2012 NeSA-Science Standard Setting Technical Report

June 26, 2012

Prepared by
Data Recognition Corporation

DATA RECOGNITION
DRC
CORPORATION


DATA RECOGNITION
DRC
CORPORATION

TABLE OF CONTENTS

1. Executive Summary

Academic Performance Levels for the science component of the Nebraska State Accountability assessments (NeSA-Science) were developed in spring 2012 by establishing cut scores that define operationally the three Performance Levels: *Below the Standards, Meets the Standards, Exceeds the Standards*. These Performance Level designations will be used by local, state, and federal accountability programs and are central to communicating to parents, teachers, and the public. The *Meets the Standards* and *Exceeds the Standards* levels are used for the *No Child Left Behind (NCLB)* Adequate Yearly Progress (AYP) proficiency goal.

The larger process comprised four events. First, a meeting was held April 12, 2012 with the Nebraska State Board of Education and other stakeholders to introduce the process and obtain feedback to ensure an effective, defensible process. Second, a *Contrasting Groups* survey of science specialists and teachers was conducted in spring 2012 to obtain the teachers' overall perception of the proficiency level of their own students, independent of the state assessment. Third, a *Bookmark* Standard Setting was conducted June 26, 2012 in Lincoln, Nebraska, after the operational data were available. Finally, recommendations of the Contrasting Groups and Bookmark processes were presented to the State Board of Education July 10–11, 2012. The purpose of this meeting was for the State Board of Education to formally establish the Performance Levels. This report specifically documents the Bookmark and Contrasting Groups portions of the process.

The Bookmark method (*Lewis, Mitzel, & Green, 1996*) is, perhaps, the most philosophically consistent with criterion-referenced, standards-based[1] assessments like the NeSA. Bookmark is an *item-based* method. It requires panelists to determine which items can be successfully answered 67% of the time by students at the Performance Level boundaries. The Contrasting Groups method (*Cizek & Bunch, 2007, chapter 8*) is *student-based* which asks teachers to place students into one of the three Performance Levels based on their knowledge of the students from their classrooms without considering the assessment. The success of either approach requires an in-depth understanding of the skills and knowledge required at each level. This shared understanding is expressed in *Performance Level Descriptors* (Appendix A).

To assist the State Board of Education in determining appropriate cut scores, DRC presented the results of both studies, the Bookmark and the Contrasting Groups, as well as a composite of the two studies for consideration. The State Board of Education approved cut scores that were closest to the composite of the two studies.

[1] The term *standard* is used in two different senses in this area. *Content standards* are written descriptions of the goals and expectations for learning and instruction at each grade level. *Performance standard*s, which are the focus of this report, define the levels of achievement necessary for each Performance Level. In some contexts, the term *performance standard* is interchangeable with *cut score*.

Board-Approved Cut Scores

The final State Board of Education approved cut scores and the percentage of spring 2012 students expected to be in each Performance Level are shown in Table 1.1.1. These values in the scale score metric will be used for all grades and will not change from year to year. The *Raw Score Ranges* may vary from year to year, depending on the difficulty of the specific form, and the *Percent in Each Performance Level* will vary, depending on the proficiency of the students at that time.

Table 1.1.1 State Board of Education Approved Standard Setting Results

	Scale Scores			Logit Cut points		Raw Score			Percent in Each Performance Level		
Grade	Below	Meets	Exceeds	B/M	M/E	Below	Meets	Exceeds	Below	Meets	Exceeds
5	1 to 84	85 to134	135 to200	-0.4971	1.0580	1 to 30	31 to 43	44 to 50	33.6%	52.0%	14.4%
8	1 to 84	85 to134	135 to200	-0.4543	1.0378	1 to 35	36 to 50	51 to 60	32.4%	52.1%	15.4%
11	1 to 84	85 to134	135 to200	-0.5407	1.3130	1 to 32	33 to 51	52 to 60	33.0%	54.0%	13.1%

Cut scores are defined in a logit metric, which, like scale scores, are also fixed. Logits are related to percentage correct scores but are preferred because they are not tied to a specific test form and will not change from year to year. This ensures a consistent definition of the Performance Levels even if different test forms vary in difficulty. For reporting purposes, logits are converted into the scale scores, which is mathematically equivalent but more user-friendly.

The meaning of the logit and scale score values will not change in the future, but the raw score ranges may shift slightly to reflect the variation in item and form difficulty; a more difficult form will require fewer correct responses and an easier form will require more. With a stable scale score cut point, changes in the percentage of students in each proficiency level will reflect changes in student proficiency and not changes in form difficulty.

2. Introduction

2.1 Background

In January 2009, the Nebraska Department of Education contracted with Data Recognition Corporation (DRC) to provide and operate a computerized information system to support the administration, record keeping, and reporting for statewide student assessment and accountability under the direction of the Department of Education.

NeSA Content Areas and Grade Levels: Legislative Bill (LB) 1157 (http://uniweb.legislature.ne.gov/FloorDocs/Current/PDF/Slip/LB1157.pdf), passed by the 2008 Nebraska Legislature, requires a single statewide assessment of the Nebraska academic content standards for writing, reading, mathematics, and science in Nebraska's K-12 public schools. The new assessment system is named NeSA (Nebraska State Accountability) with NeSA-Reading for reading assessments, NeSA-Mathematics for mathematics assessments, and NeSA-Science for science assessments. The NeSA-Science assessments were administered operationally in grades 5, 8 and 11 for the first time in the spring of 2012.

Phase-In Schedule for NeSA: The Nebraska Department of Education prescribed the assessments starting in the 2009-2010 school year to be phased in as shown in Table 2.1.1. The state used the expertise and experience of in-state educators to participate in the design and development of the new statewide assessment system. The Nebraska Department of Education developed the NeSA-Reading, NeSA-Mathematics, and NeSA-Science tests for use in the state accountability system and was charged with setting student academic Performance Level standards on the NeSA-Reading, NeSA-Mathematics, and NeSA-Science tests.

Table 2.1.1: NeSA Administration Schedule

Content Area	Administration Year		Grades
	Field Test	Operational	
Reading	2009	2010	3 through 8 and one high school grade
Mathematics	2010	2011	3 through 8 and one high school grade
Science	2011	2012	5, 8, 11

The Nebraska Department of Education required standard setting procedures to determine student academic Performance Levels for the NeSA-Reading and NeSA-Mathematics assessments administered to each of grades 3 through 8 and 11and the NeSA-Science assessments to each of grades 5, 8 and 11. DRC, with the assistance of the Nebraska Department of Education, organized and facilitated the Standard Setting events.

For all NeSA assessments, there are three student Performance Levels: *Below the Standards, Meets the Standards,* and *Exceeds the Standards,* requiring two cut points. For federal reporting

purposes, *Proficiency* is defined as students performing at *Meets the Standards* and *Exceeds the Standards* levels.

2.2 Purpose and Objectives of NeSA and Standard Setting

NeSA tests measure student performance on the state-adopted academic standards to:

1. promote student learning,
2. identify areas in which students, schools, or school districts need additional support;
3. indicate the academic achievement for schools, districts, and the state;
4. satisfy federal reporting requirements; and
5. provide professional development to educators.

The results from NeSA-Science tests were used for reporting annual state, school, and district end-of-year performance on science standards.

Many Standard Setting methods have been proposed. These fall into two major approaches:

1. *Item-based*, which focus on what knowledge, skills, and behaviors are required to successfully respond to an item, and
2. *Student-based*, which focus on what proficiencies individual students possess.

For the NeSA, both approaches were used.

2.3 Bookmark Standard Setting Method

DRC followed a Bookmark procedure similar to the method suggested by Lewis, Mitzel, and Green (1996). Bookmark is one in a broad category of methods commonly referred to as item mapping, which focus on items rather than examinees. The essential task is to identify the items that can be answered successfully (67% likelihood) by students at the boundaries of the Performance Levels. The logit difficulty value that separates the items that borderline students can do from those they cannot do, establishes the Bookmark cut score.

All panelists were trained in a large group prior to breaking into smaller working groups. Training covered the following points:

- The Performance Levels are defined and described by the Performance Level Descriptors developed by the state with advice from Nebraska teachers and other content specialists.
- The task for the panelist is to place a bookmark between items that students at the threshold of a Performance Level have mastered and those not yet mastered.
- Students at a given cut score will have a 0.67 probability of correctly responding to a multiple-choice item at the cut score. These students will have a higher probability of success on easier items (before the bookmark) and a lower probability of success on harder items (after the bookmark).

- In placing their bookmarks, the task was to consider what students *should* know and be able to do as defined by the Performance Level Descriptors and the item content.
- Panelists were instructed to first place the bookmark separating *Below the Standards* from *Meets the Standards* levels and then place the bookmark separating *Meets the Standards* from *Exceeds the Standards*.
- Panelists were asked to record their bookmark placements on a rating form. The placements were entered into a spreadsheet program, and the median cut score was calculated for the full panel.

To begin the process, participants were asked to visualize the knowledge and skills of a student who is at the borderline between two Performance Levels based on the Performance Level Descriptors. Participants were given an Ordered Item Booklet with items ordered from least to most difficult. Panelists were also provided with supporting materials for each item including the correct response, content objective, and item sequence in the test booklets.

The task for the panelist was to proceed through the Ordered Item Booklet and ask, for each item, if the borderline student could answer correctly. Each panelist placed a bookmark in front of the page in the booklet where the borderline student had not mastered the item. *Mastery* was defined as having at least a 67% likelihood of responding correctly.

The DRC adaptation of the Bookmark procedure involved three rounds of deliberation, discussion, and feedback. These iterations are described in more detail in Section 4.

2.4 Contrasting Groups Standard Setting Method

The examinee-based Contrasting Groups (Cizek & Bunch, 2007) survey was included to complement the item-based Bookmark method. The survey asked the teachers to evaluate each student with whom they were familiar and indicate which Performance Level best described the student. The survey was conducted prior to the first operational administration of the NeSA-Science, so ratings would be determined by the teachers' firsthand experience with the students in the classroom, not students' performance on the test. All science teachers and specialists in Nebraska were invited to participate in the survey.

The survey was distributed online. Teachers first selected students from a roster for their own school excluding students for whom they were unfamiliar or uncertain. The instructions emphasized the importance of knowing the student and the student's status. Teachers were encouraged to omit ratings for any student for whom the teacher did not have firsthand knowledge.

The results of the survey were summarized, provided to the Bookmark panels after the initial round, and presented to State Board of Education as part of the final cut score recommendations.

2.5 Meetings with a Committee of Stakeholders and State Board of Education

DRC presented to a subgroup of Board Committee members, media, and other stakeholders on April 12, 2012. The April meeting introduced the process to the stakeholders to familiarize them

with the Standard Setting process and obtain their reactions. DRC presented an overview of the Standard Setting procedures and outlined the appropriate interpretation of the results from the studies. There was discussion of the information needed and effective methods for its interpretation.

The State Board of Education considered the results from the two standard setting events and recommendations from the Nebraska Department of Education during the July 2012 State Board of Education meeting. The goal was to formally adopt a motion establishing proficiency level cut scores for the NeSA-Science assessments based on this impact.

3. Preparation for Standard Setting

In April 2012, a Bookmark Standard Setting plan proposed by DRC was reviewed and approved by the Nebraska Department of Education and its Technical Advisory Committee. The plan described the purpose of the meeting, specifications of panelists, methodology, and potential consequences related to accountability. This section provides an overview from the plan.

3.1 Bookmark Panelist Recruitment

The Nebraska Department of Education recruited panelists for the Standard Setting process:

- In January of 2012, Dr. Pat Roschewski communicated with District Assessment Contacts, informing them of the plan for establishing NeSA-Science cut scores and the need for Nebraska educators to participate in the process.
- Information regarding the Standard Setting process was communicated to Nebraska districts in *Standards, Assessment, and Accountability Updates*.
- The Statewide Assessment Office sought nominations for participation in the Standard Setting process.
- Statewide Assessment Office members reviewed the nominations and selected participants. Three criteria were considered:
    1. Educational role.
    2. Geographic location.
    3. Knowledge and experience with the NeSA-Science.
- Applicants were notified by the Statewide Assessment Office of their selection status.

A total of 33 panelists participated in the Bookmark event. Table 3.1.1 summarizes information about characteristics of the participating panelists based on their self-reported responses to the Participant Survey. Most panelists were classroom teachers; a few were non-teacher educators, and the majority was female.

