



A Handbook for Creating Range and Target Performance Level Descriptors

M. Christina Schneider

Karla L. Egan

The National Center for the Improvement of Educational Assessment

A Handbook for Creating Range and Target Performance Level Descriptors

Testing programs have historically used a single performance level descriptor (PLD) for each grade to describe a student's level of achievement (e.g., Basic, Proficient, Advanced) on a large scale assessment (Lewis & Green, 1997; Perie, 2008). More recently states have begun to create a set of interwoven PLDs to guide (a) teacher and item writer expectations for the progressions of evidence needed to categorize student performance for a performance level, (b) participant recommendations during the standard setting process, and (c) stakeholder interpretation of student score reports. We recommend that these PLDs be developed sequentially so that the PLDs are related and consistent with the PLD development framework proposed by Egan, Schneider, and Ferrara (2012). The purpose of the PLD development framework is to enable valid inferences about student content area knowledge and skill in relation to a state's content standards measured on a large scale assessment. This document and the set of corresponding Powerpoints are meant to support the creation of an interwoven set of PLDs.

PLD Development Framework

Policy makers begin establishing the rigor of a state's performance standards through the development of PLDs. Egan, Schneider, and Ferrara (2012) proposed four stages of PLD development to correspond with the closely linked uses of PLDs—for item development, standard setting, and score reporting. More recently some states are also using PLDs designed for item development purposes to assist teachers in the classroom in better understanding the progression of skill expectations for each standard across performance levels. The types of PLDs Egan, Schneider, and Ferrara defined are Policy, Range, Target, and Reporting. The development of these type of PLDs provides a construct-centered approach to assessment design (Messick, 1994). Because the intent is to accurately represent student performance, once information from an assessment is available, it is optimal to revise PLDs as necessary (Plake, Huff & Reshetar, 2010). The types and purposes of PLDs are discussed next.

Policy PLDs

In the first stage of the PLD development framework, the state develops Policy PLDs. Policy PLDs are important communication devices for a state's vision of what it means, for example, to be college and career ready. As such, a state optimally develops a policy-based claim. This claim clearly explicates the state's intended take-away message regarding a student's achievement within each performance level. The Policy PLDs should be consistent across grades with the exception of the policy description at the high school level. Nationally, the high school model is moving toward policy-based claims in regard to student readiness for college and careers. Table 1 shows an example of a Policy Based PLD for Highly Proficient from the state of Utah.

Table 1: Utah Policy PLD for Highly Proficient

Highly Proficient

The Level 4 student is highly proficient in applying the English language arts/literacy knowledge/skills as specified in the Utah Core State Standards. The student generally performs significantly above the standard for their grade level, is able to access above grade level content, and engage in higher order thinking skills independently.

* For Secondary English 11, this level of performance also likely indicates students are well prepared for postsecondary success in language arts.

Retrieved from <http://sageportal.org/wp-content/uploads/2013/11/2013-11-21-SAGE-PLD.pdf>

Range PLDs

For each standard and performance level on an assessment, Range PLDs should explicate observable evidence of student achievement, demonstrating how the skill changes and becomes more sophisticated across performance levels. These descriptors describe the “range” of skills within and across performance levels, and they should do so for each standard. Schneider, Huff, Egan, Gaines, and Ferrara (2013) wrote that for PLDs to be the foundation of test score interpretation, they should reflect more complex knowledge, skills, and abilities (KSAs) as the performance levels increase (e.g., more complex KSAs should be expected for Advanced than for Proficient).

Learning trajectories are described increasingly in the literature as theoretical underpinnings for curriculum development, instruction, and assessment of learning. The purpose of a learning trajectory is to inform researchers and educators about general developmental pathways of learning so that they can set reasonable, achievable learning goals and provide appropriate guidance for instruction and assessment in a given content area. Smith, Wiser, Anderson, and Krajcik (2006) defined a learning progression as the description of the increasingly more sophisticated ways of reasoning in the content domain that follow one another as a student learns. Learning progressions depend upon instruction, and multiple pathways may be used to help a student develop expert reasoning in the content. Simon (1995) posited that learning progressions, while based in research in children’s learning, are partly hypothetical. Often in the research literature, learning progressions reflect how one standard is a precursor skill to another. While this is no doubt true, in large scale assessments students will often answer items correctly and incorrectly within the same standard. Therefore, for our purposes we use the term learning *trajectory* to denote within-standard reasoning complexity that increase in sophistication as the performance level increases. Such skill advancement is often related to increases in content difficulty, cognitive level increases with the content, and the context used to present a task. Thus

the focus on the trajectory for the purpose of the Range PLD is to define the trajectory of more complex reasoning within a standard.

Assessments that are being designed for the multiple purposes of measuring student achievement and growth should be derived from the combination of the learning goals and the developmental progression engendered in a learning trajectory because, as Clements and Sarama (2004) wrote, "Developmental progressions...[are] descriptions of children's thinking and learning...and a related, conjectured route through a set of instructional tasks" (p. 83). The outcome of instructional tasks or assessment tasks should be the same: observable evidence of what students know and can do. In terms of learning progressions, the Common Core Standards may be thought of as the learning goals for students at each grade level, and the Range PLDs may be considered developmental trajectories—evidentiary statements regarding children's observable thinking and skills as they pass along the path to the learning goals. In the development of Range PLDs the state defines the expected learning trajectory, which is useful to teachers, but it also aligns the trajectory with its vision for student performance in terms of mastery of the content. An example of a Range PLD for two standards in Utah is shown in Table 2.

