THE
COLORADO
EDUCATION
INITIATIVE

STUDENT PERCEPTION SURVEY TECHNICAL REPORT

August 29, 2013

Colorado's Student Perception Survey (SPS) was developed by The Colorado Education Initiative as part of a grant from the Bill & Melinda Gates Foundation; independent analyses of Colorado's SPS have established that it is a fair, valid, and reliable measure. The items on Colorado's SPS were adapted in part from items made available for non-commercial use through the Measures of Effective Teaching (MET) Project, also funded by the Gates Foundation.

CEI is grateful to Dr. Ron Ferguson and Cambridge Education for their pioneering work in the area of student perception surveys and the development of the Tripod Project® surveys and its 7Cs™ framework for teaching, which were key components of the MET Project. For more information on the Tripod Project, please visit http://tripodproject.org/.

Table of Contents

List of Tables and Figures

Tables

Figures

Section 1: Conceptualization and Early Development

Consistent with nationwide trends, the passage of a new educator evaluation law (Colorado Senate Bill 10-191, the *Ensuring Quality Instruction through Educator Effectiveness Act*) radically altered the existing system of teacher evaluation in Colorado. S.B. 10-191 requires annual evaluation of all teachers and principals based on their professional practices and the academic growth of their students. The law also changed the way teachers earn non-probationary status: Instead of completing three years of satisfactory service based on principal observation and recommendation, they must now must earn an effective or highly effective rating for three consecutive years, based on both professional practices and student growth.

S.B. 10-191 also strongly encourages the use of student perception data, although the particular method of engagement is left unspecified:

> *"Districts are strongly encouraged to gather student perceptions of their learning experience in order to provide teachers with feedback on their performance. Where appropriate, districts are encouraged to use student perception data as part of the multiple measures used to evaluate teacher professional practice."* (Source: 1 CCR 301-87, 7.04)

To address this need – as part of a grant funded by the Bill & Melinda Gates Foundation to integrate reform efforts focused on standards, assessment, and evaluation – The Colorado Education Initiative has engaged key stakeholders from across the state, including teachers, district and state staff, and students, to develop and pilot a student perception survey. Our primary goal in this work is to ensure districts have reliable, fair, and valid ways of assessing teacher effectiveness, and to ensure teachers receive the feedback and support that they need to improve.

At the core of this work is our belief that teachers care about improving their instructional practice, and they care about their students. Historically, however, direct feedback from students about their experiences with teachers has seldom been collected for educators to use to improve their practice. However, research has shown that student perception data is in many cases more valuable than classroom observations and student growth on standardized tests. In fact, the combination of student perception, classroom observation and student growth provides the most clear and comprehensive picture of teacher performance. Student Perception Surveys provide a unique form of actionable feedback that districts, schools and teachers can use to inform practice.

CEI has created Colorado's Student Perception Survey (SPS) to be an effective and reliable tool that offers fair, relevant, and actionable student feedback to teachers. We engaged in a rigorous

development and pilot process to create a survey that is specifically designed with input from Colorado teachers and students.

Throughout the pilot process, we worked closely with 16 partner districts to identify best practices regarding survey administration, and we have incorporated those best practices in every step of this planning guide.

Our 16 partner districts represented a range of diversity, including:

- Districts with diverse geographies, including large urban districts, rural districts, and mountain districts
- Districts with varying levels of technological access: For example, one of our partners provides a tablet for every student in the district, while several others only have one computer lab for all schools in the district to share.
- Districts with highly variable support staff structures, including some that have designated tech, data, and communications staff, and others where the superintendent might also be a principal, data coordinator, or director of human resources.

The pilot period – and the hard work done by our 1,400 teachers who participated – was strategically designed to gather lessons learned, teacher feedback, and best practices to provide educators with the information they need to be able to trust and rely on their results.

1.1. Format of Student Perception Survey

Colorado's Student Perception Survey is a 34-question instrument that asks students about their classroom experiences. For each item, students are asked to indicate on a frequency scale ("Never," "Some of the Time," "Most of the Time," or "Always") how often they experience certain teaching behaviors (e.g., "*My teacher explains difficult things clearly*").

The survey items are organized by four elements, developed over the course of the pilot through analyses of the underlying relationships between items:

1. **Student Learning**: How teachers use content and pedagogical knowledge to help students learn, understand, and improve.
2. **Student-Centered Environment**: How teachers create an environment that responds to individual students' backgrounds, strengths, and interests.
3. **Classroom Community**: How teachers cultivate a classroom learning community where student differences are valued.
4. **Classroom Management**: How teachers foster a respectful and predictable learning environment.

Appendix A includes the full list of final survey items, mapped to the associated survey element.

Relationship to Teacher Quality Standards

Colorado's Student Perception Survey has been intentionally linked to the Colorado Teacher Quality Standards, and therefore can be used with any rubric for evaluating teacher effectiveness that aligns with the Colorado Teacher Quality Standards.

The first element, **Student Learning**, primarily relates to Standards I (Know Content) and III (Facilitate Learning) and the associated professional practices in the Rubric for Evaluating Colorado's Teachers. Student Learning comprises of fifteen survey items, which is more than any other element. These items relate to both content and pedagogy. Through the survey development process, we learned that students experience instruction in a way that blends these two concepts. As such, districts, schools, and teachers will want to explore both of these aspects of teacher practice as they use survey results from this element to identify strengths and inform opportunities for professional growth.

The remaining three elements, **Student-Centered Environment**, **Classroom Management**, and **Classroom Community**, primarily relate to Standard II (Establish Environment), as these are all crucial aspects of establishing a safe, inclusive, and respectful learning environment for a diverse population of students. These survey elements capture multiple aspects of classroom climate, from mutual respect and community-building to establishing routines that result in maximizing on-task behavior.

It is important to note that the survey elements are not mutually exclusive. Teaching is a complex profession and statements from the survey apply to multiple professional practices and across the Colorado Teacher Quality Standards.

Grade-Level Distinctions

Colorado's Student Perception Survey has been designed for two grade spans: grades 3-5 and grades 6-12. Both instruments have are composed of the elements described above, and both map to the Colorado Teacher Quality Standards.

The two instruments have many questions in common, although each was designed with the specific developmental and linguistic needs of students at that age in mind. Both were also tested extensively through cognitive interviews and pilot administrations to ensure that all items function well for the students for whom they are designed.

Section 2: Pilot Design & Development

2.1. Initial Development and Piloting

Overview

CEI undertook a rigorous instrument development process, largely informed by Wilson's (2003, 2005) Construct Modeling approach to measurement. The survey development process and associated timeline is outlined in Table 1 below.