Table 3.1.1 Panelist Summary

Demographic		Science
Grade teacher reported*	5	13
	8	10
	11	9
Gender	Male	10
	Female	23
Ethnicity	White/non-Hispanic	31
	Latino/Hispanic	1
	Multi-racial/Ethnic	1
Role	Other	0
	Teacher	31
	Educator	2
Region*	Rural	14
	Urban	8
	Suburban	9
Experience	0 - 5 years	5
	6 - 10 years	7
	11 - 15 years	5
	16 – 20 years	6
	21 – 25 years	5
	26 – 30 years	4
	31 – 35 years	0
	> 36 years	1

* Not all panelists responded to this question

3.2 Roles and Responsibilities

A successful Standard Setting requires the concerted and coordinated efforts of many people including staff from the Nebraska Department of Education and DRC, and, most importantly, the panelists. Each group has its unique and critical roles and responsibilities.

Panelists—brought their individual educational experience and expertise about Nebraska students, science instruction, and the Nebraska curriculum. Their knowledge of science instruction and curriculum in Nebraska and their familiarity with Nebraska students forms the foundation for the validity of the performance standards.

Nebraska Department of Education—The Nebraska Department of Education staff convened the meeting and introduced the NeSA-Science program and the importance of Standard Setting.

The Nebraska Department of Education staff monitored the progress of each panel and fielded questions on the assessment and test content and on any policy concerns.

DRC Staff—facilitated the sessions and provided logistical and technical support.

Psychometric Lead—conducted the training session and monitored progress and results throughout.

Test Development Specialist—assisted as needed with the Performance Levels and covered questions about test content.

Project Management—maintained security of materials through check-in and check-out procedures, liaison with hotel facility staff, and overall coordination of meeting logistics.

Room Facilitators—reviewed procedures for the panelists, kept the process moving on schedule, explained results, and facilitated the sessions.

Statistical Analyst—entered the panelists' bookmark ratings and performed the necessary statistical analyses.

3.3 Materials Preparation

Workshop materials were prepared by DRC. The materials available to panelists during the workshop included:

- Training Materials
- Operational Test Forms
- Ordered Item Booklet
- Performance Level Descriptors
- Item Map
- Item Separation Map
- Participant Rating Forms

Training materials comprised a much reduced test and related materials that were otherwise identical to the materials to be used in the actual process. The training materials were based on released items and item data from the Nebraska item bank.

Science Performance Level Descriptors were originally developed by the Nebraska Department of Education with assistance from educators. A complete statement of the Performance Level Descriptors is included in Appendix A.

3.4 Ordered Item Booklet

The critical information was in the Ordered Item Booklet. Each Ordered Item Booklet contained all items in the grade in order of item difficulty from least to most difficult, based on item difficulties obtained from the spring 2012 NeSA-Science administration. Table 3.4.1 displays the number of items/score points per grade on the operational forms. Item Separation Charts for each grade are included in Appendix E.

Table 3.4.1: Number of Score Points in Ordered Item Booklet

Content	Grade	No. of Score Points in the OIB
Science	5	50
	8	60
	11	60

The task presented to the panelists was to identify the item in the Ordered Item Booklet for which the student on the boundary between two Performance Levels can no longer answer the item correctly with reasonable certainty. The required level of mastery was defined operationally as a probability of success of 0.67. With the Rasch model, the choice of the mastery level does not affect the ordering of the items, but it does affect which scale score aligns with the bookmarked item.

The Rasch model for dichotomous items (Wright & Stone, 1979) defines the probability of success as:

1. $p = \frac{e^{b-d}}{1+e^{b-d}}$.

With a little algebra, p = 0.67 implies the logit cut score is shifted by *0.69 logits* from the logit difficulty of the bookmarked item:

2. $(b-d) = ln\frac{0.67}{1-0.67} = ln(2) = 0.69$.

4. Standard Setting Procedures

4.1 Contrasting Group Procedures

An examinee-based Contrasting Groups survey was included to complement the item-based Bookmark method. All Nebraska science teachers were invited to participate in the survey, which was presented online. The task for the teachers was to evaluate each student with whom the teacher was familiar and indicate the Performance Level that best described the student. The survey was conducted prior to the first operational administration of the NeSA-Science, so ratings were determined by the teachers' firsthand experience with the students in the classroom, not their performance on the test. The Performance Levels were defined by the Performance Level Descriptors, which were available online for review at any point in the process.

The teachers were asked to select students from their own classes and schools. The instructions emphasized the importance of knowing the student and the student's status. Teachers were encouraged to omit ratings for any students for whom they did not have firsthand knowledge.

Recruitment: In January 2012, the Nebraska Department of Education and DRC contacted Nebraska District Assessment Coordinators (DAC) to solicit their cooperation in the study that would bring teachers' knowledge of science instruction and an understanding of their students together. The DAC were first asked to provide contacts for these science teachers and specialists.

In February 2012, DRC sent an initial invitation to teachers. This invitation asked for their participation in an online study that would use their professional judgment to help establish the Performance Levels for the NeSA-Science. The estimated time for completing the survey was less than 30 minutes; all responses were confidential. Potential participants were also given the schedule for the survey and the training sessions.

A follow-up email with the online conference dates (via WebEx™), sign-on instructions, times available, and information about DRC's online delivery system was sent to the participating teachers on March 1, 2012.

Training: DRC hosted seven online conferences to introduce teachers to the online Contrasting Groups survey. The online conferences were interactive, allowing teachers to pose questions and seek immediate clarification. Typically, the sessions lasted fifteen to twenty minutes. For teachers who were unable to attend any of the online conference sessions, the Nebraska Department of Education placed the training materials on its website on March 14, 2012.

The training covered the details of navigating the survey website, saving the work, returning after interruptions, and submitting the ratings. Each teacher was asked to:

- Use the school and district rosters provided to create a personal class roster with 25-30 students representing all Performance Levels.
- Follow the instructions repeated at the top of each page of the survey.
- Read and refer back to the Performance Level Descriptors in the course of the survey.
- Complete the survey as soon as possible after training, but no later than March 23, 2012.

Table 4.1.1: Online Conference Training Schedule

SESSION	DATE	TIME
1	Wednesday, March 7,2012	7:00 – 7:30 AM
2	Wednesday, March 7, 2012	3:00 – 3:30 PM
3	Thursday, March 8, 2012	4:00 – 4:30 PM
4	Friday, March 9, 2012	3:00 – 3:30 PM
5	Monday, March 12, 2012	2:30 – 3:00 PM
6	Monday, March 12, 2012	3:30 – 4:00 PM
7	Tuesday, March 13, 2012	2:00 – 2:30 PM

The instructions reminded teachers that they should not include students with whom they had little experience, nor did they need to rate students, even if selected, if they were uncomfortable assigning the student to a Performance Level for any reason.

Survey Results: A total of 188 teachers participated in the survey. The initial target number was 100 per grade. Feedback from the participants indicated the task was easier and took less time than they expected. A brief survey soliciting teachers' opinions on the Contrasting Groups task was requested and results are presented in Appendix J. The participation breakdown by grade is given in Table 4.1.2.

Table 4.1.2: Contrasting Groups Participation by Grade

Grade	Number of Teachers	Number of Students Rated
5	70	2612
8	53	3028
11	65	2293
Total	**188**	**7933**

Appendix F provides detailed summaries of the survey, including student breakouts by gender, ethnic group, teacher rating, and performance level.

4.2 Modified Bookmark Procedure

The Bookmark process, including training, was completed on Tuesday, June 26, 2012. The outline and agenda for the Bookmark event are presented in Appendix B.1. The teachers were placed in three grade-grouped panels: lower, middle, and high school. The intent of the grade groupings was to ensure panelists worked with content with which they were familiar while giving each panel more breadth, and the result more continuity across grades. The groupings and timing are diagrammed in Appendix B.2.

Training was conducted with a single trainer for a single large group of the three panels. Training materials included:

- Performance Level Descriptors
- Sample Ordered Item Booklet

- Sample Item Map
- Sample Item Separation Chart
- Sample Rating Form

Participants were told that:

- Their bookmark placement should reflect their own opinions and not the group consensus;
- They should contribute their own personal experience and expertise to the group discussion and recommendation;
- They would have the opportunity to discuss, reconsider, and revise their placements in later rounds, and
- All materials and discussions were secure and were not to leave the meeting room.

The critical objective of the training was to ensure the panelists understood the task being presented to them. Components included an overview of their role in the process, a detailed description of all steps in the Bookmark method, and a practice exercise based on a short test form drawn from released NeSA-Science items. The point of the practice exercise was to provide hands-on experience with the steps and allow the panelists to receive any additional explanation they needed. A copy of the slides used for training is presented in Appendix C.

The actual Bookmark process included three iterations (rounds) of individual judgments, large group discussions between rounds, and opportunities to revise individual judgments. After the first and second rounds, panelists had the opportunity to review impacts in the form of percentage of students in each Performance Level, resulting from the group recommendation. In addition, panels for grade 8 were shown relevant NAEP statistics.

After the training and practice exercise, the panelists broke into the smaller groups by grade. The process began with the panelists working through the spring operational form of NeSA-Science. This task was included to give panelists a direct appreciation of the students' NeSA-Science experience. They were encouraged to take notes concerning their impressions of the items. Then a review of the Performance Level Descriptors specific to that grade was provided to sharpen the understanding of what was expected of students at each level. Panelists were encouraged to highlight the language differentiating the Performance Levels. After a short discussion and clarifications, the actual Bookmark placement work began.

Round 1. In Round 1, participants reviewed the Ordered Item Booklets independently to ensure the initial bookmarks were independent of other panelists' opinions. During this review, panelists were asked to determine the knowledge, skills, and competencies required to respond correctly to each progressively more difficult item and when the requirements of the items exceeded the capabilities of the borderline students. It was emphasized that the work for this round was to be individual.

The bookmarks were to be placed so that the borderline student has mastered the items before the bookmark and not those after the bookmark. To reduce counter-productive discussion about the placement of specific items in the Ordered Item Booklet, panelists were reminded that the placement was empirical based on the spring assessment and that they should focus on the progression of items rather than the details of individual items.

Round 2. The results from Round 1 were presented and explained at the beginning of Round 2. The bookmark page numbers for each panelist, the median page number of the full panel, the distribution of cut scores for each Performance Level, and the impact data were reviewed with the panelists. The impact data was the percentage of students placed in each Performance Level based on 2012 NeSA-Science student performance and panelists' Round 1 recommendations. Panelists were then asked to provide rationales for their Round 1 placements and discuss what skills and knowledge were required. During the discussion, there was no attempt to achieve consensus; the bookmark placements were to reflect the opinions of the individual panelists.

After the group discussion, panelists were given the opportunity to revise their bookmark placements. The individual locations were again collected and used to calculate revised cut scores and impact data for the full panel.

Round 3. Panelists reviewed Round 2 results and the relevant Contrasting Groups data. When applicable to the grade, the NAEP (grade 8) data were also provided. Again, panelists were instructed to explain the thinking for their Round 2 placements in terms of the skills and knowledge required. Following the discussion, the panelists made any final adjustment to their individual placements. These ratings were recorded and used to produce the final group recommendation.

4.3 Merging Bookmark and Contrasting Groups

The item-based Bookmark method was the designated method of record. The Bookmark results were the crux of the recommendation to the State Board of Education. The recommendation was developed by experts on education in Nebraska, primarily classroom teachers, from their understanding of the Performance Level Descriptors and their assessment of the knowledge, skills, and behaviors required by the operational items; and after receiving extensive training on the process and the Performance Level Descriptors.

The Contrasting Groups survey involved a different sample from the same population of experts. The focus for this method was on students known to the teacher and on the Performance Level best describing each of those students, independent of any assessment. The Performance Level Descriptors were available on demand as a pop-up for the participants in the Contrasting Groups, and there was group (online) training to ensure a common understanding of the Performance Level Descriptors.

The final recommendation to the State Board of Education was based on a composite that used both sets of data.

5. Analyses and Results

5.1 Overview

Summaries of the NeSA-Science Performance Level Standard Setting process are provided in Tables 5.1.1-4. The tables include the four options discussed with the State Board of Education.

1. Bookmark
2. Contrasting Groups
3. Average of Bookmark and Contrasting Groups
4. Board Approved

The scale score metric is the most public and is used for all reports. The minimum scale score for Meets the Standards was set to 85 and for Exceeds the Standards to 135, matching the values established in 2010 for NeSA-Reading and 2011 NeSA-Mathematics. These were derived from the logit standards following Standard Setting so that the logit standards approved by the State Board of education translate to scale scores of 84.5 and 134.5, respectively.