Table 2: Sample Range PLDs from Utah

		For grade-appropriate, low complexity texts, the Level 1 student	For grade-appropriate, low-to-moderate complexity texts, the Level 2 student	For grade-appropriate, moderate-to-high complexity texts, the Level 3 student	For grade-appropriate, high-complexity texts, the Level 4 student
Reading Literature					
Range	RL. 3.1	asks and answers explicit questions to demonstrate understanding of a text.	asks and answers explicit questions to demonstrate understanding text, referring to the text as the basis for answers.	asks and answers questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answers.	asks and answers complex questions to demonstrate understanding of a text, referring explicitly to the text as the basis answers.
Range	RL. 3.2	identifies details to recount stories; identifies explicitly stated central messages, lessons, or morals and identify details.	identifies key details to recount stories; determines central messages, lessons, or morals.	recounts stories, including fables, folktales, and myths from diverse cultures; determines the central message, lesson, or moral and explains how it is conveyed through key details in the text.	explains details to recount stories; determines implicitly stated central messages, lessons, or morals; and explains how they are conveyed through key details.

Retrieved from <http://www.schools.utah.gov/assessment/Adaptive-Assessment-System/English-Language-Arts/DRAFTELAandLiteracyGrade3PLDs.aspx>

When a state has content standards and Range PLDs in place, the last step for making the educational system cohesive and transparent is to develop a curriculum with instructional tasks and resources that move students through the range of skills located in the trajectory for a standard and provide students with multiple opportunities to learn. When both instructional activities and tasks and assessment items are based upon the same trajectory of expectations, opportunities to measure what students know and can do and teacher effectiveness are greatly enhanced. Because a trajectory can be either confirmed or disconfirmed when test data is reviewed, it is recommended that the Range PLDs and corresponding test specifications and instructional materials be edited after the initial test administration.

Target PLDs (Standard Setting PLDS)

Target PLDs define the state's minimum policy and content-based expectations (e.g., what it means to be Proficient) for a performance level. Target PLDs are the lower-bound descriptions of the performance level, and they are used to guide the cut score recommendation. These descriptors target the skills all Proficient students, for example, should have in common. They are distilled from the Range PLDs and describe the minimum amount of information, for example, that the barely Proficient student should know and be able to do beyond that of a Basic student as shown in Table 3.

Table 3: Sample Target PLD

Proficient Target Student
Proficient students, located <i>just</i> at the cut score, are able to read, analyze, and make connections to texts of different genres. They are able to cite multiple specific pieces of evidence to support their conclusions as well as the inferred main of idea in unfamiliar texts that they read. They are able to use multiple strategies to determine the meaning of unfamiliar words such as using root words, affixes, prefixes, and context clues. Proficient students, located <i>just</i> at the cut score, are ready to be successful in the next grade.

Because standard setting participants are striving to differentiate the skills that separate the highest performing Basic student from the lowest performing Proficient student, these descriptions are shorter and describe *just* the skills that appear at the threshold between the two performance levels. Another way to think of the Target PLD is that these PLDs describe the extra skills that a minimally Proficient student should know that are just beyond those possessed by the highly Basic student. The minimally Proficient student is able to demonstrate the skills of the highly Basic student and a small portion of the skills found in the Proficient Range PLDs. Defining the extra skills beyond those identified in the Basic Range PLD that are characteristic of the break in skills between the two performance levels is the content conceptualization of the cut score on an assessment. The cut score is the lowest score a student can obtain to be categorized as Proficient, for example.

Reporting PLDs

Reporting PLDs are optimally created after final cut scores are adopted. Reporting PLDs represent the reconciliation of the Target PLDs with the final cut scores. The Target PLDs reflect a state's *initial* expectation for minimal student performance within a performance level whereas the Reporting PLDs reflect actual student performance based upon the final approved cut scores. The reporting PLDs define the appropriate inferences stakeholders may make based upon the student's test score in relation to *the final approved cut scores*. A state should clearly explicate in the Reporting PLDs whether the target student (the student located at the cut score) or the typical student is being described when short summary descriptors are placed upon score reports. Often due to scheduling requirements, Reporting PLDs are developed in advance of the final

approved cut scores. In such cases, it is wise to note to the public that after the first administration of the assessment some edits to the Reporting PLDs may be made. More useful to teaching and learning is to examine the Range PLD trajectory and compare the hypothesized evidence of what students know and can do to the actual empirical evidence to determine if the trajectory has been confirmed or disconfirmed. It is recommended that the Range PLDs be edited after the initial test administration using this information and made accessible to teachers and test developers.

The advent of Range PLDs, has added additional complexities to the PLD development process. Therefore, a guide to building Range and Target PLDs for use on large scale assessments would likely be helpful to states as they work to build assessments with the multiple purposes of measuring student achievement and growth, producing data to inform teacher and school effectiveness, and supporting instructional decision making. These purposes require that the test construct and the type of evidence needed to determine student proficiency be clearly defined and transparent.