Table 1. *Survey development process*

Process/Survey Development Task	Timeline
Construct definition	April – May 2012
Item development, and mapping of items to the construct	May 2012
Item/construct review, including feedback from districts, the Colorado Department of Education , and a teacher work group	May 2012
Psychometric field test, to establish baseline psychometric properties and refine instrument as needed before Use Pilot	June 2012
Analysis of data from psychometric field test & instrument revisions	July – Aug 2012
Think-alouds/cognitive interviews[1]	August 2012
Fall pilot – Administer the survey to a collection of 16 districts statewide	November 2012
Analysis of results from use pilot – Analyze survey data to inform revisions and changes to the instrument	Dec – April 2012
Teacher feedback survey – Administered to teachers in 12 of the participating districts	January 2013
Teacher focus groups (Round 1) – Teachers were convened to discuss the instruments and recommended changes and preferences for reporting formats	March 2013
Spring validation pilot administration	April – May 2013
Prepare guidance for districts and teachers regarding analysis/use of survey data	May 2013
Release fall & spring teacher, school, and district reports	May 2013
Analysis of results from spring pilot, and examine the appropriateness of using SPS results to inform teacher evaluation	May – August, 2013
Teacher focus groups (Round 2 & 3) – Teachers were convened to discuss the pilot process, lessons learned, and to inform the development of future communication materials	June & August 2013
Prep & release full SPS toolkit, a free and publically-available toolkit to support districts in implementing the student survey independently	June & August 2013

[1] Although it is a bit unorthodox to conduct think-alouds after the initial field test, we structured the development process this way to accommodate three rounds of pilot testing (the psychometric field testing, the fall pilot, and the final spring pilot).

Construct Definition

As described above, Colorado's Student Perception Survey was developed to help teachers reflect on effective classroom practices and improve their practice. In short, our underlying belief is that students know an effective classroom when they experience one, and we sought to develop an instrument that would tap students' experiences in classrooms and perceptions about their teachers' practices.

Relevant Research. The largest and most recent inquiry into the use of student feedback in assessing teacher practice is the Measures of Effective Teaching (MET) Project. The (MET) project was a research partnership funded by the Bill and Melinda Gates Foundation that engaged 3,000 teacher volunteers and dozens of independent research teams. The project's goal was to build and test measures of effective teaching to find out how evaluation methods could best be used to tell teachers more about the skills that make them most effective and to help districts identify and develop great teaching.

The MET study tested several measures that can be used to evaluate the array of a teacher's contribution to student learning. One of these measures was a student perception survey, in which students responded to questions about teacher actions. The MET study had two significant findings around student perception surveys:

- When student surveys are combined with observation and student growth data, these three measures inform and predict future effectiveness better than any of them alone.
- Scores from student perception survey results are correlated with student achievement gains.

On a broader scale, the use of student feedback more generally has been shown to impact both teachers and students positively. On the one hand, teachers can learn about patterns in their teaching that they may not have been aware of and how those approaches impact student learning. On the other hand, students are given a forum in which they can be heard, and this emphasis on student voice promotes both reflection and responsibility on the part of the students.

Item Development and Review. After conducting a review of the limited literature related to student perception surveys, we contacted districts across the country who were piloting, developing, or implementing some form of student surveys to gain some sense of the scope and scale of student survey work across the country. We decided ultimately to develop a unique survey – one that pulled from the extraordinary work being done across the country, but that was specifically aligned to the Colorado Teacher Quality Standards and that could be tested locally with Colorado teachers. Initial item development was largely informed by the items available for non-commercial use from the MET study. These items were adapted and new items developed in collaboration with a variety of CEI, district, and state representatives, all with prior classroom

experience (at both the elementary and secondary level). In total, these individuals had expertise in:

- Instruction and pedagogy
- Curriculum and assessment,
- Teacher education,
- Teacher professional development,
- Data analysis and statistics.

Two state representatives without direct classroom experience were also included on the survey development team: One individual was selected because of her strong psychometric background – including prior experience at a large assessment company – and with the other had previous experience developing and piloting a student perception surveys for a large, urban district in Colorado.

Item developers were instructed to develop items that would best capture students' impressions of and perceptions about their classroom teachers. The item development group based their work on the results from the MET study (i.e., by including items that showed the greatest item discrimination in the publically available research reports from MET) and on the Colorado Teacher Quality Standards (e.g., by including additional items specifically designed to measure the Colorado standards).

A work group of current Colorado teachers formally reviewed preliminary survey instruments; and CEI researchers made revisions based on these teachers' recommendations prior to piloting the perception surveys in the psychometric field test.

2.2. Psychometric Field Test Findings & Instrument Revision

In June 2012, the psychometric field test of the Student Perception Surveys was administered to 311 Colorado students, including students in three grade spans: grades 6-12 (N=109), grades 3-5 (N=152), and grades ECE-2 (N=50).[2] The sample of students was taken from the classrooms of 18 teachers who volunteered to participate. Following this psychometric field test, rigorous item analyses were conducted to support item revision and instrument design. A brief summary of the preliminary results from these analyses follows.

Preliminary analyses from the psychometric field test suggested that the survey instruments for students in grades 3-5 and 6-12 demonstrated sound psychometric properties. P-values are a measure of item difficulty ranging from 0 to 1, and can be interpreted as total percent of points earned on a given item; for example, on the student perception survey, p-values are calculated by

[2] The survey for grades ECE-2 was ultimately tabled for a variety of reasons, including generally less positive feedback from the field and statistical difficulties with the instrument.

dividing the average score across students by the total points possible. The calculated p-values on the psychometric field test varied by instrument, ranging from 0.17 to 0.77 for grades 6-12, and 0.19 to 0.86 for grades 3-5 – both within typical ranges reported in large-scale standardized assessments. Item discriminations (corrected item-total or "point-biserial" correlations) ranged from 0.01 to 0.60; several items displayed properties not typically reported for large-scale assessments, and were revised accordingly. Finally, Cronbach's α for the grade 3-5 and the grade 6-12 instruments were exceptionally high – 0.84 and 0.92 respectively.

Nevertheless, several revisions were made to the instrument to eliminate items that were not performing well and improve the overall reliability of the instrument. For example, on the grade 3-5 instrument, "*My teacher doesn't let me give up when the work is hard*" was revised to "*When the work is too hard, my teacher helps me keep trying*" because it was only weakly correlated with other items on the original instrument. Similarly, the item "*My teacher wants us to understand what we learn, not just memorize facts*" was revised to "*In this class, it is more important to understand the lesson than to memorize the answers*."

2.3. Think-Aloud Interviews

One important step in the development of assessments and survey instruments is the investigation of response processes (i.e., how respondents think about and interpret items, and the associated cognitive processes when responding to items). One way to explore these processes is through *think-aloud* interviews.