The raw score ranges are specific to 2012 exam and will vary slightly from year to year with minor differences in form difficulty. For each Performance Level, the minimum raw score is the lowest score for which the corresponding logit is greater than or equal to the logit standard for that level. This determination is made in the logit metric to avoid rounding issues.

The logit metric is the native Rasch metric and is the basis for all calculations beginning with the construction of the Ordered Item Booklets and the derivation of the standards from the panelists' recommendations.

The final table in this section, 5.1.4, presents the 2012 impacts (percent in each Performance Level) of the four options discussed. These impacts were the focus of the State Board of Education's attention.

Table 5.1.1: Scale Score Ranges by Performance Level for Four Options

	Option 1 – Bookmark (BMK)			Option 2 - Contrasting Groups (CG)			Option 3 - Average of BMK & CG			Option 4 – Board Approved		
	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds
5	1 to 95	96 to 122	123 to 200	1 to 74	75 to 145	146 to 200	1 to 84	85 to 134	135 to 200	1 to 84	85 to 134	135 to 200
8	1 to 84	85 to 122	123 to 200	1 to 84	85 to 151	152 to 200	1 to 84	85 to 134	135 to 200	1 to 84	85 to 134	135 to 200
11	1 to 83	84 to 109	110 to 200	1 to 84	85 to 148	149 to 200	1 to 84	85 to 130	131 to 200	1 to 84	85 to 134	135 to 200

Table 5.1.2: Raw Score Ranges by Performance Level for Four Options

	Option 1 – Bookmark (BMK)			Option 2 - Contrasting Groups (CG)			Option 3 - Average of BMK & CG			Option 4 – Board Approved		
	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds
5	1 to 33	34 to 40	41 to 50	1 to 26	27 to 44	45 to 50	1 to 30	31 to 43	44 to 50	1 to 30	31 to 43	44 to 50
8	1 to 35	36 to 47	48 to 60	1 to 35	36 to 53	54 to 60	1 to 35	36 to 50	51 to 60	1 to 35	36 to 50	51 to 60
11	1 to 31	32 to 43	44 to 60	1 to 32	33 to 54	55 to 60	1 to 32	33 to 50	51 to 60	1 to 32	33 to 51	52 to 60

Table 5.1.3: Logit Performance Standards for Four Options

	Option 1 – Bookmark (BMK)			Option 2 - Contrasting Groups (CG)			Option 3 - Average of BMK & CG			Option 4 – Board Approved		
	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds
5		-0.133	0.706		-0.861	1.410		-0.497	1.058		-0.497	1.058
8		-0.422	0.681		-0.486	1.394		-0.454	1.038		-0.454	1.038
11		-0.557	0.399		-0.524	1.689		-0.541	1.044		-0.541	1.313

Table 5.1.4: Percent 2012 NeSA-Science Students by Performance Level for Four Options

	Option 1 – Bookmark (BMK)			Option 2 - Contrasting Groups (CG)			Option 3 - Average of BMK & CG			Option 4 – Board Approved		
	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds	Below	Meets	Exceeds
5	44.6%	29.0%	26.4%	21.1%	68.1%	10.8%	33.6%	52.0%	14.4%	33.6%	52.0%	14.4%
8	32.4%	41.6%	26.0%	32.4%	60.2%	7.4%	32.4%	52.1%	15.4%	32.4%	52.1%	15.4%
11	30.6%	32.7%	36.7%	33.0%	61.4%	5.6%	33.0%	51.2%	15.8%	33.0%	54.0%	13.1%

5.2 Contrasting Groups Analyses

The Contrasting Groups method asked teachers to evaluate students in their own classes and assign Performance Levels to each based on the Performance Level Descriptors without considering performance on the NeSA. After the assessment, two pieces of data were available about students who had been rated on the Contrasting Groups survey: first, the Performance Level assigned by the teacher and second, the observed number correct on the NeSA-Science. There was a strong relationship between these two pieces of data: students with a low number correct tended to be placed in the *Below the Standards* level and students with high number correct scores in the *Exceeds the Standards* level. Table 5.2.1 shows the relevant portion for the grade 5 NeSA-Science data. This table tabulates the number of students at each NeSA-Science number correct score that teachers assigned to each Performance Level.

Table 5.2.1: Extracted from Grade 5 Contrasting Groups Results

Number Correct	Logit Ability	Teacher Rank			B+M	M+E	Log Odds	
		Below	Meets	Exceeds			Meets	Exceeds
25	-0.994	29	28	1	57	29	0.00	-1.76
26	-0.901	31	22	1	53	23	-0.13	-1.72
27	-0.808	26	30	2	56	32	0.09	-1.45
28	-0.715	32	33	2	65	35	0.04	-1.51
29	-0.622	23	38	6	61	44	0.28	-1.01
30	-0.527	26	43	4	69	47	0.26	-1.24
31	-0.431	38	60	4	98	64	0.23	-1.39
32	-0.334	35	54	8	89	62	0.25	-1.05
33	-0.235	37	48	3	85	51	0.14	-1.45
34	-0.133	23	74	10	97	84	0.56	-0.99
35	-0.028	21	70	7	91	77	0.56	-1.11
36	0.080	25	78	19	103	97	0.59	-0.73
37	0.192	14	77	16	91	93	0.82	-0.75
38	0.310	15	96	14	111	110	0.87	-0.90
39	0.434	9	99	17	108	116	1.11	-0.80
40	0.565	12	82	31	94	113	0.97	-0.48
41	0.706	9	90	33	99	123	1.14	-0.48
42	0.858	3	81	50	84	131	1.64	-0.23
43	1.026	3	79	42	82	121	1.61	-0.29
44	1.214	1	55	56	56	111	2.05	0.00
45	1.429	2	44	50	46	94	1.67	0.04
46	1.685		45	53	45	98		0.07

The estimated cut score for each Performance Level from the Contrasting Groups survey is the point on the scale for which the likelihood of the higher Performance Level(s) surpasses the likelihood of the lower level(s). For example, of the students rated in the Contrasting Groups survey, 58 students have a number correct score of 27. Of these, 26 were rated as *Below the Standards*, 30 as *Meets the Standards*, and two as *Exceeds the Standards*. The log odds of belonging to *Meets or Exceeds the Standards* rather than *Below the Standards*, given a number correct of 27, is

$$\log(\frac{(30+2)}{26}) = 0.09$$

In this example, the odds (second column from right) of a student being at the *Meets or Exceeds the Standards* level instead *of Below the Standards* becomes greater than zero at a Number Correct score of 27. This means that the likelihood of level *Meets or Exceeds the Standards* becomes more likely than the level *Below the Standards* between raw scores 26 and 27, which correspond to logits of -0.901 and -0.808. Similarly, the line between *Meets the Standards* and *Exceeds the Standards* falls between raw scores 44 and 45. The logit cut-point is in the range of 1.214 and 1.429. There is some ambiguity about the exact logit value of the cut score because the exact point will fall between two raw scores and because there will typically be some fluctuation in the observed counts. These cases can be resolved using a combination of interpolation and smoothing.

This is illustrated graphically for grade 5 in Figure 5.2.1 below. The scale score cut point between *Below the Standards* and *Meets the Standards* is the point at which the red line crosses the blue line. For *Meets the Standards* and *Exceeds the Standards*, it is the point at which the green line crosses the red line. No number correct score in general will pass exactly through the intersection of the two curves. A very good approximation to the intersection can be obtained by a simple linear interpolation. Similar graphs are presented for grades 8 and 11 in Figure 5.2.2 and 5.2.3. Detailed tables of the Contrasting Groups results are presented in Appendix G.

Figure 5.2.1: Grade Five Contrasting Groups Results

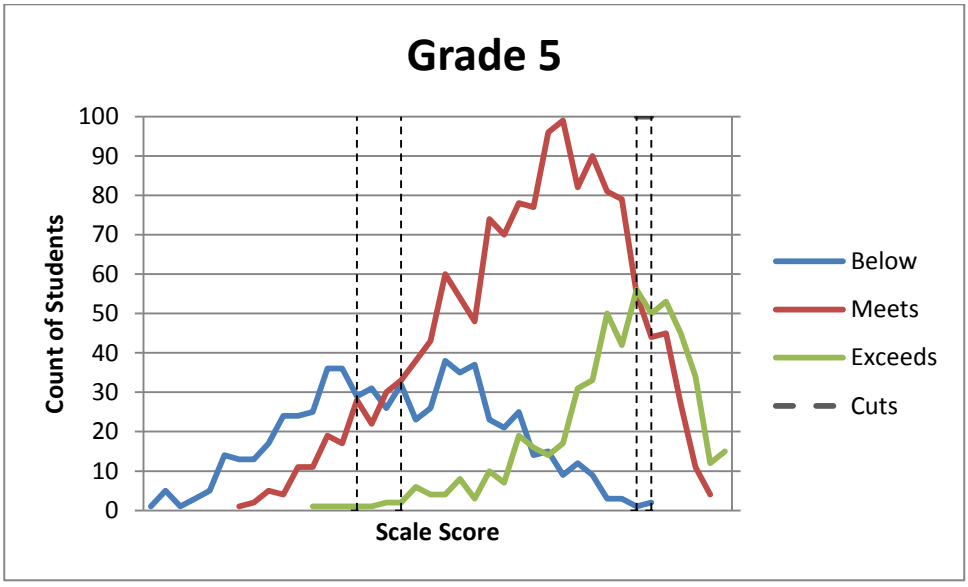


Figure 5.2.2: Grade Eight Contrasting Groups Results

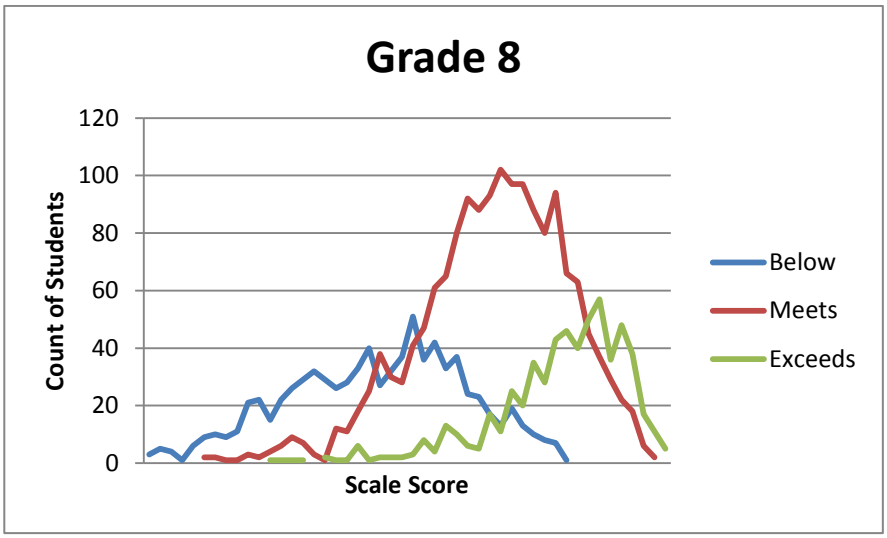
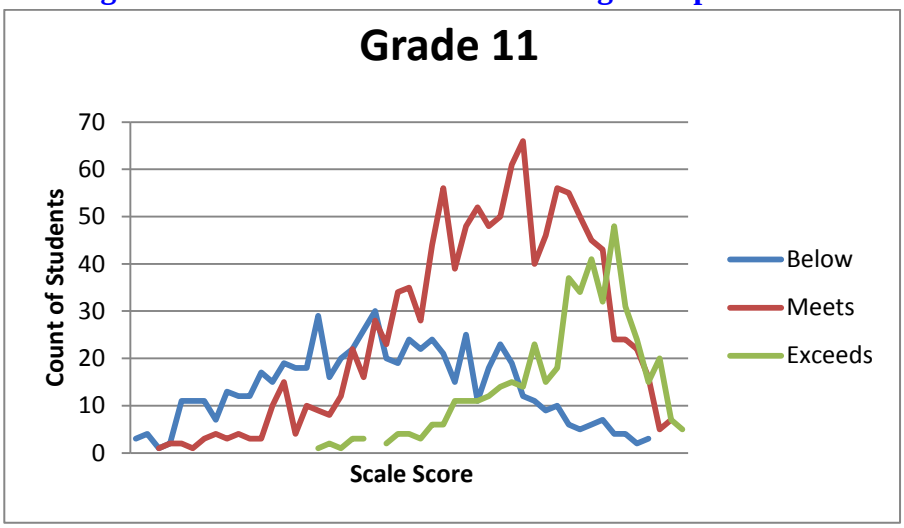


Figure 5.2.3: Grade Eleven Contrasting Groups Results



5.3 Bookmark Analyses

The Bookmark method asks the panelists to move through the items in order of increasing difficulty and place a bookmark between the pages separating items the borderline student has mastered from items this student has not mastered. The borderline student is a student whose proficiency just surpasses the Performance Level Descriptors for the lower level. Mastery was defined as a 0.67 likelihood of succeeding on the item. For the NeSA, two bookmarks are required: between *Below the Standards* and *Meets the Standards* and between *Meets the Standards* and *Exceeds the Standards*. Proficiency levels are again defined by the Performance Level Descriptors, given in Appendix A, as they were for Contrasting Groups.

