
PLDs:

Petty Little Details or
Policy-Laden Drivers

Marianne Perie

Presentation for CAS SCASS

February 4, 2009



Introduction

- This presentation will
 - Focus on PLDs and the descriptions of proficient
 - Discuss when and how PLDs are used
 - Describe differences across states
 - Provide a framework for developing PLDs

The Rise and Fall of the PLD

- The first use of PLDs were actually descriptive, also called scale anchoring (NAEP Long Term Trend, 1984)
 - Set the cut score first, then describe the types of items students above the cut score are likely to be able to do that students below the cut scores are less likely to be able to do
 - Also the process behind original CTB Bookmark when descriptors were written after bookmark was placed
- Then there was a move towards writing the PLDs first to set policy



When to Write PLDs

- Some still claim that PLDs need to be written after cut scores are set to best describe the levels indicated by the cut scores
- But that is a responsive, descriptive approach
 - Okay for some settings but not for tests of accountability
- When we are asking students/teachers/ schools to attain a certain level of performance, we must describe that level in words and then find the cut score that matches it
 - And design a test that measures it well

Use of PLDs

- The standards and assessment guidance indicates that PLDs need to be written prior to and used for standard setting.
- PLDs are instrumental to the validity and defensibility of the standard-setting process
- Many researchers argue that the descriptors should be written early in the test development process and be used in developing test blueprints and item specifications.

Why Does this Matter Now that We've Designed our Assessments?

- Consider recent research comparing percent proficient across states...
 - Research focuses on rigor of cut scores
 - Fails to examine rigor of the definitions of each level
 - (Also fails to examine the rigor of the content standards or the alignment of state standards to NAEP standards, but that is a topic for another day)

Favorite Quote

- *Saying that all students must be at the proficient level or above by 2014, but leaving the definition of proficient achievement to the states has resulted in so much state-to-state variability in the level of achievement required to meet the proficient standard that “proficient” has become a meaningless designation. (Linn, 2005, p. 14)*

Webster's Definition of Proficient

- Proficient: “performing in a given art, skill, or branch of learning with correctness and facility (adj.); an expert (n.).”

Multiple Ways States Describe “Proficient”

- Satisfactory achievement
- Adequate understanding of the on-grade content
- Solid understanding of challenging subject matter
- Competency indicating preparation for the next grade level
- Ability to apply on-grade standards capably
- Acceptable command of grade-level content and processes
- Ability to apply concepts and processes effectively
- Solid academic performance...competency with challenging subject matter
- Solid academic performance...prepared for the next grade
- Mastery of grade-level standards
- High level of achievement...ability to solve complex problems

~taken from Beck, 2003

Proficient

- Everything in NCLB is based on 100% proficiency, but the law never addresses what proficiency means
 - Each state defines proficiency
 - Result is 52 definitions and little comparability
- Performance level descriptors (PLDs) have powerful influence over multiple areas of assessment
 - Cut scores
 - Item development
 - Reporting



So How Do We Develop Meaningful PLDs?

- Start early in the test development process
- Determine the minimum number of levels necessary to meet interpretive goals
- Assign labels to these levels
- Write policy definitions for each level
 - Describes the level of rigor
 - Same across all grades and subjects
- Add content and grade-specific details for each assessment for a full PLD
- Optional: Flesh out the PLDs further after the cut score has been set by adding specific examples

Number and Name of Levels

- Choose the fewest performance levels needed to fulfill your purpose
 - Too many are difficult to support unless the test is very long (reduces measurement power)
 - Meaningful distinctions become difficult with too many levels
 - Currently states use between 3 and 6 levels for NCLB tests
- Name the levels carefully as the names themselves carry meaning



Beck Guidelines for Naming Levels

■ AVOID

- ❑ Nebulous, unclear, or unreasonable terms or oxymorons (needs improvement, reasonable mastery)
- ❑ Normative terms (average, typical)
- ❑ Moving terms (nearly X, approaching the standard, emerging, progressing) as they apply to all parts of the level, making it more difficult to distinguish borderline performance
- ❑ Non-educational terms (normal, inadequate, novice/apprentice)
- ❑ Non-parallel terms (Outstanding, Pass, Warning)



Create Policy Definitions

- What are they?