In think-alouds, respondents are asked to "talk aloud about what they are thinking while they are actually responding to the item" (Wilson, 2005). In August 2012, we conducted think-aloud interviews with students about the student perception survey, helping us ensure that students interpreted our items as we intended.

Think-Aloud Participants

Sample size is typically quite small when conducting think-alouds and cognitive interviews. Nielson (1994) suggested that as few as five participants can yield sufficient information. Despite the small sample size, it is important that sampling of participants is purposeful: "Subjects are chosen as representatives of particular subsets of students deemed important to the project" (Johnstone, Bottsford-Miller, & Thompson, 2006). For the purposes of the student perception survey, it is critical that we examine how the survey functions for students traditionally underserved in public education – low income students, students of color, English Language Learners, and students with disabilities. Some of these analyses can be conducted using student-level survey responses from the psychometric field test, but the think-alouds provide rich qualitative data unavailable via quantitative analyses.

For each grade span (3-5, 6-12), we selected students to ensure we had representation from each of the following groups:

1. Students eligible for free or reduced-price lunch
2. Students of color
3. English Language Learners (who would not take the Spanish version of the instrument)
4. Students with disabilities
5. Students without disabilities who are proficient in English (comparison group)

We also made efforts to ensure geographic diversity with regard to region (e.g., across school districts,) and urbanicity. Table 2 describes the students who participated in the think-alouds in the fall of 2012.

Table 2. *Think-aloud participants*

Description/Characteristics	Percent
Grade Level	
Elementary	46%
Secondary	54%
Gender	
Female	38%
Male	62%
Race	
White	46%
Hispanic	31%
Black/Multi-racial	8%
Asian	15%
English Language Learners	23%
Students with Disabilities	23%

Student think-aloud interviews were scheduled across the Front Range (the Denver metro area extending from South Denver to Boulder) with 13 students – six in elementary school and seven in middle or high school.

Think-Aloud Protocol

Elementary think-aloud interviews took approximately 30 minutes, and secondary interviews took between 30-45 minutes each. All interviews were audiotaped and later transcribed, and were administered in a quiet, private location (e.g., an office or classroom).

Students were initially instructed that “thinking-aloud” may be new or unfamiliar, and that the information would not be used for any other purpose than improving the survey instrument; they

were also reassured that their confidentiality would be maintained and that no one – including the teacher about whom they were responding – would ever see their responses on the instrument. Before the students were given the Student Perception Survey, they were given a warm-up question to introduce the think aloud process: "Try to visualize the place where you live, and think about how many windows there are in that place. As you count the windows, tell me what you are seeing and thinking about" (Willis, 1994). Following this warm-up, students read aloud each survey item, explaining the cognitive process they used to reach an answer. They were also asked to record their answer on the instrument. When necessary, simple follow-up questions were asked (e.g., "Can you say more about that?" and "Can you explain why you chose the answer you did?").

Results from the Think-Aloud Interviews

Overall, the results from the cognitive interviews were very positive and encouraging. Overwhelmingly, students of all backgrounds were able to engage with the questions and provide thoughtful feedback on their teachers' instructional practices. A handful of items were revised to increase readability and understandability, particularly to eliminate colloquialisms (e.g., "speak up") and increase the accessibility of the vocabulary for English Language Learners (e.g., eliminating the word "input").

Despite these minor changes, students responded thoughtfully, supporting the notion that students can respond to questions about teaching quality and instruction. For example, when asked to respond to the item, "In my class, we learn things that matter to me," a female 7th grade ELL student responded "Always" and explained: "[My teacher] made the people who speak Spanish feel more important because we participated… we could teach about our culture. [It] taught us to trust in ourselves." An 8th grade Latino student shared the following: "I say most of the time [about the item "*My teacher knows when we understand the lesson and when we do not"*] … because one of my friends didn't understand and when she asked if we all understood, he didn't say anything [and she didn't know and kept on going]."

2.4. Fall Pilot Overview

Revisions were made to the instrument based on results from the psychometric field test and the think-aloud study. The version of the survey piloted in the fall can be found in Appendix B and in Tables 4 and 5, and is referred to throughout this section as the "fall pilot instrument."

Pilot Design

Several priorities informed our pilot design. First, for each teacher, it was critical to survey a *sample* of students; this not only helps to ensure that the results are fair and valid, but also prevents students in grades 6-12 from being over-surveyed (i.e., by responding about each of

their teachers each year). Second, it was critical that student confidentiality be protected. And, finally, it was critical that the survey be accessible to as many students as possible. To ensure this, we invested in a robust translation process, including back translation. We also required districts to provide accommodations as necessary, and provided them with a list of approved accommodations, largely informed by those allowed by the state during the Transitional Colorado Assessment Program (TCAP), the state's standardized assessment system.

Participants

Just over 40,000 student surveys were collected from the fall pilot. Those surveys were collected from 16 Colorado districts, representing a mix of rural, suburban, and mountain regions. In total, students responded about over 1,400 teachers in 86 schools. Table 3 below outlines the 16 district participants (and one additional participant that has joined the spring administration), including demographic and enrollment data.

Table 3. *Pilot district demographics & enrollment data, 2010*

District	Student N	Teacher N	% FRL	% Minority	% ELL	% H.S. Grad.
Centennial	248	21	82%	92%	13%	53%
Eagle County	6,181	477	44%	54%	37%	81%
Thompson	15,310	962	32%	23%	3%	80%
Archuleta	1,492	88	51%	28%	8%	80%
Bayfield	1,405	92	28%	19%	2%	89%
Dolores RE-2	309	22	51%	15%	0%	91%
Dolores RE-4	689	53	36%	14%	2%	90%
Ignacio	751	67	51%	63%	5%	56%
Mancos	369	32	58%	26%	8%	79%
Montezuma-Cortez	2,929	188	62%	46%	10%	56%
Silverton	65	8	72%	32%	26%	100%
Del Norte	567	51	55%	55%	4%	95%
Salida	1070	85	41%	20%	3%	94%
Miami-Yoder	294	27	51%	20%	2%	82%
Estes Park	1159	89	33%	28%	14%	88%
Boulder*	29526	1864	18%	29%	10%	85%

** Boulder only participated in the spring pilot and only included a sample of schools. Surveys were also administered to a small number of teachers in Adams-Arapahoe, Greeley, Harrison, and Denver Public Schools as part of another research project focused on first year teachers.*

Results & Instrument Revision

Preliminary analyses of item difficulty, discrimination, and reliability from the fall pilot instrument were, once again, promising. Cronbach's Alpha improved from the psychometric field test – to 0.92 for the 3-5 fall pilot instrument and 0.96 for 6-12 – indicating that revisions to the items/instruments generally increasing instrument reliability.