Considerations for Creating Range PLDs

Range PLDs and Target PLDs should be created in separate workshops because the PLDs have different purposes and conceptualizations. Educators who have been focused on Range PLD development find it difficult to change their cognitive focus to Target PLD development immediately after crafting the Range PLDs. Because Target PLDs are built from the Range PLDs, the work should be sequenced. To that end, the Range PLD is optimally created first. We have created corresponding Powerpoints to illustrate how to train panelists to create the Range PLDs and Target PLDs. Generally the Range PLD development process will take five days, including editing time. We recommend educators draft the Range PLDs initially and the state, along with its vendor, edit the Range PLDs for clarity of voice and cohesiveness across grade levels and content areas. The Target PLD creation process optimally occurs as one of the first steps of the standard setting workshop.

Developing Range PLDs

To develop Range PLDs expert educators that represent the teaching population of the state should be convened. The educators should be demographically diverse and be drawn from an educator pool with experience in using the state standards and with teaching students with the following characteristics: gifted education, special education, English language learner education, and post-secondary education. Because of the desire to describe college and career readiness (CCR), for those courses in which CCR policy claims are being considered, at least two members of the panel should represent two-year and four-year post-secondary institutions. We also recommend test development experts be present and, if possible, they should participate in the development of the Range PLDs.

The educators and test development experts should be divided into grade and content area panels based on their expertise and teaching experience. It is possible to have one panel per grade; however, to promote cohesiveness across grades we generally find it

beneficial to build grade-band panels that develop Range PLDs for one grade the first day and an adjoining grade the second day.

The grade-level panel can be divided into subgroups of three, with each subgroup assigned one-third of the state standards for their grade. Each subgroup then parses each standard they are assigned into a trajectory of observable student evidence that should be expected for that standard across the performance levels using a standardized template developed in Excel, for example. On the third through fifth days, a smaller panel should engage in the editing work. For this portion of the work we recommend that the group be comprised of content experts at the state level along with the vendor.

A sample agenda from a three-day workshop is provided in Table 4. Note, editing is expected to occur beyond the workshop.

Table 4. Agenda for a Range PLD Writing Workshop

Day	Time	Activity
Day 1	8:00 AM	Welcome and Policy Vision—State Agency
	8:30 AM	Range PLD Training
	9:30 AM	Move to Breakout Rooms, Break, and Technology Testing
	10:00 AM	PLD Development
	12:00 PM	Lunch
	12:45 PM	Continue PLD Development
	2:30 PM	Break
	4:30 PM	End Day 1

Table 4. Agenda for a Range PLD Writing Workshop Continued

Day 2	8:00 AM	Complete Set 1 PLDs
	9:00 AM	Break
	9:15 AM	Begin Set 2 PLDs
	12:00 PM	Lunch
	12:45 PM	Continue PLD Development
	2:30 PM	Break
	4:00 PM	Evaluation
	4:15 PM	Collect Materials
	4:30 PM	Workshop Ends
Day 3	8:00 AM	Across Grade Coherence Review
	8:30 AM	Reviewing Across Grades 3, 4, 5
	9:45 AM	Break
	10:30 AM	Reviewing Across Grades, 6, 7, 8
	12:00 PM	Lunch
	12:45 PM	Continue
	1:00 PM	Reviewing Across Grades 9, 10, 11
	3:00 PM	Evaluation
	3:15 PM	Collect Materials
	3:30 PM	Workshop Ends

Characteristics of Range PLDs

Range PLDs should possess the following characteristics.

1. The state's policy vision should be clearly embedded in the Range PLD.

Embedding Policy PLDs into the Range PLDs sends a clear signal for teaching and learning because the state's vision of what it means to be Proficient (or college and career ready) is described. The Policy PLD supports the Range PLD and assessment program in several ways. First, the state's intended take-away message regarding a student's achievement within each performance level becomes the centerpiece of how an evidence trajectory is aligned with a performance level. Second, the embedded Policy PLD supports a research agenda that can be validated to confirm or disconfirm its veracity. If a performance level claim purports that students have the necessary skills to be successful in entry-level, credit bearing courses in 2-year or 4-year postsecondary institutions, a number of validity studies can be conducted to gather evidence to support this claim. If students are ready to move to the next grade, student course grades may be collected to determine the probability a student has of earning a C or better, for example, in the subsequent grade.

Table 5 shows how embedded Policy PLDs can also sent important policy messages for teaching and learning. For Highly Proficient students, for example, the Policy PLD implies that teachers should expect students to work at cognitively higher levels of thinking than stipulated in the standard with grade level content, and teachers may move these students to the next grade level standards when and where appropriate. This gives teachers important information about ways to adapt instruction for the highest performing students.