The bookmarks placed by the panelists are summarized in Table 5.3.1. The values are page numbers in the Ordered Item Booklets.

Table 5.3.1: Bookmark Page Number Medians and Standard Errors

	Number of Panelists	Round 1		Round 2		Round 3	
		B/M	M/E	B/M	M/E	B/M	M/E
Grade 5	13						
Median		13	33	14	33	23	44
Std Dev		4.1	5.7	3.2	3.9	3.5	4.6
SE (med)		1.4	2.0	1.1	1.4	1.2	1.6
Grade 8	11						
Median		15	38	16	40	23	50
Std Dev		4.2	6.4	2.9	2.9	3.7	4.2
SE (med)		1.6	2.4	1.1	1.1	1.4	1.6
Grade 11	9						
Median		15	43	15	43	15	47
Std Dev		8.1	6.2	5.6	3.4	3.4	2.2
SE (med)		3.4	2.6	2.3	1.4	1.4	0.9

Each page number in the Ordered Item Booklet represents an item location and the item has a logit difficulty estimate. Any logit difficulty can be translated into a logit ability corresponding to the 0.67 likelihood. The Rasch probability will be 0.67 when the person's ability exceeds the item's difficulty by natural log of 2 because:

1. $Prob(correct) = \frac{e^{b-d}}{1+e^{b-d}} = \frac{e^{ln2}}{1+e^{ln2}} = \frac{2}{1+2} = 0.67.$

Consequently, the upper bound on the recommended logit cut score is the minimum logit ability that is higher than the bookmarked item's difficulty plus *ln(2) = 0.693*. This is an upper bound because the bookmark is actually placed before the item's page in the Ordered Item Booklet and all that is known about the panelist's implied standard is that it is no higher than the bookmarked item. The logit is rounded up again to align with the minimum raw score on the operational test that is equal to or greater than the logit implied by the bookmark.

5.4 Recommendation and Approval of State Board of Education

The recommended Bookmark cut scores involved two additional considerations: the trends across years and the standard errors of measurement. The trend across years was introduced to maintain a coherent progression of percentage at or above a level from grade to grade. The standard errors of measurement reflected the variability in the testing process and were used to restrict the size of the adjustments made for cross-year smoothing. This was used to develop recommendations that were consistent with all information provided by teachers and panelists.

The State Board of Education reviewed the results from both the Bookmark and Contrasting Groups studies. DRC also presented the third option of a simple, unweighted average of the logit cuts from the two studies. The average was computed in the logit metric and the percentage of students in each Performance Level determined from the raw-to-logit conversion table and observed frequencies. Summary values for the cut scores and impacts are shown in Table 5.4.1.

Table 5.4.1: Logit and 2012 Raw Score Cut points for NeSA-Science

	Logit Cut points		2012 Raw Score Ranges by Performance Level			Percent in Each Performance Level		
Grade	B/M	M/E	Below	Meets	Exceeds	Below	Meets	Exceeds
5	-0.4971	1.0580	1 to 30	31 to 43	44 to 50	33.6%	52.0%	14.4%
8	-0.4543	1.0378	1 to 35	36 to 50	51 to 60	32.4%	52.1%	15.4%
11	-0.5407	1.3130	1 to 32	33 to 51	52 to 60	33.0%	54.0%	13.1%

The scale score metric was derived from the logits so that the minimum scale score for *Meets the Standards* was 85 and the minimum score for *Exceeds the Standards* was 135 for all grades. The calculations for the NeSA-Science scale score conversion are in Table 5.4.2.

Table 5.4.2: Conversion of Logits to Scale Scores

	Logit Cutpoints		Scale Score Ranges by Performance Level			Logit to Scale Score Conversion	
Grade	B/M	M/E	Below	Meets	Exceeds	Slope	Intercept
5	-0.4971	1.0580	1 to 84	85 to 134	135 to 200	32.15095	100.49331
8	-0.4543	1.0378				33.50958	99.73252
11	-0.5407	1.3130				26.97256	99.09502

5.5 Panelists' Survey Evaluation Results

The final step of the Bookmark Standard Setting process was asking the panelists to complete an evaluation on the Standard Setting meeting itself. This information was used to assess the panelists' impression of the validity of the process and their confidence in the result. A copy of the instrument is included in Appendix H and a summary of the results is included Appendix I.

References

Cizek, G. J., & Bunch, M. B. (2007). Standard setting: a guide to establishing and evaluating performance standards on tests. Thousand Oaks, CA: Sage.

Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996). Standard setting: a bookmark approach. In D. R. Green (Chair), *IRT-Based standard-setting procedures utilizing behavioral anchoring*. Symposium conducted at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.

Wright, B. & Stone, M. (1979). *Best test design*. Chicago: MESA Press.

Appendices

Appendix A: NeSA-Science Performance Level Descriptors

The Performance Level Descriptors provide meaning to the scale score metric and give a qualitative description of the numeric scores. The attached Performance Level Descriptors were used by the panelists during both the Bookmark Standard Setting and the Contrasting Groups study. The labels used for the levels were *Below the Standards, Meets the Standards,* and *Exceeds the Standards.*

Nebraska State Accountability-Science (NeSA-S) Performance Level Descriptors

Grade 5

Below the Standards	**Meets the Standards**	**Exceeds the Standards**
Overall student performance in science reflects *unsatisfactory* performance on the standards and *insufficient* understanding of the content at fifth grade. A student scoring at the Below the Standards level *inconsistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.	Overall student performance in science reflects *satisfactory* performance on the standards and *sufficient* understanding of the content at fifth grade. A student scoring at the Meets the Standards level *generally* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.	Overall student performance in science reflects *high academic* performance on the standards and *a thorough* understanding of the content at fifth grade. A student scoring at the Exceeds the Standards level *consistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.
A student at this level ***inconsistently***:	A student at this level ***generally***:	A student at this level ***consistently***:
• Identifies testable questions. • Identifies factors that may impact an investigation. • Identifies appropriate selection and use of scientific equipment. • Develops a reasonable explanation based on collected data. • Identifies physical properties of matter. • Identifies forces on motion. • Recognizes that energy transfers. • Identifies characteristics of living and nonliving things. • Identifies inherited characteristics and life cycle of living things. • Identifies components of an ecosystem. • Identifies adaptations made by plants and animals to survive. • Identifies changes of objects in the sky. • Identifies Earth's materials and structure. • Identifies energy sources on Earth. • Identifies changes in Earth's surface.	• Identifies testable questions. • Identifies factors that may impact an investigation. • Identifies appropriate selection and use of scientific equipment. • Develops a reasonable explanation based on collected data. • Describe the physical properties of matter and its changes. • Identifies the influence of forces on motion. • Identifies signs of energy transfer. • Compares the characteristics of living and nonliving things. • Identifies variations of inherited characteristics and life cycles. • Describes relationships within an ecosystem. • Describes changes in organisms over time. • Describes characteristics, patterns, and changes of objects in the sky. • Describes Earth's materials, structure, and processes. • Describes the effects of energy changes on Earth. • Describes changes in Earth's surface.	• Identifies testable questions. • Identifies factors that may impact an investigation. • Identifies appropriate selection and use of scientific equipment. • Develops a reasonable explanation based on collected data. • Compares physical properties of matter. • Compares the influence of forces on motion. • Compares energy transfers. • Compares how parts of organisms function to meet basic needs. • Compares variations of inherited characteristics and life cycles. • Compares relationships within an ecosystem. • Compares changes in organisms over time. • Compares characteristics, patterns, and changes of objects in the sky. • Compares Earth's materials, structure, and processes. • Compares the effects of energy changes on Earth. • Compares changes in Earth's surface.

Nebraska State Accountability-Science (NeSA-S) Performance Level Descriptors

Grade 8

Below the Standards	**Meets the Standards**	**Exceeds the Standards**
Overall student performance in science reflects *unsatisfactory* performance on the standards and *insufficient* understanding of the content at eighth grade. A student scoring at the Below the Standards level *inconsistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences. A student at this level ***inconsistently***: • Identifies testable questions that lead to predictions and scientific investigations. • Identifies and controls variables that impact a scientific investigation. • Identifies appropriate selection and use of scientific equipment. • Develops logical inferences based on collected data and accounts for non-relevant information. • Analyzes investigations for quality, accuracy, and relevancy. • Identifies the particulate nature of matter. • Identifies forces and motion. • Identifies energy systems. • Identifies structure of living organisms. • Identifies types of reproduction. • Identifies components of an ecosystem. • Identifies characteristic of organisms. • Identifies components of the solar system. • Identifies Earth's structure and processes. • Identifies energy in Earth systems. • Identifies changes in Earth over time.	Overall student performance in science reflects *satisfactory* performance on the standards and *sufficient* understanding of the content at eighth grade. A student scoring at the Meets the Standards level *generally* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences. A student at this level ***generally***: • Identifies testable questions that lead to predictions and scientific investigations. • Identifies and controls variables that impact a scientific investigation. • Identifies appropriate selection and use of scientific equipment. • Develops logical inferences based on collected data and accounts for non-relevant information. • Analyzes investigations for quality, accuracy, and relevancy. • Describes the particulate nature of matter including physical and chemical interactions. • Describes forces and motion. • Describes how energy systems and matter interact. • Describes the structure and function of living organisms. • Describes the relationship between reproduction and heredity. • Describes populations and ecosystems. • Identifies characteristics of organisms that help them survive. • Describes Earth and the solar system. • Describes Earth's structure, systems, and processes. • Describes energy in Earth's system. • Describes changes in Earth over time.	Overall student performance in science reflects *high academic* performance on the standards and *a thorough* understanding of the content at eighth grade. A student scoring at the Exceeds the Standards level *consistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences. A student at this level ***consistently***: • Identifies testable questions that lead to predictions and scientific investigations. • Identifies and controls variables that impact a scientific investigation. • Identifies appropriate selection and use of scientific equipment. • Develops logical inferences based on collected data and accounts for non-relevant information. • Analyzes investigations for quality, accuracy, and relevancy. • Describes the particulate nature of matter by comparing and contrasting physical and chemical interactions. • Predicts the impact of balanced and unbalanced forces acting on objects. • Evaluates interactions between energy and matter. • Evaluates the interactions between structure and function of living organisms. • Describes the relationship between reproduction and heredity. • Analyzes interactions between populations and ecosystems. • Evaluates survival of organisms based on characteristics. • Analyzes interactions between Earth and the solar system. • Analyzes interactions among Earth's structure, systems, and processes. • Analyzes energy's impact on Earth systems.