Tables 4 and 5 (on the following pages) present two estimates for each item that were used to evaluate overall item performance and make instrument revisions: (a) the corrected item-total correlations, and (b) the reliability estimate (α) should that item be deleted. Several items were flagged from these analyses, and from additional IRT analyses that were conducted but are not presented here for parsimony.

The item "*School work in this class is too easy*" was included on the fall pilot instrument because teachers expressed a concern that student responses would be driven almost entirely by the perceived rigor (or lack thereof) of the course. Analyses suggested, however, that this item was not correlated to the total score (point-biserial $r = 0.02$), which suggests that student responses are not strongly driven by the perceived difficulty of the course. This item was also not highly correlated ($r > 0.25$) with *any* other item on the fall pilot instrument. As such, it was removed from the spring pilot instrument for both grade spans.

Several other items were altered and/or removed on the basis of analyses from the fall pilot. "*We waste time in this class*," was changed to "*Our class stays busy and does not waste time*" for elementary grades and eliminated entirely for secondary. "*Student behavior in this class makes the teacher angry*" was changed to "*The way students behave in this class makes it hard to learn*" for both grade spans.[3] The secondary item "*The classroom materials, pictures, words, books, and art reflect my cultural background*" was edited for clarity to "*Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background.*"

Teacher Engagement and Feedback Throughout the Preliminary Development Process

We were committed to engaging teachers and students throughout the survey's development and pilot process. During the development process, more than 100 Colorado teachers were given the opportunity to provide formal feedback on the appropriateness and utility of the questions. The teachers who participated in the psychometric field test were also given the opportunity to provide formal feedback about the instrument and its administration. We conducted additional teacher interviews and focus groups following the fall 2012 administration; those teachers were

[3] The revised item ("*The way students behave in this class makes it hard to learn*") was ultimately cut from the final Student Perception Survey because it did not perform well in the spring pilot.

Table 4. *Item summary statistics – fall 2012 pilot grades 3-5*

Item	Item-Total Correlation	α if Item Deleted
The schoolwork we do helps me learn.	0.551	0.914
The schoolwork we do is interesting.	0.477	0.914
What I learn in this class is useful to me in my real life.	0.451	0.915
I get bored in this class. (*negatively-worded item)*	0.472	0.915
In this class, we learn a lot almost every day.	0.533	0.914
My teacher makes sure that we think hard about things we read and write.	0.464	0.915
When the work is too hard, my teacher helps me keep trying.	0.594	0.913
In this class, it is more important to understand the lesson than to memorize the answers.	0.365	0.916
My teacher uses a lot of different ways to explain things.	0.505	0.914
My teacher knows when we understand the lesson and when we do not.	0.522	0.914
Our classroom materials and supplies have a special place and things are easy to find.	0.406	0.915
My teacher tells us what we are learning and why.	0.499	0.914
My teacher wants us to share what we think.	0.450	0.915
Students feel comfortable sharing their ideas in this class.	0.467	0.915
My teacher talks to me about my work to help me understand my mistakes.	0.596	0.913
My teacher writes notes on my work that help me do better next time.	0.461	0.915
My teacher builds on things we learn in other classes, subjects, and years.	0.508	0.914
My teacher cares about me.	0.601	0.913
If I am sad or angry, my teacher helps me feel better.	0.629	0.912
My teacher would notice if something was bothering me.	0.601	0.913
We waste time in this class. (*negatively-worded item)*	0.331	0.916
Students in my class are respectful to our teacher.	0.385	0.916
My classmates behave the way my teacher wants them to.	0.414	0.915
All of the kids in my class know what they are supposed to be doing and learning.	0.422	0.915
Students behave so badly in this class that it slows down our learning. (recode)	0.290	0.917
The people we learn and read about in this class are like me.	0.388	0.916
My teacher teaches us to respect people's differences.	0.534	0.914
In this class, I feel like I fit in.	0.568	0.913
I feel like an important part of my classroom community.	0.578	0.913
My teacher knows what my life is like outside of school.	0.437	0.915
My teacher knows what is important to me.	0.629	0.912
School work in this class is too easy. (*negatively-worded item)*	0.046	0.920
I ask for help when I need it.	0.454	0.915
I feel like I do a good job in this class.	0.465	0.915
The schoolwork we do helps me learn.	0.551	0.914
The schoolwork we do is interesting.	0.477	0.914
What I learn in this class is useful to me in my real life.	0.451	0.915

Table 5. *Item summary statistics – fall 2012 pilot grades 6-12*

Item	Item-Total Correlation	α if Item Deleted
My teacher makes learning enjoyable.	0.750	0.954
What I learn in this class is useful to me in my real life.	0.593	0.955
My teacher teaches things that are important to me.	0.651	0.954
My teacher knows the things that make me excited about learning	0.723	0.954
I get bored in this class. (*negatively-worded item)*	0.644	0.954
In this class, we learn a lot every day.	0.600	0.955
In this class, it is more important to understand the lesson than to memorize the answers.	0.492	0.955
When the work is too hard, my teacher helps me keep trying.	0.738	0.954
My teacher accepts nothing less than my best effort.	0.559	0.955
My teacher knows when we understand the lesson and when we do not.	0.685	0.954
If I don't understand something, my teacher explains it a different way.	0.715	0.954
My teacher explains difficult things clearly.	0.737	0.954
My classroom is organized and I know where to find what I need.	0.520	0.955
Students feel comfortable sharing their ideas in this class.	0.617	0.955
My teacher respects my opinions and suggestions.	0.724	0.954
In this class, we have a say in what we learn and do.	0.556	0.955
My teacher talks to me about my work to help me understand my mistakes.	0.699	0.954
My teacher writes notes on my work that help me improve.	0.540	0.955
When we study a topic, my teacher makes connections to other subjects or classes.	0.585	0.955
My teacher cares about me.	0.755	0.954
My teacher pays attention to what all students are thinking and feeling.	0.759	0.954
My teacher would notice if something was bothering me.	0.699	0.954
We waste time in this class. (*negatively-worded item)*	0.463	0.955
Students in this class treat the teacher with respect.	0.529	0.955
The students behave the way my teacher wants them to.	0.542	0.955
Student behavior in this class makes the teacher angry. (recode)	0.340	0.956
The classroom materials, pictures, words, books, and art reflect my cultural background.	0.434	0.956
My teacher respects my cultural background.	0.551	0.955
My teacher respects me as an individual.	0.692	0.954
Students in this class respect each other's differences.	0.455	0.955
In this class, I feel like I fit in.	0.606	0.955
I feel like an important part of this classroom community.	0.677	0.954
My teacher knows what my life is like outside of school.	0.520	0.955
My teacher knows what is important to me.	0.684	0.954
School work in this class is too easy. (*negatively-worded item)*	0.020	0.958
I ask for help when I need it.	0.522	0.955
I feel like I do a good job in this class.	0.537	0.955

given the opportunity to react to their experience in the fall, provide feedback on the instrument itself, and also suggest reporting formats. Teachers who participated in the fall pilot were also given an online survey about the instrument and process of administration. On this survey we received feedback from over 1,100 teachers across the state.