Table 5: Sample Policy PLDs from Utah Embedded into a Hypothetical Assessment Claim

	Below Proficient	Approaching Proficient	Proficient	Highly Proficient
Policy PLD	The student generally performs significantly below the standard for <i>the grade</i> , is likely able to <i>partially</i> access grade-level content and engage with higher-order thinking skills with extensive support	The student generally performs slightly below the standard for <i>the grade</i> , is likely able to access grade-level content and engage in higher-order thinking skills with some independence and support.	The student generally performs at the standard for <i>the grade</i> , is able to access grade level content, and engage in higher order thinking skills with some independence and minimal support.	The student generally performs significantly above the standard for <i>the grade</i> , is able to access above grade level content, and engage in higher order thinking skills independently.
Reading Claim 1	The Below Proficient Student's ability to read closely and analytically to comprehend a range of increasingly complex literary texts is significantly below the standard. The student is likely able to access a restricted range of literary texts and engage in higher order thinking skills with extensive support.	The Approaching Proficient Student's ability to read closely and analytically to comprehend a range of increasingly complex literary texts is slightly below the standard. The student is likely able to access a range of literary texts and engage in higher order thinking skills with some independence and support.	The Proficient Student's ability to read closely and analytically to comprehend a range of increasingly complex literary texts is at the standard. The student is likely able to access a range of increasingly complex literary texts and engage in higher order thinking skills with some independence and support.	The Highly Proficient Student's ability to read closely and analytically to comprehend a range of increasingly complex literary texts is above the standard. The student is likely able to access an above grade range of increasingly complex literary texts and engage in higher order thinking skills independently.

2. The Range PLD should define differences in content across performance levels rather than the frequency with which students respond to content.

The use of frequency-based terms such as *partially demonstrates*, *adequately demonstrates*, and *consistently demonstrates* as a defining phrase to separate performance levels is insufficient in describing what students actually should know and be able to do and does not describe a trajectory of content acquisition that helps to guide teacher instructional decisions or item writing design decisions. Therefore, such frequency terms should generally be avoided. Oftentimes such language occurs in reading or English language arts when PLD writers are trying to define when students are able to respond and make inferences from text, as an example. Describing the expected text complexity of the stimulus for the items, the genre of text, the interest level of the text, the use of familiar versus unfamiliar vocabulary, vicinity of supporting details (are they clustered or sparsely spread in the text) or context describe what makes the content easier or more difficult and should be used to differentiate performance across performance levels.

3. Range PLDs should describe contextual or scaffolding characteristics.

Contextual or scaffolding characteristics of items and instructional tasks can elicit relevant student skills in the content area. This information is often central to understanding why a student is or is not able to respond to an item; thus, discussing the contextual or scaffolding characteristics that need to be present in order for the student to demonstrate the skill is an important characteristic of Range PLDs. Students may be able to respond to informational texts, for example, when that text has organizational structures or graphics that help the student locate and interpret the information. Without such scaffolds, the students may not be able to independently re-read and locate the necessary details to answer a question.

Similarly in mathematics for performance task items, as an example, asking the student to respond to the stimulus with multiple questions or item demands can be more difficult if the questions are presented together without a space to answer each question separately. If multiple task questions are presented in a paragraph followed by an area to write, students are likely to be less successful than if asked a question, given a space to respond, and then asked the next question. Descriptors such as “the student solves multistep or multi-question tasks when scaffolding is provided for each step” provides important information to both item writers and teachers with regard to the conditions under which the student is able to successfully show evidence of his or her learning.

4. The Range PLD content is coherent within and across grades.

Range PLDs should describe coherently increasing expectations for achievement across performance levels within and across grade levels. We should see

expectations for student skills becoming more rigorous as the performance levels increase within a grade, *and* we should see related skills/standards within a grade growing similarly across objectives. We should also see trajectories of growth in standards across grades (Hess & Kearns, 2011a; 2011b). Standards from one grade often increase in content rigor across grades; therefore, growth from *Proficient* in one grade to *Proficient* in the next, as an example, should be evaluated to ensure the progression is cogent.

When developing Range PLDs, the first step is having a grade-level committee develop the learning trajectory within grade. However, a second step in their development is to have a grade span committee compare skill growth across grades for similar standards to ensure that growth across years is described meaningfully, accurately, and the construct increases in rigor in a way that makes sense.

In English language arts, in particular, at the beginning of the development process it is important to ensure that there is a clear vision regarding how students will interact with texts during instruction and assessment. Differences in student abilities can be supported by using texts of differing complexity coupled with tasks that are used to elicit evidence regarding what students can do. Students may receive text at the grade-level only, in which case, students scoring in the lower performance levels would have little ability to analyze the text. They might be able to understand a line of a poem whereas moderate performing students may glean a main idea of a poem, but they may not be able to interpret the author's multiple themes. Or, students may receive texts targeted to their independent reading levels so texts may be below, slightly below, on grade, or above grade level. As a result of this latter testing model, all students may be expected to demonstrate many, though not all, of the skills in the standards and therefore, the performance levels may be primarily driven by the sophistication of the texts the student is able to analyze and their abilities to, for example, analyze figurative language while making connections to allusions to other texts, history, or knowledge from other disciplines. These conceptualizations should be fleshed out prior to the beginning of workshop, and they are central to the test development process.

5. The Range PLD describes what students should be able to do using observable, measurable verbs.

Descriptors that use implicit verbs require that item writers and teachers infer their meaning. When explicit verbs are used, teachers, parents, and stakeholders have a better sense of what the student can actually do (at least within some reasonable probability). For example, writing "The Level 2 student recognizes the main idea" does not lead to a clear conceptualization of what the student should do to demonstrate his or her skill. Rather, "The Level 2 student is able to tell or

show the main idea” clearly describes the evidence that needs to be collected to measure the skill.