Nebraska State Accountability-Science (NeSA-S) Performance Level Descriptors

Grade 11

Below the Standards	**Meets the Standards**	**Exceeds the Standards**
Overall student performance in science reflects *unsatisfactory* performance on the standards and *insufficient* understanding of the content at eleventh grade. A student scoring at the Below the Standards level *inconsistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.	Overall student performance in science reflects *satisfactory* performance on the standards and *sufficient* understanding of the content at eleventh grade. A student scoring at the Meets the Standards level *generally* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.	Overall student performance in science reflects *high academic* performance on the standards and *a thorough* understanding of the content at eleventh grade. A student scoring at the Exceeds the Standards level *consistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.
A student at this level ***inconsistently***:	A student at this level ***generally***:	A student at this level ***consistently***:
• Identifies a testable hypothesis that guides a scientific investigation. • Identifies and manages variables and constraints. • Identifies appropriate selection and use of scientific equipment. • Analyzes and interprets data and evaluates models and explanations. • Evaluates scientific investigations. • Identifies matter in terms of its structure and composition. • Identifies field forces. • Identifies energy systems in matter. • Identifies organic molecules, sub-cellular structures, and cellular functions. • Identifies DNA and its role in heredity. • Identifies the flow of energy between organisms and their environment. • Identifies types of adaptations necessary for survival. • Identifies components of the universe. • Identifies the relationships between Earth's structure and processes. • Identifies relationships between sources of energy and Earth's systems. • Identifies the Law of Superposition.	• Identifies a testable hypothesis that guides a scientific investigation. • Identifies and manages variables and constraints. • Identifies appropriate selection and use of scientific equipment. • Analyzes and interprets data and evaluates models and explanations. • Evaluates scientific investigations. • Describes matter in terms of its structure, composition, and conservation. • Describes the nature of field forces and their interactions with matter. • Describes energy systems relating to the conservation and interaction of energy and matter. • Describes the chemical basis of the growth, development, and maintenance of cells. • Describes the molecular basis of reproduction and heredity. • Describes, on a molecular level, the cycling of matter and the flow of energy between organisms and their environment. • Describes the theory of biological evolution. • Describes the known universe. • Investigates the relationships among Earth's structure, systems, and processes. • Describes the relationships among the sources of energy and their effects on Earth's systems. • Explains the history and evolution of Earth.	• Identifies a testable hypothesis that guides a scientific investigation. • Identifies and manages variables and constraints. • Identifies appropriate selection and use of scientific equipment. • Analyzes and interprets data and evaluates models and explanations. • Evaluates scientific investigations. • Analyzes structure, composition, and conservation of matter. • Analyzes interactions between field forces and matter. • Analyzes interactions between energy systems and matter. • Analyzes the chemical basis of the growth, development, and maintenance of cells. • Analyzes the molecular basis of reproduction and heredity. • Analyzes the cycling of matter and the flow of energy between organisms and their environment. • Analyzes the theory of biological evolution and the diversity of life. • Analyzes the formation of the universe. • Analyzes the relationships among Earth's structure, systems, and processes. • Analyzes the relationships between sources of energy and their effects on Earth's systems. • Analyzes the history and evolution of Earth.

Appendix B: Meeting Agenda

Appendix B.1 Agenda

NeSA-S
Nebraska Bookmark Standard Setting Meeting
Agenda

Monday, June 25, 2012

Hotel Check-in for those traveling long distances

Tuesday, June 26, 2012 *(times are approximate depending on work completion)*

8:00 – 8:30 Light Breakfast and Check-in Lincoln Room

8:30 – 10:30 Training in Large Group Room Lincoln Room

10:35 – 12:00 Grade Group Breakouts, round 1

Science Grade	*Room*
5	*Board Room*
8	*Omaha Room*
11	*Capitol Room*

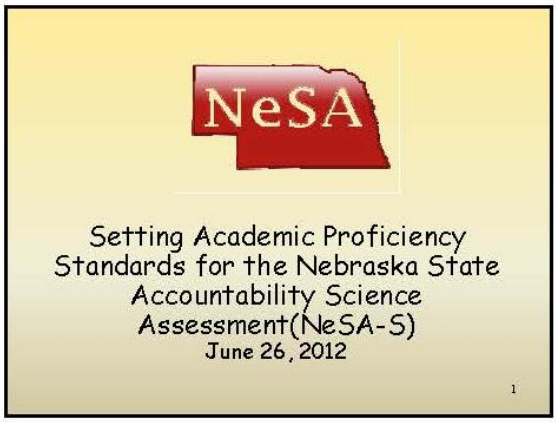
12:00 – 1:00 Lunch in Lincoln Room

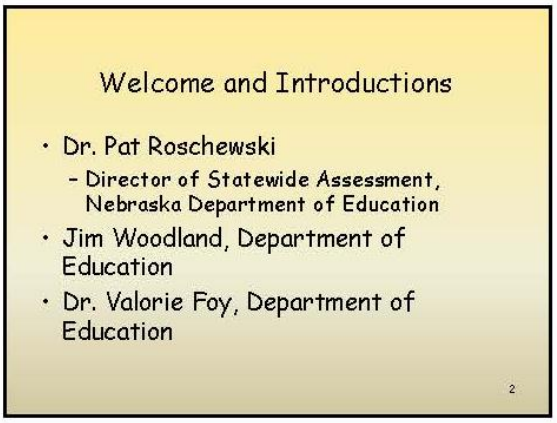
1:00 – Completion Complete rounds 2 and 3 of Bookmark process
(Afternoon break will be determined by completion of round 2 and will be held in the Lincoln Room.)

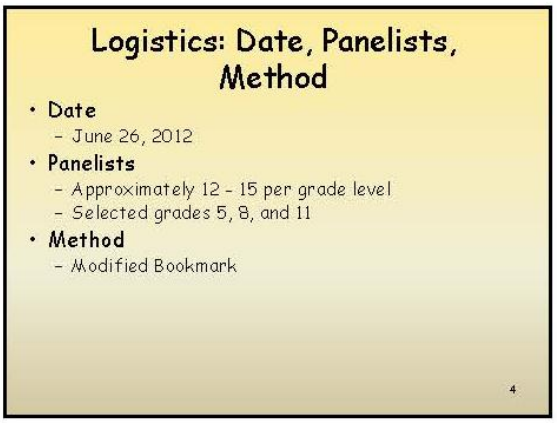
Appendix B.2: Timeline

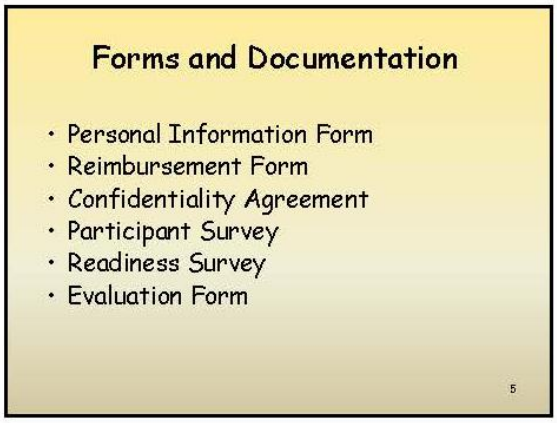
Science	Monday		
June 26, 2012	**Room 1**	**Room 2**	**Room 3**
8:00 AM **8:15 AM**	Breakfast/Check-in		
8:30 AM **8:45 AM** **9:00 AM** **9:15 AM** **9:30 AM** **9:45 AM** **10:00 AM** **10:15 AM**	Training Large Group		
10:30 AM	Move to grade level rooms		
10:45 AM **11:00 AM**	**Grade 5** Take test	**Grade 8** Take test	**Grade 11** Take test
11:15 AM **11:30 AM** **11:45 AM**	PLD review	PLD review	PLD review
12:00 PM **12:15 PM** **12:30 PM** **12:45 PM**		Lunch and Analysis	
1:00 PM **1:15 PM** **1:30 PM** **1:45 PM**		R1 OIB review and Bookmark placement	
2:00 PM		Break and Analysis	
2:15 PM **2:30 PM** **2:45 PM**		R1 Feedback and Discussion	
3:00 PM **3:15 PM** **3:30 PM**		R2 Bookmark Adjustments	
3:45 PM		Break and Analysis	
4:00 PM **4:15 PM** **4:30 PM** **4:45 PM**		R2 Feedback and Discussion Adding in NAEP data as available	
5:00 PM		R3	

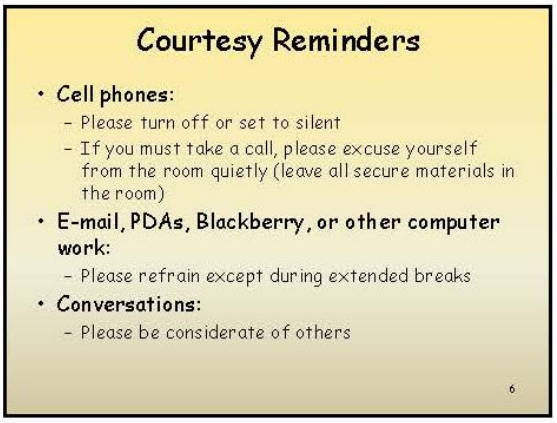
Appendix C: Setting Academic Proficiency Standards PowerPoint











Purpose of the Meeting

- To recommend NeSA-S cutscores that categorize students into one of three performance levels:
  - Exceeds the Standards
  - Meets the Standards
  - Below the Standards
- To articulate these expectations across grades 5, 8 and 11

7

Schedule

- Tuesday Break-out Groups:
  - Grade 5
  - Grade 8
  - Grade 11 (panelists from grade 11 and Higher Education)

8

Methodology

Modified Bookmark

- One in a broad category of methods commonly referred to as item mapping that focuses on items rather than examinees
- Places emphasis on what a student should know and be able to do

9

Step 1: How do students demonstrate their proficiency?

- Take the test
  - Provides panelists a feel for students' testing experience on the operational administration of the NeSA Science

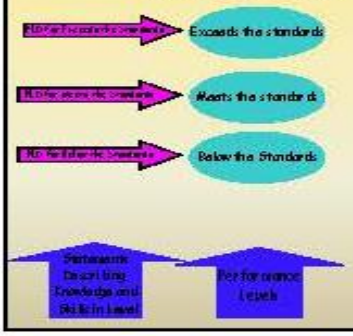
10

Step 2: What do the performance levels mean?

- Performance Level Descriptors (PLDs) are:
  - Statements that describe the knowledge and skills expected at each of the three achievement levels
  - Unique to each grade and subject
  - **Middle** of the level; not the borderline students

11

The Relationships Between PLDS and Performance Levels

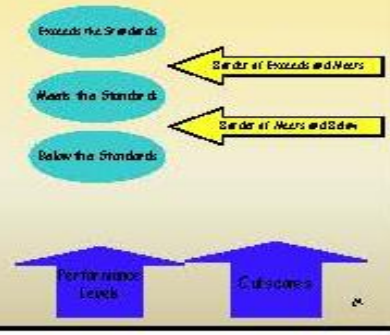


12

Step 3: What Defines Borderline Students?

- Visualize Nebraska students who are:
  - Just barely leaving one level, and
  - Just barely entering the next higher level

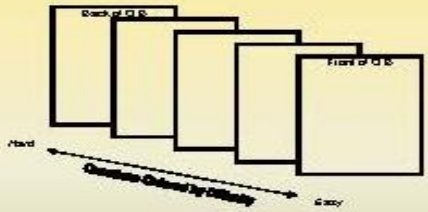
The Relationships Between Performance Levels and Cutscores



Step 4: What are we expected to do?

- Panelists are presented with operational test questions ordered from easiest to most difficult

Ordered Item Booklet (OIB)



*Based on Spring 2012 data

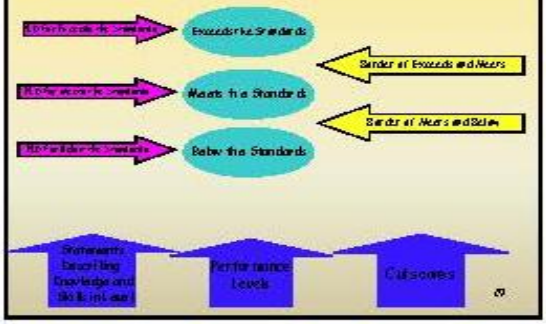
Step 5: Where do we draw the line?

- Panelists are asked to make judgments about which items students at the borderline between two performance levels are able to get correct and which ones they are not.
  - Criterion: "67 or more out of 100"
- Place the "bookmark" on the first item that does **not** meet the criterion.

What happens to our bookmarks?