Throughout the development process, items were changed or eliminated when they demonstrated poor statistical properties (e.g., they did not correlate well with the overall instrument) and in some cases when teachers found them to be inappropriate or confusing to teachers. For example, between the fall and spring administration, we removed the item "I get bored in this class" because many teachers found it troubling.

2.5. Spring Validation Pilot Overview

Pilot Design

The final pilot administration occurred in the spring of 2013 across 16 Colorado districts (see Table 5 above); this final pilot was designed and conducted primarily as a validation study of the Student Perception Survey . Again, Appendix B presents a comparison of the three versions of each instrument: the field test instrument, fall pilot instrument, and final Student Perception Survey. The version of the instrument tested in the spring validation pilot is exactly the same as the final SPS, with one notable exception: The item "*The way students behave in this class makes it hard to learn*" was removed.

Participants

Grades 3-5. For the Grade 3-5 Student Perception Survey, there were 8,715 student survey assignments in 16 Colorado districts. Tables 6 and7 present the number of surveys assigned by grade level and district. Thompson School District contributed the greatest number of students (35.3 percent of the total sample), followed by Eagle County and Montezuma-Cortez (19.6 and 8.6 percent, respectively). Within the population of assigned surveys, students were distributed evenly in grades 3 through 5.

Table 6. *Spring pilot participants & population, by grade, grades 3-5*

Grade	Number of Responses	Percent of Population
Grade 3	2822	32.4%
Grade 4	2877	33.0%
Grade 5	3016	34.6%
Total	8715	100.0%

Table 7. *Spring pilot participants & population, by district, grades 3-5*

District Name	Number of Responses	Percent of Population
Archuleta County 50 JT	433	5.0%
Bayfield 10 JT-R	626	7.2%
Boulder Valley RE 2	116	1.3%
Centennial R-1	80	0.9%
Del Norte C-7	198	2.3%
Dolores County RE No.2	60	0.7%
Dolores RE-4a	320	3.7%
Eagle County RE 50	1708	19.6%
Ignacio 11 JT	279	3.2%
Mancos RE-6	106	1.2%
Miami/Yoder 60 JT	98	1.1%
Montezuma-Cortez RE-1	747	8.6%
Park (Estes Park) R-3	508	5.8%
Salida R-32	340	3.9%
Silverton	19	0.2%
Thompson R2-J	3077	35.3%
Total	8715	100%

Of the 8,715 surveys assigned to students in grades 3-5, 1,187 (13.6 percent) were never initiated online. These surveys are most likely missing due to student absences, although there could be a small number that are missing for other reasons (e.g., students declined participation in the survey, students were assigned a teacher they did not have and did not follow directions to indicate that on the survey). An additional 63 surveys were invalidated or removed for a variety of reason (see Table 13 in the "Summary of invalidated data" discussion below), leaving 7,465 records for a total response rate of 85.7 percent.

Among respondents, however, item-level response rates are much higher, despite the fact that all items were optional in the online system. Figure 1 presents the distribution of total number of items completed. The overwhelming majority (78.8 percent) responded to all 35 questions; 92.8 percent skipped two or fewer questions.

Figure 1. *Number of items completed by grade 3-5 respondents*

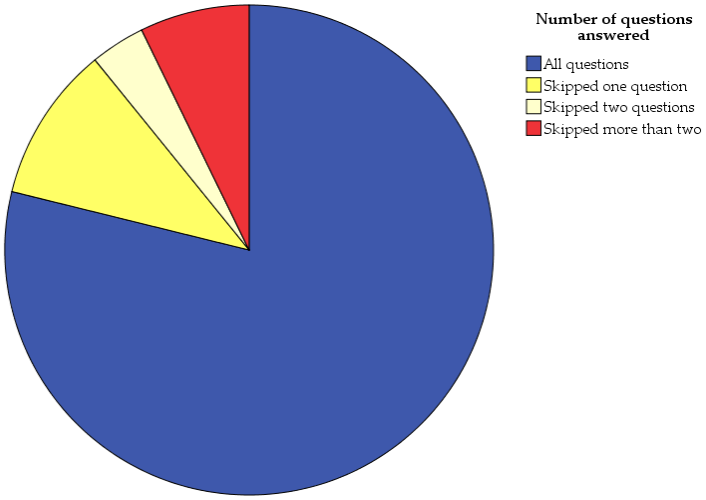


Table 8 presents item-level response rates for grades 3-5, which are also quite high (ranging from 95.0 to 99.0 percent). Response rates are slightly lower for three items that explicitly ask about reading and/or written work: "*The people we learn and read about in this class are like me*," "*My teacher writes notes on my work that help me do better next time*," and "*My teacher makes sure that we think hard about things we read and write*." Empirical results suggest that these items are missing most often when students are responding about specialist teachers (e.g., P.E., music).

Table 8. *Spring pilot response rates by item, grades 3-5*

Item	Number of Responses	Item response rate
1. The schoolwork we do helps me learn.	7380	98.9%
2. The schoolwork we do is interesting.	7373	98.8%
3. What I learn in this class is useful to me in my real life.	7357	98.6%
4. My teacher knows what makes me excited about learning.	7385	98.9%
5. In this class, we learn a lot almost every day.	7375	98.8%
6. My teacher makes sure that we think hard about things we read and write.	7091	95.0%
7. When the work is too hard, my teacher helps me keep trying.	7368	98.7%
8. In this class, it is more important to understand the lesson than to memorize the answers.	7243	97.0%
9. My teacher uses a lot of different ways to explain things.	7355	98.5%
10. My teacher knows when we understand the lesson and when we do not.	7374	98.8%
11. Our classroom materials/supplies have a special place & things are easy to find.	7366	98.7%
12. In this class, we learn to correct our mistakes.	7334	98.2%
13. My teacher tells us what we are learning and why.	7392	99.0%
14. My teacher wants us to share what we think.	7352	98.5%
15. My teacher asks questions to be sure we are following along.	7336	98.3%