6. The state’s item alignment methodology is explicitly found in the within-grade PLD progression.

States use different item alignment policy models within K–12 high stakes testing, but Webb’s (2005) depth of knowledge procedure is currently the most common process used to conduct item-to-standard alignment, in particular from a cognitive perspective. The item alignment policy model a state uses on the year-end assessment it develops and administers should be embedded into the PLD development process in two ways. First, the alignment methodology should provide an implicit link to the expectations for student performance that should appear in the PLDs. Second, the alignment methodology prescribes the cognitive language that should be used.

In the most common policy model, 50% of a test’s items have to be aligned at or above the cognitive level of the state standard to achieve depth of knowledge consistency (Webb, Herman, and Webb, 2007). Under this model, the state standards represent the minimum a student should know and be able to do at the end of the year. That is, it may be reasonable to depict in PLDs that students are expected to demonstrate knowledge, skills, and abilities with the content beyond what is specified in the standards from a cognitive perspective at the performance level of critical importance for accountability purposes.

Other states use a policy model requiring that items on the year-end assessment match the level specified in the state standard. Under this model, the state standards represent the cumulating expectations regarding what a student should know and be able to do at the end of the year. The point to consider here is that *Proficient*, for example, may implicitly mean reaching the standard or going beyond the standard depending upon the item alignment methodology a state uses.

The state’s item alignment methodology should influence the cognitive language found in the performance level descriptors. That is, if items are to be created to the cognitive level of the standard then this should be depicted in the PLDs by holding the cognitive level constant across performance levels while the rigor of the content increases. If items are to be created above the cognitive level of the standard then this should be depicted in the higher performance levels by increasing the cognitive level across performance levels while also increasing the rigor of the content increases. At times, the standard may be more than one statement. In such cases the statements can be sequenced to show an expected trajectory of performance. Table 6 provides a Grade 4 Mathematics example.

Table 6: Example of embedding cognitive language from a standard to a PLD: Make a line plot to display a data set of measurements in fractions of a unit ($\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$). Solve problems involving addition and subtraction of fractions by using information presented

		Below	Approaching	Proficient	Advanced
		Grade 4			
Range	Make a line plot to display a data set of measurements in fractions of a unit ($\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$). Solve problems involving addition and subtraction of fractions by using information presented in line plots	Plot a line graph to a real-world problem using fractions $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{1}{8}$.	Plot a line graph to a real-world problem using fractions $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{1}{8}$ and interpret data from a line plot to solve to solve one-step addition problem with like denominators.	Plot a line graph to a real-world problem using fractions $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{1}{8}$ and interpret data from a line plot to solve to solve two-step addition problem with like denominators.	Plot a data set to a real-world problem using fractions $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{1}{8}$ and interpret data from a line plot to solve a multi-step word problems involving subtraction with like denominators

Considerations for Building Target PLDs

Target PLDs define the expected performance of a student who just enters a performance level and as a result they represent the lower border of the Range PLDs. Target PLDs help refine the hypotheses underlying the Range PLDs by asking educators to identify those skills that constitute the transition point of moving from one performance level to the next. Therefore, Target PLDs help transition from the breadth of the Range PLDs to the narrow focus of those skills most essential to possess to be considered for example, Proficient.

The importance of Target PLDs to the standard setting process cannot be underestimated. Target PLDs provide standard setting panelists with a common understanding of the minimum expectations for entry into a performance level, and these expectations are the foundation for the cut score recommendations. Because panelists arrive at a cut-score recommendation workshop with many different expectations for students, standard setting facilitators must quickly and efficiently assist participants in using the expectations found in the Range PLDs to build a common short description of what each performance level represents as well as what is at the heart of that expectation. For example, in English language arts perhaps what is at the heart of proficiency is the student's ability to make inferences or in mathematics to see and use relationships between inverse operations to solve patterns. When entering the standard setting workshop participants must study and synthesize the Range PLDs and draw inferences about what is common across the standards within a performance level that describes the salient knowledge and skills that move a child from one performance level

to the other. Next participants must determine the level of the content for which students must demonstrate those skills. For example, in English language arts if it is the ability to make inferences that are the hallmark of the Proficient student, the panelists must define the minimum content based expectations for this skill. If students are able to demonstrate inferences in narrative and informational texts but not in poetry would the panel consider the student Proficient? Or in mathematics if it is the ability of a Grade 3 student to see and use relationships between inverse operations to solve problems is using the relationship between addition and subtraction and not multiplication and division sufficient for the panel to consider the student Proficient?

Developing Target PLDs

Creating the Target PLDs is optimally accomplished on Day 1 of the Standard Setting. To develop Target PLDs, panelists should begin studying the Policy and Range PLDs to understand the intent of the sponsoring agency regarding the testing program. It is the job of the panelists to operationalize this intent through Target PLDs and cut scores.

Next, the panelists should study and begin to annotate the Range PLDs. How this is accomplished likely differs by content area. Across content areas however panelists analyze the cognitive expectations, content information, and the context found in the Range PLDs.