- Assert a policymaker's position on the desired level of performance or rigor intended at each level
- Not linked to content
- Same across all assessments in one program
- Help with comparability

- How do I develop them?

- Use a similar set of words that are memorable and that distinguish clearly among the performance levels
- Write definitions that are concise and clear



Finally, Develop Full PLDs

- PLDs express the knowledge and skills required to achieve each level of performance for a specific assessment and are linked directly to the content standards for that assessment.
- Best drafted by committee of content experts with policy guidance
- Start with the policy definitions and expand these definitions in terms of specific knowledge, skills, and abilities required at each level for each subject for each grade



Materials Needed

- Policy definitions
 - Content standards
 - Test blueprints
 - Sample items and rubrics
-
- *Be careful to write PLDs to content standards but within test blueprints or specifications. Do NOT write to the item level*

Example of Names States Use

- Below Basic
- Basic
- Proficient
- Advanced
- Unsatisfactory
- Limited knowledge
- Satisfactory
- Advanced
- Approaches Standard
- Meets Standard
- Exceeds Standard
- Did not meet the standard
- Met the standard
- Commended performance
- Level I
- Level II
- Level III
- Level IV



Example of a Policy Definition

■ NAEP Proficient

- ❑ *Solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.*
- ❑ (This is a very rigorous definition, which may not be appropriate for other assessments.)



Example from a State Assessment

- Advanced: Superior academic performance indicating an in-depth understanding and exemplary display of the skills included in [State's] Academic Standards;
- Proficient: Satisfactory academic performance indicating a solid understanding and adequate display of the skills included in [State's] Academic Standards;
- Basic: Marginal academic performance, work approaching, but not yet reaching, satisfactory performance, indicating partial understanding and limited display of the skills included in [State's] Academic Standards; and
- Below Basic: Inadequate academic performance that indicates little understanding and minimal display of the skills included in [State's] Academic Standards.



Example of a Full PLD: Grade 4 Writing

Proficient

- As a proficient writer, the student establishes a central focus, generally organizes and connects ideas, and includes some supporting details. The student demonstrates some variety in sentence structure and word choice and uses basic conventions of print.

Advanced

- As an advanced writer, the student establishes and develops a central focus, organizes and connects ideas, and elaborates on supporting details coherently. The student varies sentence structures, chooses words effectively, and uses conventions of print.

Challenges in Writing Good PLDs

- PLDs are written as if the student should be able to do everything but tests are compensatory meaning strong performance in one area overcomes weak performance in another
- We want to use the “sometimes” approach but this makes standard setting and interpretation of reports difficult
- We need to find the appropriate balance between keeping PLDs general enough to apply to multiple forms but specific enough to provide useful information to our stakeholders

Issue of PLDs and Comparability

- Within a state, using policy definitions will help ensure similar meaning across assessments in terms of rigor
- Across states, there are no common policy definitions and rigor varies dramatically
 - In my opinion, this means that comparisons of percent proficient across states are meaningless and should not be done

Compare Two State Definitions of Proficient...

State A

- This level denotes demonstration of solid academic performance on challenging subject matter reflected by the content standards. This includes knowledge of subject matter, application of such knowledge to real-world situations, and content relevant analytical skills.

State B

- Performance at this level indicates that the student has partial success with the challenging content of the standards but performance is inconsistent. A [Proficient] student answers many of the questions correctly but is generally less successful with questions that are most challenging.



Question to Ponder

- If researchers are going to continue to compare states to one another, is it important to adopt a common policy definition of Proficient?
 - States could still supplement that definition with state-specific content for the full PLD
 - It would ensure that the rigor of the Proficient level was the same across states

Thank You

- Full paper published in EM:IP, volume 27, issue 4 (Winter 2008)
- If you have other questions, you can reach me at mperie@nciea.org