- These collective judgments determine the recommended cutscores that separate:
  - ***Meets the Standards*** students from ***Below the Standards*** students
  - ***Exceeds the Standards*** students from ***Meets the Standards*** students

The Relationships Among PLDs, Performance Levels, and Cutscores



19

Overview of Process

- Three rounds of individual judgments
  - Group discussions
  - Opportunities to revise judgments
  - Data will be presented at the beginning of Rounds 2 and 3
    - For example, the percent of students that would fall into each of the three performance levels based on the group recommendation from the previous round

20

Articulation Across Grades: The Process

- Results from Rounds 1, 2, and 3 will be presented for all grades as they become available
- Contrasting Groups results will be presented for grades 5, 8 and 11
  - Participation from educators across the state
  - Used educator placements of students in their classrooms and NeSA results to determine cutscores
  - After Round 2
- NAEP Data results presented for Grade 8

21

Panelists' Roles and Responsibilities

- Satisfy yourself that you have contributed to a group recommendation that is based on your experience and professional judgment

22

Outcomes

- The recommendations from this meeting will be presented to the State Board for review, along with other relevant information
- Final cutscores will be established and approved by the State Board
- Final, Board-approved cutscores may not be the same as the group recommendations from this meeting

23

Test Security

- Check in
  - Distribution of training materials
- Security
  - ALL materials must remain in the room.
  - For unscheduled breaks, please notify the Room Facilitator.
  - No discussion related to any of the secure materials outside of the rooms, including breaks, and lunch
  - Check out
  - Turn in materials at the end of each session.
  - Materials will be returned to you at the beginning of each session

24

Return Times

- Lead Facilitator will indicate when you should return (e.g., after breaks and lunch)
- Times may be different for each panel

25

Training Materials

- Sample:
  - NeSA-S Performance Level Descriptors (grade 8 Science)
  - Five items (grade 8 Science)
  - Item Map
  - Item Separation Chart
  - Participant Rating Form
  - Results

26

Performance Level Descriptors

- Performance level descriptors (PLD's) describe the level of knowledge and skills required at each performance level.
  - Below the Standards
  - Meets the Standards
  - Exceeds the Standards

27

NeSA-S Performance Level Descriptors

Two levels

- Policy definitions provide descriptors for each level of proficiency
- Policy definitions with full descriptors communicate content expectations at each grade level assessed (number sense, geometric concepts, algebraic concepts, data analysis/probability concepts)

28

Performance Level Descriptors
Policy Statements

Exceeds the Standards	Overall, a student performance in science reflects *high-academic* performance on the standards and *a thorough* understanding of the content of the given grade. A student scoring at the Exceeds the Standards level *consistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.
Meets the Standards	Overall, a student performance in science reflects *satisfactory* performance on the standards and *sufficient* understanding of the content of the given grade. A student scoring at the Meets the Standards level *generally* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.
Below the Standards	Overall, a student performance in science reflects *unsatisfactory* performance on the standards and *insufficient* understanding of the content of the given grade. A student scoring at the Below the Standards level *inconsistently* draws on a broad range of scientific knowledge and skills in the areas of inquiry, physical, life, and Earth/space sciences.

29

Performance Level Descriptors
Full Descriptors

Exceeds the Standards	Describes the particulate nature of matter by comparing and contrasting physical and chemical processes. Analyzes interactions between Earth and the solar system.
Meets the Standards	Describes the particulate nature of matter including physical and chemical processes. Describes Earth and the solar system.
Below the Standards	Identifies the particulate nature of matter. Identifies components of the solar system.

30

Bookmark Training
(cont.)

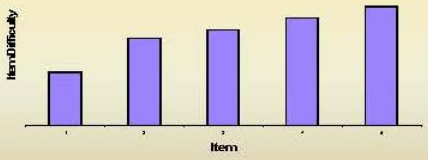
- The Item Map

OIB Page	Item Type	Key	Standard	Round 1	Round 2	Round 3
1	MC	B	3			
2	MC	C	2			
3	MC	C	3			
4	MC	A	5			
5	MC	C	4			

31

Bookmark Training
(cont.)

- Item Separation Chart



32

Placing the Bookmark: The Region of Uncertainty

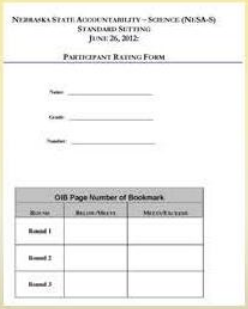
- Identify ***groups*** of items that are probable choices based on your judgment
- Do not focus on a single item
- Utilize the "item separation chart"

33

Bookmark Training
(cont.)

- Participant Rating Form
  - For the formal part of the meeting, two bookmarks will be placed in the OIB
    - 1st: Below the Standards/Meets the Standards
    - 2nd: Meets the Standards/Exceeds the standards

34


NEBRASKA STATE ACCOUNTABILITY – SCIENCE (NeSA-S)
STANDARD SETTING
JUNE 26, 2012

PARTICIPANT RATING FORM

OIB Page Number of Bookmark

35

Training
(cont.)

- Placing a Bookmark
  - Go through the OIB page by page and assess whether a borderline ***Meets the Standards*** student, according to the PLDs, has a sufficient probability of answering each item correctly
    - For multiple-choice (MC) items, ***sufficient*** is .67

36

Training (cont.)

- Does a borderline ***Meets the Standards*** student have at least a .67 probability of answering this item correctly?
- If yes, turn the page and make the same judgment about the next most difficult item

37

Training (cont.)

- Continue until you reach an item that the borderline ***Meets the standards*** student would **not** have a .67 chance of answering correctly
- ***Key Point***
  - Place your post-it bookmark on this page, which represents the first item that the borderline student would not answer correctly 67 percent of the time
    - *Reminder:* Region of Uncertainty!
  - Record this page number on your Rating Form

38

Practice Exercise

39

Practice Steps

- Review the five sample Science questions
- Review the NeSA-S PLDs for Meets the Standards and Below the Standards
- Visualize a student just barely out of the Below the Standards level and just barely into the Meets the Standards level
- Indicate on the Sample Rating Form the first item you judge your borderline student would get correct *less* than 67 percent of the time

40

Done?

41

Show of Hands!

42

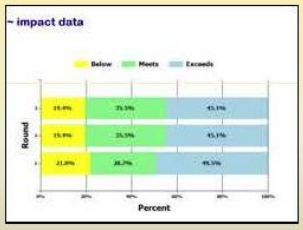
Sample Results

		Round 1		Round 2		Round 3	
ID	Table	Meets	Exceeds	Meets	Exceeds	Meets	Exceeds
1	1	22	38	26	41	27	42
2	1	20	37	27	41	28	41
3	1	22	39	30	40	31	40
4	1	20	38	31	40	30	40
5	1	23	37	30	39	29	40
6	1	22	36	30	39	29	41
7	1	25	37	29	41	28	42
8	1	28	39	31	41	30	42
9	1	27	35	30	43	32	43
10	1	26	38	31	40	30	41

43

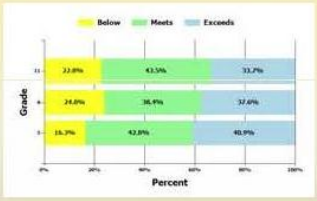
Sample Results (cont.)

Percent of Students in Performance Levels



44

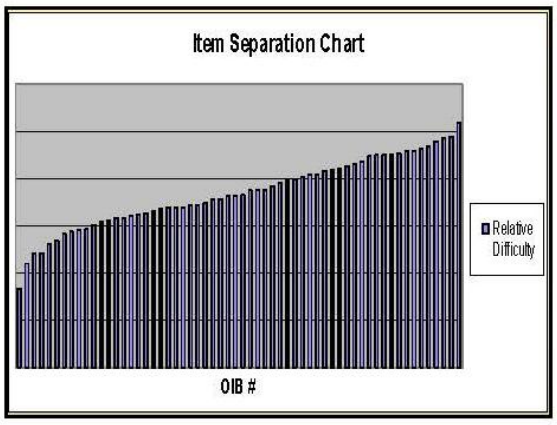
Science Results Across Grades Percent of Students in Performance Levels

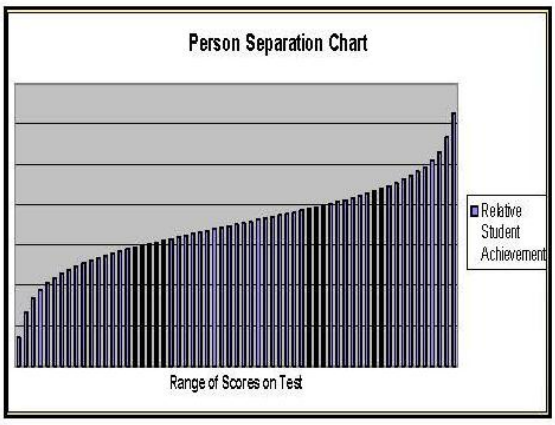


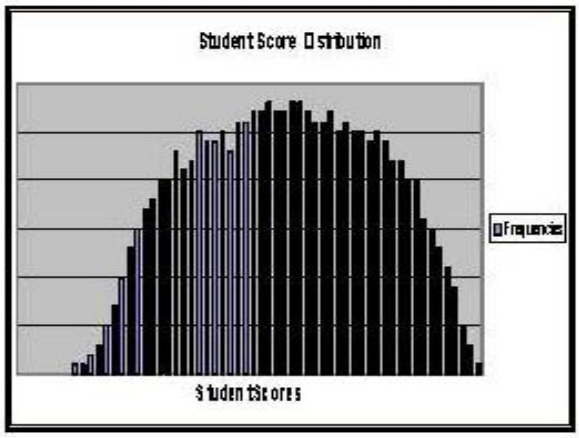
45

Psychometrics: An Introduction

46







What You May And May Not Discuss Outside Of This Meeting

- You may discuss:
  - The processes used
  - PLDs
- You may **not** discuss:
  - The results
  - The contents of the secure materials
    - Items

Questions?

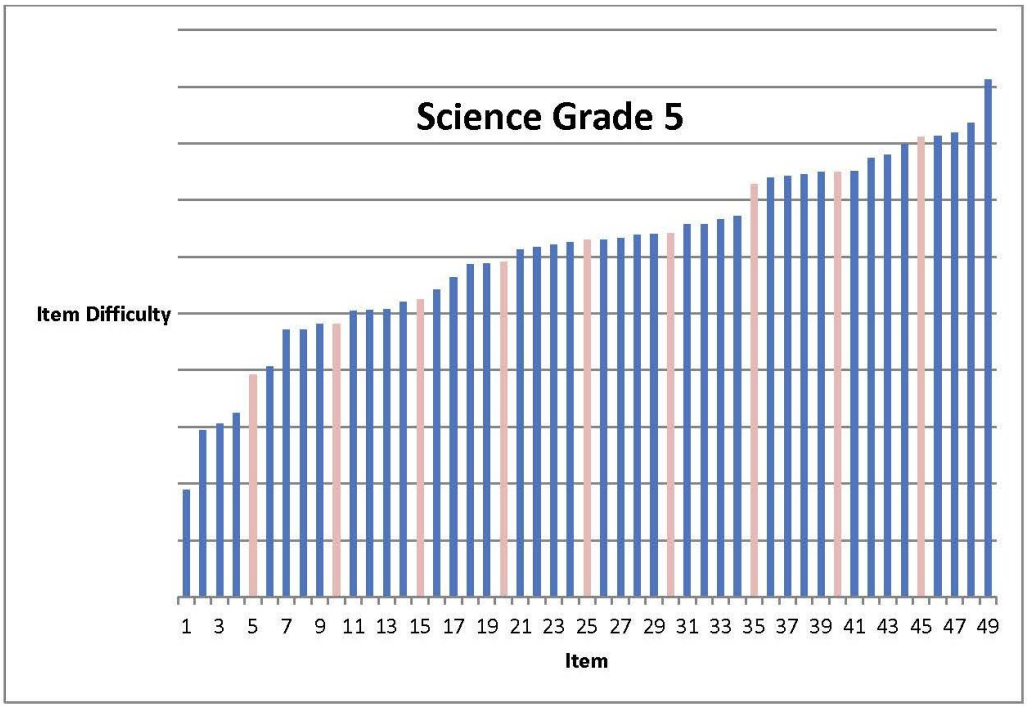
What's Next?