Table 7: English Language Arts/Literacy (ELA/L)

		Below	Approaching	Proficient	Highly Proficient
Policy		The Level 1 student is below proficient in applying the English language arts/literacy knowledge/skills as specified in the Utah Core State Standards. The student generally performs significantly below the standard for the grade level and is likely able to partially access grade-level content and engages with higher-order thinking skills with extensive support.	The Level 2 student is approaching proficient in applying the English language arts/literacy knowledge/skills as specified in the Utah Core State Standards. The student generally performs slightly below the standard for the grade level and is likely able to access grade-level content and engages in higher-order thinking skills with some independence and support.	The Level 3 student is proficient in applying the English language arts/literacy knowledge/skills as specified in the Utah Core State Standards. The student generally performs at the standard for the grade level, is able to access grade-level content, and engages in higher-order thinking skills with some independence and minimal support.	The Level 4 student is highly proficient in applying the English language arts/literacy knowledge/skills as specified in the Utah Core State Standards. The student generally performs significantly above the standard for the grade level, is able to access above grade-level content, and engages in higher-order thinking skills independently.
		For a grade-appropriate, low complexity texts, the Level 1 student	For a grade-appropriate, low-to-moderate complexity texts, the Level 2 student	For a grade-appropriate, moderate-to-high complexity texts, the Level 3 student	For grade-appropriate, high-complexity texts, the Level 4 student
Reading Literature					
Range	RL.3.1	asks and answers <u>explicit</u> questions to demonstrate understanding of a text.	asks and answers explicit questions to demonstrate understanding text, <u>referring to the text as the basis for answers.</u>	asks and answers questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answers.	asks and answers <u>complex</u> questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answers.

Range	RL.3.2	identifies details to recount stories; identifies explicitly stated central messages, lessons, or morals and identify details.	identifies key details to recount stories; determines central messages, lessons, or morals.	recounts stories, including fables, folktales, and myths from diverse cultures; determines the central message, lesson, or moral and explains how it is conveyed through key details in the text.	explains details to recount stories; determines implicitly stated central messages, lessons, or morals; and explains how they are conveyed through key details.
-------	--------	--	---	---	---

Source: <http://www.schools.utah.gov/assessment/Adaptive-Assessment-System/English-Language-Arts/DRAFTELAandLiteracyGrade3PLDs.aspx>

Table 7 shows a sample of a Grade 3 ELA/L Range PLD from Utah. The first standard is “asks and answers questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answer.” The Range PLD has the following trajectory from easy to difficult with key points underlined and in red below.

- asks and answers explicit questions to demonstrate understanding of a text.
- asks and answers explicit questions to demonstrate understanding text, referring to the text as the basis for answers.
- asks questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answer
- asks and answers complex questions to demonstrate understanding of a text, referring explicitly to the text as the basis for answers.

Restated, students move from answering explicit questions without supporting their answers to answering explicit questions with textual support as they progress from Below Proficient to Approaching Proficient. What is notable in the Proficient descriptor is that the questions are not restricted to explicit questions. Thus, the conclusion to be drawn is that Proficient students are answering questions that require an inference by referring to the text. Finally, Highly Proficient students are answering complex questions, signifying that the required inferences may require understanding layers of meaning within the text. Should the remaining standards develop similarly and coherently, we can expect the break in skills that differentiate the Approaching Proficient student from the Proficient student to be the ability to answer questions that require students to make inferences. Note as well, the state has documented that Proficient students should be engaging in this skill with grade-appropriate, moderate- to high-complexity text. Thus, it is likely that the ability to make an inference should be a key

descriptor in drafting the Target PLD for the Proficient level. During the standard setting process participants will likely center their focus on identifying items that are inference based and isolating which moderate to high complexity text panelists identify as being the text a minimally proficient student should be able to analyze accurately. To accomplish the process of distilling the Range PLDs to the Target PLDs, facilitators should have standard setting panelists engage in a half-day process similar to the following. Note this process is demonstrated in the corresponding Powerpoint slides.

Step 1: Annotate

The panelists should discuss each skill and annotate or highlight the Range PLDs text that is different across performance levels. For ELA highlighting or underlining is likely optimal, no matter if the process is completed electronically or via paper and pencil. Where multiple skills are present within one standard descriptor, as is common in science and mathematics, annotation into *just* Proficient (P-), average Proficient (P), or above Proficient (P+) using the annotation process shown in Figure 2 for Approaching and Proficient is likely helpful to participants.

Note when annotation is used, the process is done for each performance level. For example, Table 8 shows panelists have discussed and indicated that “creates a model showing the sun as an energy source that results in evaporation” is a skill expected of the just Approaching Proficient student (A-), whereas “recognizes relative percentages of water found in various locations on Earth” is a skill expected of the average Approaching Proficient student (A). It is important to note that within each standard, participants do not need to find descriptors for the just, average, and average student. For some standards with difficult content, it may be that only average or high students within that range should be expected to show evidence of understanding the content.