Item	Number of Responses	Item response rate
16. Students feel comfortable sharing their ideas in this class.	7346	98.4%
17. My teacher talks to me about my work to help me understand my mistakes.	7334	98.2%
18. My teacher writes notes on my work that help me do better next time.	7174	96.1%
19. My teacher talks about things we learn in other classes, subjects, and years.	7229	96.8%
20. My teacher cares about me.	7299	97.8%
21. If I am sad or angry, my teacher helps me feel better.	7232	96.9%
22. My teacher would notice if something was bothering me.	7344	98.4%
23. Our class stays busy and does not waste time.	7394	99.0%
24. Students in my class are respectful to our teacher.	7378	98.8%
25. My classmates behave the way my teacher wants them to.	7387	99.0%
26. All of the kids in my class know what they are supposed to be doing & learning.	7383	98.9%
27. The way students behave in this class makes it hard to learn.*	7338	98.3%
28. The people we learn and read about in this class are like me.	7150	95.8%
29. My teacher teaches us to respect people's differences.	7283	97.6%
30. In this class, I feel like I fit in.	7357	98.6%
31. I feel like an important part of my classroom community.	7341	98.3%
32. My teacher knows what my life is like outside of school.	7312	98.0%
33. My teacher knows what is important to me.	7282	97.5%
34. I ask for help when I need it.	7391	99.0%
35. I feel like I do a good job in this class.	7360	98.6%

* This item was ultimately removed from the final version of the SPS instrument

Grades 6-12. For the Grade 6-12 Student Perception Survey, there were 30,627 surveys assigned to students in 16 Colorado districts and one online e-school. Tables 9 and 10 present the number of surveys assigned by grade level and district. Within the population of assigned surveys, students were distributed evenly in grades 6-12.

Table 9. *Spring pilot participants & population, by grade, grades 6-12*

Grade	Number of Responses	Percent of Population
Grade 6	4360	14.2%
Grade 7	4394	14.3%
Grade 8	4419	14.4%
Grade 9	4595	15.0%
Grade 10	4620	15.1%
Grade 11	4186	13.7%
Grade 12	4053	13.2%
Total	30627	100%

Thompson School District contributed the greatest number of students (47.9 percent of the total sample), followed by Eagle County and Montezuma-Cortez (12.2 and 7.6 percent, respectively).

Table 10. *Student Perception Survey population, by district, grades 6-12*

District Name	Number of Responses	Percent of Population
Archuleta County 50 JT	1421	4.6%
Bayfield 10 JT-R	1399	4.6%
Boulder Valley RE 2	1390	4.5%
Centennial R-1	198	0.6%
Del Norte C-7	536	1.8%
Dolores County RE No.2	241	0.8%
Dolores RE-4a	772	2.5%
Eagle County RE 50	3738	12.2%
Ignacio 11 JT	690	2.3%
Mancos RE-6	339	1.1%
Miami/Yoder 60 JT	298	1.0%
Montezuma-Cortez RE-1	2336	7.6%
Park (Estes Park) R-3	1263	4.1%
Salida R-32	1210	4.0%
BOCES E-School	64	0.2%
Silverton	55	0.2%
Thompson R2-J	14677	47.9%
Total	30627	100%

Of the 30,627 surveys assigned to students in grades 6-12, 8,010 (26.7 percent) were never initiated online. These surveys are most likely missing due to student absences, although there could be a small number that are missing for other reasons (e.g., students declined participation in the survey, students were assigned a teacher they did not have and did not follow directions to indicate that on the survey). An additional 790 surveys were invalidated or removed for a variety of reason (see Table 16 in the "**Summary of invalidated data**" section below), leaving 21,827 records for a total response rate of 71.3 percent.

Among secondary respondents, item-level response rates are slightly higher than among elementary students. Figure 2 presents the distribution of items completed. The overwhelming majority (87.1 percent) responded to all 35 questions; 97.3 percent skipped two or fewer questions.

Figure 2. *Number of items completed by grade 6-12 respondents*

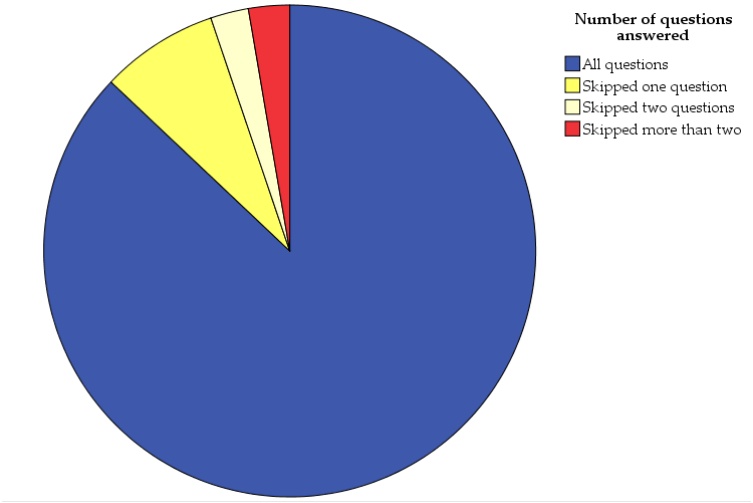


Table 11. *Spring pilot response rates by item, grades 6-12*

Item	Number of Responses	Item response rate
1. My teacher makes learning enjoyable.	21800	99.9%
2. What I learn in this class is useful to me in my real life.	21680	99.3%
3. My teacher teaches things that are important to me.	21710	99.5%
4. My teacher knows the things that make me excited about learning.	21728	99.5%
5. In this class, we learn a lot every day.	21757	99.7%
6. In this class, it is more important to understand the lesson than to memorize the answers.	21686	99.4%
7. When the work is too hard, my teacher helps me keep trying.	21692	99.4%
8. My teacher accepts nothing less than my best effort.	21702	99.4%
9. My teacher knows when we understand the lesson and when we do not.	21685	99.3%
10. If I don't understand something, my teacher explains it a different way.	21691	99.4%
11. My teacher explains difficult things clearly.	21704	99.4%
12. My classroom is organized and I know where to find what I need.	21691	99.4%
13. Students feel comfortable sharing their ideas in this class.	21682	99.3%
14. My teacher respects my opinions and suggestions.	21649	99.2%
15. In this class, we have a say in what we learn and do.	21653	99.2%
16. My teacher talks to me about my work to help me understand my mistakes.	21674	99.3%
17. My teacher writes notes on my work that help me improve.	21603	99.0%
18. When we study a topic, my teacher makes connections to other subjects or classes.	21630	99.1%
19. My teacher cares about me.	21517	98.6%