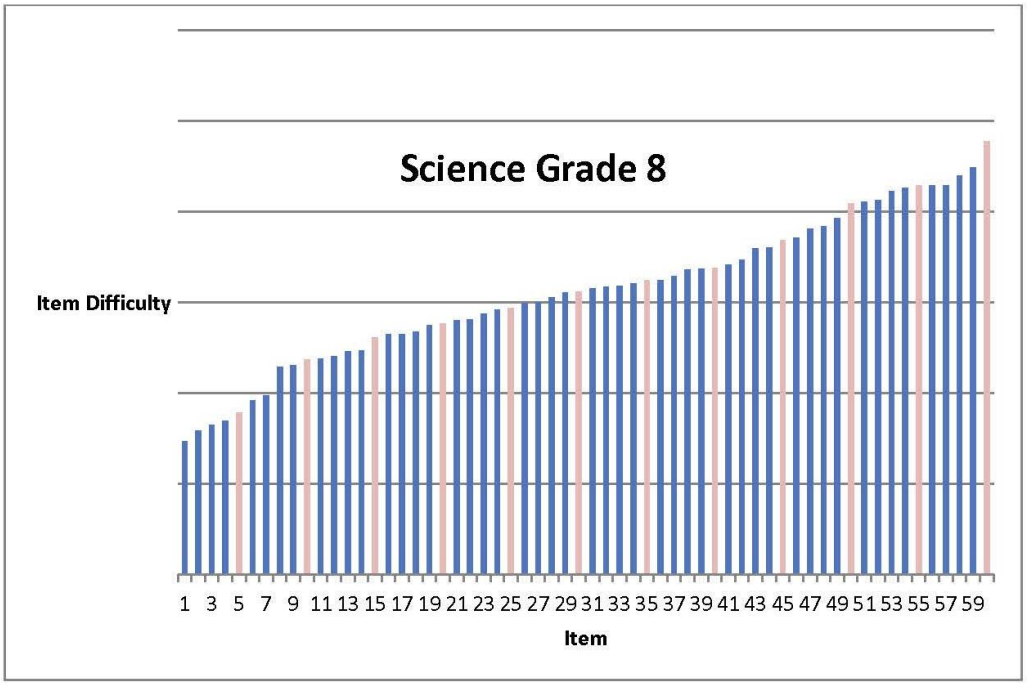
- Return evaluation and readiness survey sheets to staff at back of the room
- Proceed to breakout rooms as directed

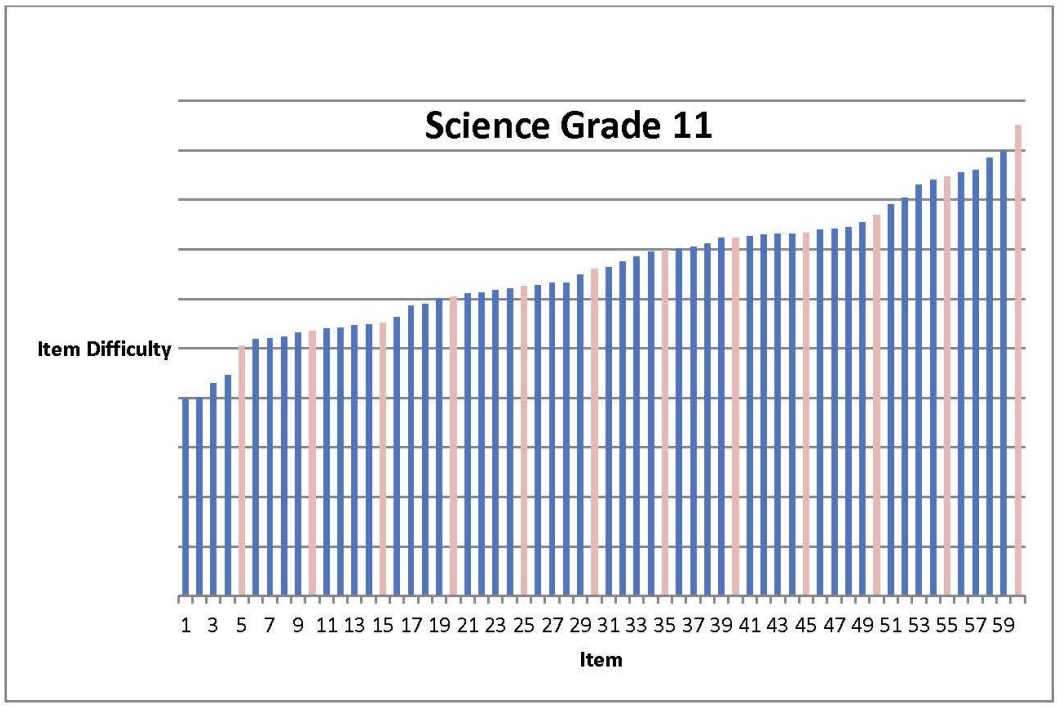
Appendix D: Impacts by Bookmark Round

Science	Below the Standards	Meets the Standards	Exceeds the Standards
Grade 5			
Round 1	24.0	28.3	47.7
Round 2	24.0	28.3	47.7
Round 3	44.6	29.0	26.4
Grade 8			
Round 1	20.0	28.0	52.0
Round 2	24.6	27.0	48.4
Round 3	32.4	41.6	26.0
Grade 11			
Round 1	30.6	32.7	36.7
Round 2	30.6	32.7	36.7
Round 3	30.6	32.7	36.7

Appendix E: Item Separation Maps






Science Grade 11
Item Difficulty
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59
Item

Appendix F: Contrasting Groups Summaries

Table F.1: Overall Contrasting Group Summary Data

Group	Grade 5		Grade 8		Grade 11	
	State	Teacher Rated	State	Teacher Rated	State	Teacher Rated
Student Count						
Overall	21683	2612	20803	3028	20815	2293
Gender						
Male	11046	1292	10735	1595	10656	1138
Female	10637	1320	10068	1433	10159	1155
Ethnicity						
African American	1448	96	1316	51	1306	70
American Indian	344	33	268	49	238	12
Hispanic	3658	404	3210	406	2827	245
Asian	471	59	392	53	437	47
White	15100	1961	15029	2405	15435	1860
Teacher Rating						
Below		634		814		639
Meets		1439		1616		1180
Exceeds		539		598		474
Performance Level						
Below	7281	671	6746	638	6861	432
Meets	11270	1486	10846	1753	11229	1425
Exceeds	3132	455	3211	637	2725	436

Table F.2: Agreement between Teacher Ratings and Final Performance Level Status

Grade 5		Teacher Rating		
		Below	Meets	Exceeds
Report Card	Below	**386**	265	20
	Meets	244	**988**	254
	Exceeds	4	186	**265**

Grade 8		Teacher Rating		
		Below	Meets	Exceeds
Report Card	Below	**443**	175	20
	Meets	370	**1153**	230
	Exceeds	1	288	**348**

Grade 11		Teacher Rating		
		Below	Meets	Exceeds
Report Card	Below	**288**	133	11
	Meets	325	**860**	240
	Exceeds	26	187	**223**

Appendix G: Contrasting Groups Analyses

Grade 5 Science Contrasting Groups

Raw Score	Logit Ability	Below	Meets	Exceeds	B+M	M+E	Log Odds Meets	Log Odds Exceeds
1	-6.630		1		1	1		
3	-4.410				0	0		
4	-3.972				0	0		
6	-3.385				0	0		
7	-3.163				0	0		
8	-2.969				0	0		
9	-2.795				0	0		
10	-2.637				0	0		
11	-2.492	1			1	0		
12	-2.356	5			5	0		
13	-2.228	1			1	0		
14	-2.106	3	1		4	1	-0.48	
15	-1.991	5			5	0		
16	-1.880	14			14	0		
17	-1.772	13	1		14	1	-1.11	
18	-1.668	13	2		15	2	-0.81	
19	-1.567	17	5	1	22	6	-0.45	-1.34
20	-1.468	24	4		28	4	-0.78	
21	-1.371	24	11		35	11	-0.34	
22	-1.275	25	11	1	36	12	-0.32	-1.56
23	-1.181	36	19	1	55	20	-0.26	-1.74
24	-1.087	36	17	1	53	18	-0.30	-1.72
25	-0.994	29	28	1	57	29	0.00	-1.76
26	-0.901	31	22	1	53	23	-0.13	-1.72
27	-0.808	26	30	2	56	32	0.09	-1.45
28	-0.715	32	33	2	65	35	0.04	-1.51
29	-0.622	23	38	6	61	44	0.28	-1.01
30	-0.527	26	43	4	69	47	0.26	-1.24
31	-0.431	38	60	4	98	64	0.23	-1.39
32	-0.334	35	54	8	89	62	0.25	-1.05
33	-0.235	37	48	3	85	51	0.14	-1.45
34	-0.133	23	74	10	97	84	0.56	-0.99
35	-0.028	21	70	7	91	77	0.56	-1.11
36	0.080	25	78	19	103	97	0.59	-0.73
37	0.192	14	77	16	91	93	0.82	-0.75
38	0.310	15	96	14	111	110	0.87	-0.90
39	0.434	9	99	17	108	116	1.11	-0.80

Grade 5 Science Contrasting Groups

Raw Score	Logit Ability	Below	Meets	Exceeds	B+M	M+E	Log Odds Meets	Log Odds Exceeds
40	0.565	12	82	31	94	113	0.97	-0.48
41	0.706	9	90	33	99	123	1.14	-0.48
42	0.858	3	81	50	84	131	1.64	-0.23
43	1.026	3	79	42	82	121	1.61	-0.29
44	1.214	1	55	56	56	111	2.05	0.00
45	1.429	2	44	50	46	94	1.67	0.04
46	1.685		45	53	45	98		0.07
47	2.004	1	27	45	28	72	1.86	0.21
48	2.441		11	34	11	45		0.49
49	3.166		4	12	4	16		0.48
50	4.391			15	0	15		

Grade 8 Science Contrasting Groups

Raw Score	Logit Ability	Below	Meets	Exceeds	B+M	M+E	Log Odds Meets	Log Odds Exceeds
0	-6.504				0	0		
1	-5.279				0	0		
2	-4.557				0	0		
3	-4.122				0	0		
5	-3.552				0	0		
6	-3.340	1			1	0		
8	-2.991				0	0		
9	-2.843				0	0		
10	-2.706				0	0		
11	-2.580				0	0		
12	-2.461				0	0		
13	-2.349	3			3	0		
14	-2.243	5			5	0		
15	-2.141	4			4	0		
16	-2.043	1			1	0		
17	-1.949	6			6	0		
18	-1.858	9	2	1	11	3	-0.48	-1.04
19	-1.769	10	2		12	2	-0.70	
20	-1.683	9	1		10	1	-0.95	
21	-1.598	11	1		12	1	-1.04	
22	-1.516	21	3		24	3	-0.85	

Grade 8 Science Contrasting Groups

Raw Score	Logit Ability	Below	Meets	Exceeds	B+M	M+E	Log Odds Meets	Log Odds Exceeds
23	-1.434	22	2		24	2	-1.04	
24	-1.354	15	4	1	19	5	-0.48	-1.28
25	-1.275	22	6	1	28	7	-0.50	-1.45
26	-1.197	26	9	1	35	10	-0.41	-1.54
27	-1.119	29	7	1	36	8	-0.56	-1.56
28	-1.042	32	3		35	3	-1.03	
29	-0.965	29	1	2	30	3	-0.99	-1.18
30	-0.888	26	12	1	38	13	-0.30	-1.58
31	-0.811	28	11	1	39	12	-0.37	-1.59
32	-0.734	33	18	6	51	24	-0.14	-0.93
33	-0.657	40	25	1	65	26	-0.19	-1.81
34	-0.579	27	38	2	65	40	0.17	-1.51
35	-0.501	32	30	2	62	32	0.00	-1.49
36	-0.422	37	28	2	65	30	-0.09	-1.51
37	-0.342	51	41	3	92	44	-0.06	-1.49
38	-0.261	36	47	8	83	55	0.18	-1.02
39	-0.178	42	61	4	103	65	0.19	-1.41
40	-0.094	33	65	13	98	78	0.37	-0.88
41	-0.008	37	80	10	117	90	0.39	-1.07
42	0.081	24	92	6	116	98	0.61	-1.29
43	0.171	23	88	5	111	93	0.61	-1.35
44	0.265	17	93	17	110	110	0.81	-0.81
45	0.363	13	102	11	115	113	0.94	-1.02
46	0.464	19	97	25	116	122	0.81	-0.67
47	0.570	13	97	20	110	117	0.95	-0.74
48	0.681	10	88	35	98	123	1.09	-0.45
49	0.799	8	80	28	88	108	1.13	-0.50
50	0.925	7	94	43	101	137	1.29	-0.37
51	1.061	1	66	46	67	112	2.05	-0.16
52	1.209		63	40	63	103		-0.20
53	1.372		45	50	45	95		0.05
54	1.555		37	57	37	94		0.19
55	1.767		29	36	29	65		0.09
56	2.018		22	48	22	70		0.34
57	2.335		18	38	18	56		0.32
58	2.768		6	17	6	23		0.45
59	3.489		2	11	2	13		0.74
60	4.713			5				