Table 8: Portion of Grade 4 Science Range PLD from Utah

Below Approaching Proficient Highly Proficient					
Water Cycle					
Range	I.1 Describe the relationship between heat energy, evaporation and condensation of water on Earth.	Identifies that water is stored in various locations. Describes the sun as an energy source that results in evaporation. Identifies examples of the states of water.	Recognizes relative percentages of water found in various locations on Earth (A). Creates a model showing the sun as an energy source that results in evaporation (A-). Gives examples of the states of water pertaining to evaporation (A) and condensation (A).	Compares the locations (P) and percentages (P+) of water found in various locations on Earth. Investigates (P-) and records (P-) data showing the effect of temperature on the state of water. Records evidence of evaporation and condensation (P).	Compares and contrasts the effects of temperature change on evaporation and condensation. Collects, records, and interprets data from an experiment of changing states of water. Forms predictions of states of water from data.
Range	I.2 Describe the water cycle.	Identifies the processes of evaporation, condensation, and precipitation. Draws a simple diagram or model of the water cycle.	Describes the processes of evaporation (A-), condensation (A), and precipitation (A). Explains how water passes through the water cycle and is distributed to different locations. (A+) Constructs and labels a diagram modeling the water cycle (A).	Constructs a model of the processes of evaporation, condensation, and precipitation (P-). Identifies that evaporation occurs from people (P+), plants (P+), ice (P-), and ground water (P). Supports predictions and inferences based on the water cycle with data and evidence (P). Using provided resources, constructs a complex diagram of the water cycle including the concept that the total amount of water on Earth is constant (P).	Independently constructs a complex diagram of the water cycle. Explains how the water cycle affects human activities.

Source: <http://www.schools.utah.gov/assessment/Adaptive-Assessment-System/Science/DRAFTScienceGrade4PLDs.aspx>

Step 2: Compile

After the annotation process is completed for each Range PLD, panelists should compile the content, skills, and context associated with the *just* Approaching Proficient, Proficient, and Highly Proficient categories. These are found by taking the phrases noted with a “-” and making a list (or deleting other phrases if working electronically) for each performance level. It is possible to have participants work on all three categories, or if participants are divided into three tables to have each table focus on a category. Participants should review the list and discuss what recurring themes across the standards participants see for the borderline students. These themes should be summarized with a bullet list for presentation to the larger group and for comparison and contrast with the other descriptors as shown in the corresponding Powerpoint slides.

Step 3: Summarize

The themes for each Target Student Descriptor (e.g., Approaching Proficient) should be shared across the panel and contrasted with the other Target Student Descriptors (e.g., Proficient). During the discussion, the bulleted list can be condensed to a one-paragraph descriptor that includes the policy descriptor, the inference regarding what is at the heart of the performance level, and the content and context in which the student who has just the knowledge and skills to be Proficient is likely able to do. Participants should clearly be able to articulate how the PLDs reflect more complex KSAs as the performance levels increase (e.g., more complex KSAs should be expected for Advanced than for Proficient). It should be understood that standard setting is about identifying what students should know and be able to do, and during the standard setting process additional pieces of evidence are used to shape the understanding of what Target Students should be able to do with multiple opportunities to learn.

Characteristics of Target PLDs

Range PLDs and Target PLDs are developed at different times because they have different purposes, and the cognitive task for their development is different. The two sets of PLDs, however, should be clearly linked. Generally the Target PLD process described will take two to three hours, especially if the Range PLDs have been sent to the standard setting panel for review and study prior to the workshop. While the text above has been focused on their development, Target PLDs should possess the following characteristics.

1. Target PLDs should be connected to the desired cut score.

During the standard setting workshop participants use the skills found in Target PLDs to predict how students *should* perform on items given effective instruction that is grounded in the standards along with multiple opportunities to learn; thus, the cut scores on the test scale and the PLDs are connected. When the percentage of students who would be classified in each performance level with a particular set of cut scores (known as impact data) is reviewed as a component of the standard setting or following the standard setting and causes a change to

the cut scores, the PLDs may no longer describe the content the cut score represents. That is, the PLDs may become disconnected from the cut scores.

PLDs may become disconnected from the cut scores during the standard setting process because policy experts may be guided by different, yet equally important, considerations from those considered by content experts. Policy considerations often include the across-grade articulation of cut scores so that the performance standards form a cohesive system for accountability purposes, available funding for schools designated as low performing, political implications of cut scores (e.g. would they be considered “easy” compared to NAEP), as well as past performance in the state.

Although it is reasonable to believe there should be agreement across content areas for the sake of consistency regarding whether the performance level of critical importance may mean reaching the standard or going beyond the cognitive rigor of the standard, the use of external benchmarks to inform cut scores during the standard setting process can disconnect the content meaning of a performance level from the original policy intent. Thus, it is possible that after the final approved cut scores are in place, the implied message of meeting or going beyond the standards for the performance level of critical importance can differ by content area.

The separation of the content-based PLDs from the final approved cut scores may mean that the PLDs do not provide valid interpretations regarding what students can do for score reporting purposes or the Range PLD instructional and test development purposes. That is, the PLDs would not be as useful for teachers in shaping instructional actions as more content-referenced PLDs.

Should the PLDs that inform standard setting become disconnected from the final approved cut scores, the Target PLDs should be edited to reflect the content-based changes because these are often subsequently used as Reporting PLDs. Moreover, the items in the standard setting ordered item book (if an item mapping approach is used to recommend standards) and any others found in the item bank should be used to edit and realign the Range PLDs to the test scale. If a benchmark process is going to be used as a component of the standard setting (see Phillips 2012) then the process described in this paper is not optimal and the procedures for developing the Target PLDs should be based upon the test data and the desired cut score rigor of the policy making body at the onset of test development.