Item	Number of Responses	Item response rate
20. My teacher pays attention to what all students are thinking and feeling.	21589	98.9%
21. My teacher would notice if something was bothering me.	21593	98.9%
22. Our class stays busy and does not waste time.	21689	99.4%
23. Students in this class treat the teacher with respect.	21675	99.3%
24. The students behave the way my teacher wants them to.	21651	99.2%
25. The way students behave in this class makes it hard to learn.*	21662	99.2%
26. Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background.	20960	96.0%
27. My teacher respects my cultural background.	21048	96.4%
28. My teacher respects me as an individual.	21538	98.7%
29. Students in this class respect each other's differences.	21598	99.0%
30. In this class, I feel like I fit in.	21619	99.0%
31. I feel like an important part of this classroom community.	21583	98.9%
32. My teacher knows what my life is like outside of school.	21525	98.6%
33. My teacher knows what is important to me.	21523	98.6%
34. I ask for help when I need it.	21624	99.1%
35. I feel like I do a good job in this class.	21652	99.2%

* This item was ultimately removed from the final version of the SPS instrument

Section 3: Survey Administration & Logistics

3.1. Survey Administration Conditions Overview

Although the student perception survey is not a high-stakes assessment, survey administration is very important. Ensuring that students take the surveys under appropriate conditions will help ensure that their feedback is reliable and useful.

3.2. Operational Procedures

Proctor administration guides provided detailed information about the testing environment, procedures, security, and instructions to students. For reference, the Proctor Guide can be found here. The guides contained administration instructions and scripts in English and Spanish. Teachers and/or proctors were instructed to follow the instructions and read the accompanying directions verbatim to students. The manuals included instructions for both paper-and-pencil and online administrations, and provided an FAQ for proctors to address student questions regarding the instrument.

In general, the operational procedures for both the fall pilot and spring validation pilot were designed with several priorities in mind:

1. *Districts and schools should take steps to ensure that students know their confidentiality is being maintained.*

   At the elementary level, teachers did not proctor surveys for their own students. This was not meant to isolate the teacher or create an atmosphere of secrecy, but instead to ensure that students felt comfortable answering honestly. Due to the recommended random sampling criteria at the secondary level, teachers may have proctored a class where some students were completing a survey about them. This did not mean that the survey administration was overly-prescriptive or formal; for example, teachers could swap classes with colleagues or work with instructional assistants or other staff members to proctor the survey at a convenient time.

2. *The words and actions of teachers and administrators should also communicate to students that their answers are private and that they should feel comfortable answering honestly.*

   CEI modeled this belief in communication materials shared with relevant staff and in proctor guides and oral instructions for teachers.

3. *The sampling of students is important.*

   Ensuring an adequate and representative sample for each teacher was critical to building buy-in and attaining fair, representative results for teachers.

4. *Finally, we value instructional time.*

   Although the student surveys yield important information that can be used to inform instructional practice, it is more critical that teachers and students have access to necessary instructional time. In designing the sampling approach, we felt that it was important to limit the number of times students were surveyed (e.g., we set a target of two or fewer times) and to generally protect instructional time (e.g., by trying to set a limit that no student give up more than roughly one hour of instructional time).

Recommended Administration Windows

For administrations moving forward, [4] we recommend that districts administer Colorado's Student Perception Survey in the early part of the year to ensure that teachers receive timely and actionable feedback. The pilot districts found that a window between November and January was best in this regard.

Confidentiality of Student Records

Throughout the pilot, CEI worked diligently to ensure that the confidentiality of student records were protected in compliance with the Family Educational Rights and Privacy Act, 34 CFR & 99 et. Seq., and Colorado Administrative Rules relating to student records.

Student-Teacher Assignment

For both the fall pilot and spring validation pilot, students were pre-assigned to respond about a sample of their teachers. At the elementary level, all students were assigned to respond to the survey regarding their homeroom teacher. If districts opted to survey students about specialists at the elementary level, there were a handful of approaches (e.g., students were randomly assigned to survey one of the specialists in their school, or entire classes were randomly assigned to respond about a given specialist to ease potential confusion for young students). At the secondary level, students were randomly sampled to respond about two of their teachers; this random assignment ensured that students did not know in advance who they would be

[4] For more detailed information about how to administer and plan a Student Perception Survey, please consult the comprehensive planning guide available as part of the Student Survey Toolkit online.

responding about and that teachers/proctors did not know what teacher any given student was responding about.

Most districts piloting student surveys across the country have relied on a convenience sample of students within a select set of pre-determined class periods (e.g., all students are surveyed about their teachers during 2nd and 7th period). While there are numerous benefits to this approach, including the simplicity of administration and lack of need for pre-populated surveys with student-teacher links already assigned, we opted for a more robust sampling framework for several reasons. First, teachers in pilot districts strongly objected to the period sampling approach; they felt that this choice could unfairly skew their data should the selected period not adequately represent their entire population of students. Second, in our smaller districts, we found that the period sampling approach yielded sample sizes below the recommended threshold for many teachers and/or excluded some teachers entirely, largely due to free/prep periods during which some teachers had no assigned students. When we attempted to schedule "make-up" sessions to compensate during subsequent or preceding periods in the day, we found that some students were being surveyed five or even six times, and that those students being surveyed multiple times were disproportionately likely to receive special services (e.g., special education or ELL services). In general, this pattern is unsurprising, given that teachers providing special services often have small class sizes that make it difficult to yield a large sample size.

3.3. Testing Accommodations

In general, districts were advised to use the IEPs of students with disabilities to identify any accommodations those students might need. Decisions to exclude students from the survey should be made on a case-by-case basis when the nature and severity of a student's disability is such that the student likely would not be able to meaningfully complete the survey. Alternate arrangements should be made for these students in advance of administration day.

Teachers were encouraged to work with district survey coordinators ahead of time to identify students that may need accommodations, and make alternate arrangements if an individual proctor or classroom space was needed.

Because this survey does not measure student achievement, but instead asks for student perceptions of their teacher, we were open to most accommodations that may be needed to support the student in completing the survey. However, we asked that the teacher being assessed was not the teacher that provides these accommodations. The Proctor Protocols includes a list of standard TCAP accommodations for reference, but districts and teachers were also advised to utilize other supports as needed.

Section 4: Spring Validation Pilot – Student & Teacher Results

4.1 Final Results from Spring Validation Pilot

Creating the Analysis File & Ensuring Data Quality

Several safeguards were built into the online survey platform during the spring validation pilot to ensure data quality before attributing student responses to teachers, and to address the potential problems identified during the fall pilot. First, to allow student responses to be matched to the assigned teacher, each student was given a unique code for the teacher(s) about whom they were surveyed. Elementary students were given simple word codes, while secondary students received 4-5 digit numerical codes. Students entered these codes into the survey platform, instead of their names, to ensure confidentiality. Students were then asked to verify the teacher they were assessing. If the student-entered teacher name did not match the assigned teacher, the survey response was invalidated. Students were finally asked to verify their school name. This was done primarily to ensure that if a code was mistyped (i.e., dogw instead of dogs), we could look at the school name, teacher name, and identify the correct code to form the student-teacher link.