Grade 11 Science Contrasting Groups

Raw Score	Logit Ability	Below	Meets	Exceeds	B+M	M+E	Log Odds Meets	Log Odds Exceeds
1	-4.970				0	0		
2	-4.253				0	0		
3	-3.823				0	0		
4	-3.511				0	0		
5	-3.263				0	0		
6	-3.055				0	0		
7	-2.876				0	0		
8	-2.717				0	0		
9	-2.573				0	0		
10	-2.441				0	0		
11	-2.319	1	1		2	1	0.00	
12	-2.205	3			3	0		
13	-2.097	4			4	0		
14	-1.995	1	1		2	1	0.00	
15	-1.897	2	2		4	2	0.00	
16	-1.804	11	2		13	2	-0.74	
17	-1.714	11	1		12	1	-1.04	
18	-1.627	11	3		14	3	-0.56	
19	-1.542	7	4		11	4	-0.24	
20	-1.460	13	3		16	3	-0.64	
21	-1.379	12	4		16	4	-0.48	
22	-1.300	12	3		15	3	-0.60	
23	-1.223	17	3	1	20	4	-0.63	-1.30
24	-1.147	15	10		25	10	-0.18	
25	-1.071	19	15	1	34	16	-0.07	-1.53
26	-0.997	18	4		22	4	-0.65	
27	-0.923	18	10		28	10	-0.26	
28	-0.849	29	9	1	38	10	-0.46	-1.58
29	-0.776	16	8	2	24	10	-0.20	-1.08
30	-0.703	20	12	1	32	13	-0.19	-1.51
31	-0.630	22	22	3	44	25	0.06	-1.17
32	-0.557	26	16	3	42	19	-0.14	-1.15
33	-0.484	30	28		58	28	-0.03	
34	-0.410	20	23	2	43	25	0.10	-1.33
35	-0.335	19	34	4	53	38	0.30	-1.12
36	-0.260	24	35	4	59	39	0.21	-1.17

Grade 11 Science Contrasting Groups

Raw Score	Logit Ability	Below	Meets	Exceeds	B+M	M+E	Log Odds Meets	Log Odds Exceeds
37	-0.183	22	28	3	50	31	0.15	-1.22
38	-0.106	24	44	6	68	50	0.32	-1.05
39	-0.027	21	56	6	77	62	0.47	-1.11
40	0.054	15	39	11	54	50	0.52	-0.69
41	0.136	25	48	11	73	59	0.37	-0.82
42	0.221	11	52	11	63	63	0.76	-0.76
43	0.308	18	48	12	66	60	0.52	-0.74
44	0.399	23	50	14	73	64	0.44	-0.72
45	0.492	19	61	15	80	76	0.60	-0.73
46	0.590	12	66	14	78	80	0.82	-0.75
47	0.692	11	40	23	51	63	0.76	-0.35
48	0.800	9	46	15	55	61	0.83	-0.56
49	0.914	10	56	18	66	74	0.87	-0.56
50	1.037	6	55	37	61	92	1.19	-0.22
51	1.169	5	50	34	55	84	1.23	-0.21
52	1.313	6	45	41	51	86	1.16	-0.09
53	1.473	7	43	32	50	75	1.03	-0.19
54	1.652	4	24	48	28	72	1.26	0.23
55	1.860	4	24	31	28	55	1.14	0.04
56	2.109	2	22	24	24	46	1.36	0.00
57	2.421	3	16	15	19	31	1.01	-0.10
58	2.851		5	20	5	25		0.60
59	3.569		7	7	7	14		0.00
60	4.790			5	0	5		

Appendix H: Bookmark Panelist Evaluation Form

NEBRASKA STATE ACCOUNTABILITY-SCIENCE (NESA-S)
STANDARD SETTING MEETING
JUNE 26, 2012
EVALUATION FORM

THE PURPOSE OF THIS EVALUATION IS TO OBTAIN YOUR OPINIONS ABOUT THE STANDARD SETTING MEETING. YOUR OPINION WILL PROVIDE A BASIS FOR EVALUATING THE BOOKMARK PROCESS. PLEASE DO NOT PUT YOUR NAME ON THIS FORM. WE WANT YOUR OPINIONS TO REMAIN ANONYMOUS. AND ALSO NOTE, IN ORDER FOR YOUR ANSWERS TO BE INCLUDED PLEASE CLEARLY STATE YOUR RESPONSE.

1. Grade Level:

5 8 11

2. Circle the phrase that most accurately reflects your satisfaction with the training.

Clarity	Not at all	Somewhat	Adequate	Totally clear
Amount of Time	Way too little	Too Little	Appropriate	Too Much
Practice Exercises	Not Useful	Somewhat Useful	Useful	Very Useful

3. Check the column that most accurately reflects your level of agreement regarding the Performance Level Descriptors (PLDs).

	Strongly Disagree	Disagree	Agree	Strongly Agree
Adequate information was provided to participants regarding the PLDs.				
Adequate time was provided for participants to gain understanding of the PLDs.				
The PLDs capture what students should know and be able to do at each grade level.				
The PLDs communicate a reasonable profile of students' achievement at Below the Standards, Meets the Standards, and Exceeds the Standards.				
The PLDs were helpful in making decisions regarding cut-points.				

4. Check the column that most accurately reflects your opinion regarding the usefulness of the following materials.

Materials	Not Useful	Somewhat Useful	Useful	Very Useful
Test Booklet				
Ordered Item Booklet				
Item Separation Chart				
Item Map				
Statistical Impact Data				

5. Check the column that most accurately reflects your opinion regarding the amount of time allotted for your ratings.

Time Allotted	Too Little Time	About Right	Too Much Time
Round 1			
Round 2			
Round 3			

6. Check the column that most accurately reflects your satisfaction with the following roles.

Role	Not Satisfied	Somewhat Satisfied	Satisfied	Very Satisfied
DRC Psychometric Lead				
DRC Room Facilitator				
Other DRC Staff				

7. Check the column that most accurately reflects the level of confidence you had in determining the bookmark location for each assessment cut-point. Please only indicate confidence level for the grades in which you participated. Otherwise, leave it blank.

Grade	Cut-point Location	Not Confident	Partially Confident	Confident	Very Confident
5	Below/Meets				
	Meets/Exceeds				
8	Below/Meets				
	Meets/Exceeds				
11	Below/Meets				
	Meets/Exceeds				

8. How confident are you that the processes and methods used will produce valid results?

Not Confident Somewhat Confident Confident Very Confident

9. If you have further comments or suggestions for ways to improve the Bookmark meeting, please do so in the space below. All comments will remain anonymous.

THANK YOU FOR PARTICIPATING IN THE STANDARD SETTING MEETING.

Appendix I: Bookmark Panelist Evaluation Summary

	Grade	5	8	11
	Count	13	11	9
Training	Clarity	3.9	3	3.2
	Time allotted	3.3	3	3.0
	Excercises	3.3	3	3.1
PLD's	Adeq info	3.5	2.9	3.5
	Adeq time	3.5	3	3.1
	Capture	3.5	3.1	3.2
	Communicated	3.4	3.1	3.2
	Helpful	3.6	3.1	3.4
Materials	Test bklt	3.8	3.5	3.2
	OIB	3.8	3.7	3.7
	Item seperation	3.5	2.7	3.0
	Item map	3.7	2.7	3.2
	Stat data	3.7	3.4	3.2
Roles	PS Lead	3.9	2.9	3.3
	Rm Fac	3.8	3.1	3.6
	Other	3.8	2.7	3.2
Confidence	Below/Meets	3.1	2.6	3.0
	Meets/Exceeds	3.0	3	2.8
Process	Confident	3.0	2.7	3.0

Amount of time*	Rnd 1	2.3	2.1	1.9
	Rnd 2	2.1	2	2.2
	Rnd 3	2.0	2	2.1

*Three point scale: Too Little, About Right, Too Much

For the quantitative analyses, the categories were coded 1 to 4, except questions about "Amount of Time" were 1 to 3. Please refer to Appendix H for the precise category labels.

Appendix J: Contrasting Groups Teacher Opinion Summary

Contrasting Groups Teacher Survey 2012

All Participants	Strongly Disagree(1)	Disagree(2)	Agree(3)	Strongly Agree(4)	Average	Total # Responded
1. The PLDs describe what students should know and be able to do in each grade.	3%	3%	77%	17%	3.1	151
2. The PLDs describe what students are expected to achieve at *Exceeds the Standards* level.	2%	4%	76%	18%	3.1	151
3. The PLDs describe what students are expected to achieve at *Meets the Standards* level.	1%	5%	76%	17%	3.1	150
4. The PLDs describe what students are expected to achieve at *Below the Standards* level.	3%	5%	75%	17%	3.1	151
5. The PLDs helped me place my students in the appropriate achievement level.	3%	10%	70%	17%	3.0	149
6. Adequate information was provided during training to understand the PLDs.	2%	6%	67%	25%	3.1	147
7. The process for completing on-line student ratings was explained clearly.	2%	4%	61%	33%	3.2	150
8. Completing the student ratings was easy to do.	3%	2%	60%	35%	3.3	150
9. Completing the student ratings took reasonable time.	2%	5%	58%	36%	3.3	149

10. Did you attend any of the on-line training sessions?	Yes	111	No	39
If not, did you access the recording of the on-line training session available from NDE and DRC?	Yes	23	No	20

Participants who answered "YES" #10	Strongly Disagree(1)	Disagree(2)	Agree(3)	Strongly Agree(4)	Average	Total # Responded
1. The PLDs describe what students should know and be able to do in each grade.	3%	5%	75%	18%	3.1	111
2. The PLDs describe what students are expected to achieve at *Exceeds the Standards* level.	2%	5%	75%	18%	3.1	111
3. The PLDs describe what students are expected to achieve at *Meets the Standards* level.	1%	5%	75%	18%	3.1	110
4. The PLDs describe what students are expected to achieve at *Below the Standards* level.	2%	6%	74%	18%	3.1	111
5. The PLDs helped me place my students in the appropriate achievement level.	3%	9%	73%	15%	3.0	110
6. Adequate information was provided during training to understand the PLDs.	2%	6%	65%	27%	3.2	110
7. The process for completing on-line student ratings was explained clearly.	2%	5%	59%	35%	3.3	111
8. Completing the student ratings was easy to do.	3%	2%	60%	35%	3.3	110
9. Completing the student ratings took reasonable time.	2%	6%	59%	34%	3.2	109

Participants who answered "No" to Q#10 part 1 and "Yes" to Part 2	Strongly Disagree(1)	Disagree(2)	Agree(3)	Strongly Agree(4)	Average	Total # Responded
1. The PLDs describe what students should know and be able to do in each grade.	0%	0%	86%	14%	3.1	22
2. The PLDs describe what students are expected to achieve at *Exceeds the Standards* level.	0%	0%	77%	23%	3.2	22
3. The PLDs describe what students are expected to achieve at *Meets the Standards* level.	0%	5%	73%	23%	3.2	22
4. The PLDs describe what students are expected to achieve at *Below the Standards* level.	5%	0%	73%	23%	3.1	22
5. The PLDs helped me place my students in the appropriate achievement level.	0%	0%	71%	29%	3.3	21
6. Adequate information was provided during training to understand the PLDs.	0%	5%	65%	30%	3.3	20
7. The process for completing on-line student ratings was explained clearly.	0%	0%	73%	27%	3.3	22
8. Completing the student ratings was easy to do.	0%	0%	59%	41%	3.4	22
9. Completing the student ratings took reasonable time.	0%	0%	55%	45%	3.5	22

Participants who answered "No" to Q#10 part 1 and "No" to Part 2	Strongly Disagree(1)	Disagree(2)	Agree(3)	Strongly Agree(4)	Average	Total # Responded
1. The PLDs describe what students should know and be able to do in each grade.	6%	0%	82%	12%	3.0	17
2. The PLDs describe what students are expected to achieve at *Exceeds the Standards* level.	6%	0%	82%	12%	3.0	17
3. The PLDs describe what students are expected to achieve at *Meets the Standards* level.	6%	6%	82%	6%	2.9	17
4. The PLDs describe what students are expected to achieve at *Below the Standards* level.	6%	6%	82%	6%	2.9	17
5. The PLDs helped me place my students in the appropriate achievement level.	6%	29%	53%	12%	2.7	17
6. Adequate information was provided during training to understand the PLDs.	6%	6%	82%	6%	2.9	17
7. The process for completing on-line student ratings was explained clearly.	6%	6%	65%	24%	3.1	17
8. Completing the student ratings was easy to do.	6%	6%	65%	24%	3.1	17
9. Completing the student ratings took reasonable time.	6%	6%	59%	29%	3.1	17


DATA RECOGNITION
DRC
CORPORATION