2. The Target PLD clearly defines the student it is meant to represent.

A statement should be added to the end of the introductory phrase within each performance level similar to the following: The just (or typical) Below Proficient student is likely able to” This statement is only appropriate if the PLD is meant to be the descriptor used for standard setting (or for reporting).

3. The Target PLDs describe what students should be able to do using observable, measurable skills.

Descriptors that use implicit verbs require that teachers in the standard setting have to infer their meaning. When explicit verbs are used, teachers have a better sense of what the student can actually do (at least within some reasonable probability). For example, writing “The borderline Level 2 student recognizes the main idea” does not lead to a clear conceptualization of what the student should do to demonstrate his or her skill. Rather, “The borderline Level 2 student is able to explain the main idea” clearly describes the evidence from items teachers should find during the review of the test items.

4. The Target PLDs should reflect the content importance found in the test blueprint.

PLDs used for standard setting help support the content-based expectations panelists have of the borderline student in a performance level. Therefore, it is optimal if the proportion of content within a descriptor approximates the proportion of content a student is expected to see on the test (Perie, Hess & Gong, 2008). Panelists who see descriptions that are weighted proportionately different from the test blueprints may hesitate to recommend a cut score until they see more items in the domains that are more heavily described.

5. The Target PLDs should reflect more complex knowledge, skills, and abilities (KSAs) as the performance levels increase.

The reasoning that students use as well as the difficulty of the content should increase as performance levels increase. As students progress in their skills, we should see more difficult work handled independently and without teacher or scaffolding supports.

States that create a set of interwoven PLDs to guide (a) teacher and item writer expectations for the progressions of evidence needed to categorize student performance (b) participant recommendations during the standard setting process, and (c) stakeholder interpretation of student score reports are actively working to enable valid inferences about student content area knowledge and skill in relation to a state’s content standards measured on a large scale assessment. Once the standard setting process is complete, and a state makes edited Range PLDs accessible to teachers and test developers, there is the opportunity for districts within such states to ensure that curriculum and sets of instructional tasks teachers use related to particular standards are aligned to the content, cognitive complexity and context expectations found in the PLDs. States that release examples of tasks that exemplify some or all of the descriptors located within the Range PLD trajectory also support teachers by clarifying the interpretation of the standard by showing the types of skills and tasks that students are expected to do. Administrators, teachers, and parents may use the PLDs to better understand where the student’s current level of performance is during the school year with regard to each standard and what skills the student should demonstrate next.

References

- Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, 6(2), 81-89.
- Egan, K.L., Schneider, M.C., & Ferrara, S. (2012). Performance level descriptors: History, practice and a proposed framework. In G. Cizek (Ed.), *Setting Performance Standards*, Second Edition. 79–106. New York, NY: Routledge.
- Hess, K. & Kearns, J. (2011a). Learning progression frameworks designed for use with the Common Core Standards in Mathematics K–12. Retrieved from http://www.nciea.org/publication_PDFs/Math_LPF_KH11.pdf
- Hess, K. & Kearns, J. (2011b). Learning progression frameworks designed for use with the Common Core Standards in English language arts & literacy K–12. Retrieved from http://www.nciea.org/publication_PDFs/ELA_LPF_12%202011_final.pdf
- Lewis, D. M., & Green, R. (1997, June). *The validity of PLDs*. Paper presented at the National Conference on Large Scale Assessment, Colorado Springs, CO.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13–23.
- Perie, M. (2008). A guide to understanding and developing PLDs. *Educational Measurement: Issues and Practice*, 27(4), 15–29.
- Perie, M., Hess, K., & Gong, B. (March 2008) *Writing performance level descriptors: Applying lessons learned from the general assessment to alternate assessments based on alternate and modified achievement standards*. Paper presented at the Annual Meeting of the National Council on Measurement in Education in New York, NY.
- Phillips, G. W. (2012). The benchmark method of standard setting. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 323–346). New York, NY: Routledge.
- Plake, B. S., Huff, K., & Reshetar, R. (2010). Evidence-centered assessment design as a foundation for achievement-level descriptor development and for standard setting. *Applied Measurement in Education*, 23(4), 342–357.
- Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26(2), 114-145.

- Smith, C. L., Wiser, M., Anderson, C. W. & Krajcik, J., (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic molecular theory. *Measurement: Interdisciplinary Research and Perspectives*, 4(1&2). 1-98.
- Schneider, M.C., Huff, K.L., Egan, K.L, Gaines, M.L., & Ferrara, S. (2013). Relationships among item cognitive complexity, contextual response demands, and item difficulty: Implications for achievement level descriptors. *Educational Assessment*, 18(2). 99–121.
- Webb, N. L. (2005). *Web alignment tool (WAT): Training manual*. Draft Version 1.1. Wisconsin Center for Education Research, Council of Chief State School Officers. Retrieved on September 15, 2005, from <http://www.wcer.wisc.edu/wat/index.aspx>.
- Webb, N. M., Herman, J. L., & Webb, N. L. (2007). Alignment of mathematics' state level standards and assessments: The role of reviewer agreement. *Educational Measurement: Issues and Practice*, 26, 17–29.