additional 23 cases, there was in excess of 70% exact agreement (18 or more of the 25 judges) on the School Performance Level. These data are taken as a clear indication of panel convergence as to the most-appropriate classification of schools based on their attainment of the PIs.

## **Facilitator Evaluation**

The facilitators have conducted similar sessions to these in over 20 states. We judge the quality of a session primarily on the basis of two elements:

1. Did the panelists have the opportunity to understand the process being used, to interact with their fellow judges in an open and honest manner, and to express their judgments freely and in an informed manner? and
2. Do the results of the process “make sense,” both internally and in terms of the constructs being judged?

In terms of both of these qualitative elements, the facilitators believe that the PJP provided the reflective, informed, knowledgeable, and candid reaction to the elements of the school accountability program that was desired at the time the committee was constituted. We are comfortable with the discussions that took place and with the recommendations that resulted from these deliberations. We recommend that the pilot year recommendations of the PJP be accepted and that a similar model of setting the school accountability standards be used when the operational system is established.

## **Other PJP Session Activities**

In addition to carrying out the standard setting process outlined above, PJP members reviewed and reacted to a draft plan for an appeals process for the school accountability

program. This draft plan was prepared by Paige Fenton Hughes, who facilitated the discussion of the draft appeals process.

Finally, panelists had an abbreviated opportunity to react to and suggest additional or revised elements of the “business rules” that guided their sessions. Additional information concerning these rules and the discussions of same will be presented in the Final Report of the PJP sessions.

MDB – 10/8/12

## APPENDIX A

# AGENDA

## *Wyoming Professional Judgment Panel Meetings*

Casper, WY --- 2 - 4 October 2012

### DAY 1 (8:30 – 4:00)

- A.M.**
- Orientation / Goals / Responsibilities
  - What is “standard setting”? How this relates to PJP
  - Overview of the 4 Performance Indicators (PIs)
    - Achievement, Equity, Readiness, and Growth
    - How each PI is operationally defined
    - How these lead to a School Accountability System
  - Giving meaning to the Terminology of the standards
    - for PAWS, for the PIs, for the Accountability System
  - Methodology used for setting standards for each PI
- P.M.**
- Initial recommended standards for Achievement and Equity
  - Initial recommended standards for Growth and Readiness

### DAY 2 (8:30 – 4:00)

- A.M.**
- Review of Day 1 activities
  - Summary of Initial Recommendations; state impact
  - Discussion of Initial Recommendation among PJP
- P.M.**
- *Final* recommendations for each PI

- Transforming PI standards into School Accountability
- Components of the system
- Melding PI standards into Accountability – “matrix”
- Making the Accountability Descriptors more concrete

### DAY 3 (8:00 – 4:00)

- A.M.**
- Review of final judgments of the panel on each PI
  - Making School Performance Level judgments –  
methodology & mechanics
  - PJP Makes Initial Accountability Ratings for Schools
- P.M.**
- Summary / Discussion of Initial Accountability Ratings
  - *Final* School Performance Level recommendations
  - Additional Issues: Reactions/suggestions on the PIs  
Suggestions for an “appeals” process  
Related “business rules” for the pilot
  - Next steps in the process; Adjournment

## Appendix B

### Summary of PJP Panel Members' Recommendations by Round of Judgments by Performance Indicator

Performance Indicator	<u>Round 1 Median</u>		<u>Round 2 Median</u>		<u>Final Median</u>		
	<i>Target:</i>	<i>Meets</i>	<i>Exceeds</i>	<i>Meets</i>	<i>Exceeds</i>	<i>Meets</i>	<i>Exceeds</i>
<b>Achievement</b>		75%	85%	75%	87%	75%	87%
<b>Equity</b>		+ 5%	+ 11%	+ 6%	+ 20%	+ 6%	+ 20%
<b>CCR - % HS Graduates</b>		80%	91%	85%	*	85%	*
<b>ACT Benchmark</b>		15%	25%	19%	*	20%	*
<b>Growth (Mdn Percentile)</b>		40	58	40	60	50	60

\*During the PJP discussions of the Round 1 results, prior to making Round 2 judgments, panelists agreed to a process of jointly considering both elements of the CCR PI. This process is discussed in the body of the report. In order to Meet the Target, schools must reach one of the two CCR criteria; to Exceed the Target, schools must reach *both* CCR standards.

**Appendix C:**

**FINAL - PJP Judges' Recommended School Performance Levels**

**FINAL PANEL MEDIAN RECOMMENDATIONS ARE SHADED**

**DIRECTIONS:** For each possible combination of school outcomes on the 4 Performance Indicators, decide what overall School Performance Level you would assign to the school. Indicate your judgment by marking **X** in the appropriate column for *each* row. Your judgments would apply to all schools in Wyoming. Please judge each combination of PIs, even if such a combination seems unlikely.

Performance Indicator Standard *				School Performance Level Recommendation				Notes
ACHIEVE	EQUITY	CCR	GROWTH	Exceeding Expectations	Meeting Expectations	Partially Meeting Expectations	Not Meeting Expectations	
(non-HS)								
A	A		A					
A	A		M					
A	A		B					
A	M		A					
A	M		M					
A	M		B					
A	B		A					
A	B		M					
A	B		B					
M	A		A					
M	A		M					
M	A		B					
M	M		A					
M	M		M					
M	M		B					
M	B		A					
M	B		M					
M	B		B					
B	A		A					
B	A		M					
B	A		B					
B	M		A					
B	M		M					
B	M		B					
B	B		A					
B	B		M					
B	B		B					

(HS)								
A	A	A						
A	A	M						
A	A	B						
A	M	A						
A	M	M						
A	M	B						
A	B	A						
A	B	M						
A	B	B						
<b>Performance Indicator Standard *</b>				<b>School Performance Level Recommendation</b>				<b>Notes</b>
<b>ACHIEVE</b>	<b>EQUITY</b>	<b>CCR</b>	<b>GROWTH</b>	<b>Exceeding Expectations</b>	<b>Meeting Expectations</b>	<b>Partially Meeting Expectations</b>	<b>Not Meeting Expectations</b>	
M	A	A						
M	A	M						
M	A	B						
M	M	A						
M	M	M						
M	M	B						
M	B	A						
M	B	M						
M	B	B						
B	A	A						
B	A	M						
B	A	B						
B	M	A						
B	M	M						
B	M	B						
B	B	A						
B	B	M						
B	B	B						
(K-12/7-12)								
A	A	A	A					
A	A	A	M					
A	A	A	B					
A	A	M	A					
A	A	M	M					
A	A	M	B					
A	A	B	A					
A	A	B	M					
A	A	B	B					

A	M	A	A					
A	M	A	M					
A	M	A	B					
A	M	M	A					
A	M	M	M					
A	M	M	B					
A	M	B	A					
A	M	B	M					
A	M	B	B					
A	B	A	A					
A	B	A	M					
A	B	A	B					
A	B	M	A					
A	B	M	M					
A	B	M	B					
A	B	B	A					
A	B	B	M					
A	B	B	B					
M	A	A	A					
M	A	A	M					
M	A	A	B					
M	A	M	A					
M	A	M	M					
M	A	M	B					
M	A	B	A					
<b>Performance Indicator Standard *</b>				<b>School Performance Level Recommendation</b>				<b>Notes</b>
<b>ACHIEVE</b>	<b>EQUITY</b>	<b>CCR</b>	<b>GROWTH</b>	<b>Exceeding Expectations</b>	<b>Meeting Expectations</b>	<b>Partially Meeting Expectations</b>	<b>Not Meeting Expectations</b>	
M	A	B	M					
M	A	B	B					
M	M	A	A					
M	M	A	M					
M	M	A	B					
M	M	M	A					
M	M	M	M					
M	M	M	B					
M	M	B	A					
M	M	B	M					
M	M	B	B					
M	B	A	A					
M	B	A	M					
M	B	A	B					
M	B	M	A					
M	B	M	M					
M	B	M	B					

M	B	B	A					
M	B	B	M					
M	B	B	B					
B	A	A	A					
B	A	A	M					
B	A	A	B					
B	A	M	A					
B	A	M	M					
B	A	M	B					
B	A	B	A					
B	A	B	M					
B	A	B	B					
B	M	A	A					
B	M	A	M					
B	M	A	B					
B	M	M	A					
B	M	M	M					
B	M	M	B					
B	M	B	A					
B	M	B	M					
B	M	B	B					
B	B	A	A					
B	B	A	M					
B	B	A	B					
B	B	M	A					
B	B	M	M					
B	B	M	B					
B	B	B	A					
B	B	B	M					
B	B	B	B					

\* For PIs: **A** = Above (Exceeds) Target, **M** = Meets Target; **B** = Below Target