Secondary students were also asked to verify that the course they were assigned to respond about was in fact an instructional course. If students indicated that it was not they were redirected to the beginning of the survey and did not complete the survey for that teacher. Elementary and secondary students were also asked to indicate how long they have had their teacher and how often they have class with the teacher to ensure students were responding only about teachers with whom they actually had a course for a sufficient period of time.

Data analysis preparation included several quality checks:

- As mentioned above, students entered their grade, school name and teacher name(s). These were compared to the original district data to ensure quality.
- Duplicate codes, although rare, were removed from the data. Often these occurred when students had been kicked-out of the survey due to web connectivity issues. In these instances, only one completed survey from each student was included in the analysis.
- In a few instances, data were invalidated when students indicated that their course was no longer taught by their teacher (e.g., they currently had a long-term substitute or student teacher).

Summary of invalidated data for grades 3-5. In the elementary grades, a very small percentage of student records were eliminated. Only a handful of surveys were invalidated because a student entered a teacher name in the validation field that did not match the assigned teacher (N=12, or 0.1 percent of the total population).

In total, there were an additional 41 instances (0.5 percent of the total population) where students indicated that they had been assigned to take the survey about a teacher that they did not have.

Table 12. *Responses to grade 3-5 teacher verification question*

"How often do you have class with this teacher?"	**Number of Responses**	**Percent**
Every day	5780	66.4%
2-3 times a week	497	5.7%
Once a week	1116	12.8%
Less than once a week	80	0.9%
I do not have this teacher.	41	0.5%
Missing*	1189	13.7%
Total	8703	100.0%

** The missing responses include 1,187 surveys that were never initiated online, and 2 additional responses that were missing on this item.*

Finally, an additional 10 responses were removed from the analyses because the student began the survey, completed the introductory questions, and then did not complete any of the items about their teacher. In total, 1,250 responses (see Table 13) were removed from the analysis file, leaving 7,465 responses from students in grades 3-5.

Table 13. *Summary of responses excluded from analysis file, grades 3-5*

Reason for exclusion from analysis file	**Number of Responses**	**Percent of Population**
Mismatched student-entered teacher name	12	0.1%
Student response – I don't have this teacher	41	0.5%
Missing – Initiated survey but didn't complete any teacher questions	10	0.1%
Missing – Never initiated the survey	1187	13.6%
Total number of removed from analysis file	1250	14.3%

Summary of invalidated data for grades 6-12. Again, only a small percentage of student records were invalidated for grades 6-12; despite this, there were a higher percentage of student surveys missing in grades 6-12 than in 3-5. Thirty records (0.1 percent of the total population) were removed because students indicated the teacher that was no longer at the school (i.e., in this case, all 30 students were assigned to the same teacher who was out on long-term leave). Another 133 surveys (0.4 percent of the total population) were invalidated because a student entered a teacher name in the validation field that did not match the assigned teacher.

In total, there were an additional 564 instances (1.8 percent of the total population) where student responses to the built-in data quality items immediately terminated the survey, including when the students had been assigned to take the survey about a teacher that they did not have (N=131) or about a class where instruction did not take place (N=433).[5]

Table 14. *Responses to grade 6-12 teacher verification question*

"How long have you had this teacher?"	**Number of Responses**	**Percent**
I have had this teacher all year.	14874	48.6%
I have had this teacher all term/semester.	7252	23.7%
I have had this teacher for less than 1 month.	325	1.1%
I do not have this teacher.	134	0.4%
Missing	8042*	26.3%
Total	30627	100.0%

** Missing responses include 8,010 surveys never initiated online, and 32 responses that were missing on this item.*

Table 15. *Responses to grade 6-12 instruction course verification question*

"Is this a class where no instruction takes place (such as homeroom, advisory or office assistant)?"	**Number of Responses**	**Percent**
No, this is a class with instruction.	22001	71.8%
Yes, this is a class without instruction.	438	1.4%
Missing	8168	26.7%
Total	30627	100.0%

Finally, an additional 63 responses were removed from the analyses because the student began the survey, completed the introductory questions, and then did not complete any of the items about their teacher. In total, 8,800 responses (see Table 16 on page 32) were removed from the analysis file, leaving 21,827 responses from students in grades 6-12.

4.2 Descriptive Statistics & Classical Test Theory Results

After the spring validation pilot ended, we conducted rigorous item analyses, incorporating methodology from both classical test theory and item response theory (IRT). This section will primarily outline results from (a) basic descriptive statistical analyses and (b) classical test theory, while the subsequent section will summarize results from IRT analyses. Classical test

[5] Note: Numbers in Table 14 and 15 may vary slightly than those outlined above; when this is the case, it is because some surveys included in this these question analyses that were invalidated for other reasons.

theory allows for analysis of item-level data using the raw responses from Colorado's Student Perception Survey.

Table 16. *Summary of responses excluded from analysis file, grades 6-12*

Reason for exclusion from analysis file	Number of Responses	Percent of Population
Flagged for removal – Invalid teacher assignment	30	0.1%
Mismatched student-entered teacher name	133	0.4%
Student response – I don't have this teacher	131	0.4%
Student response – Non-instructional course	433	1.4%
Missing – Initiated survey but didn't complete any teacher questions	63	0.2%
Missing – Never initiated the survey	8010	26.2%
Total number removed from analysis file	8800	28.7%

Student-Level Results

Tables 17 and 18 below present the most common response option and mean score for each item on the Student Perception Survey for grades 3-5 and grades 6-12. At the elementary level, there are 19 items where the most common response is "Always," 12 were it is "Most of the time," three "Some of the time," and one "Never." At the secondary level, 14 items have "Always" as the most frequent response, 14 have "Most of the time," six "Some of the time," and – once again – one has "Never" as the most common response. The same item for students in grades 3-5 and 6-12 – "*My teacher knows what my life is like outside of school*" – is the one that is most likely to garner a response of "Never."[6]

This variety in response patterns is promising – a good instrument will include items at various difficulty levels that appeal to students with a variety of perceptions. It is desirable, from a psychometric perspective, to include items that are both difficult and relatively easy. The statistics provided in Tables 17 and 18 describe difficulties for each item on each of the instruments (grades 3-5 and 6-12). Histograms for each item show the distributions across response options, and are presented in Appendix C and D. In general, the items are slightly negatively skewed (with more students responding "Always" and "Most of the time" than "Some of the time" and "Never"), although overall there is good variability both within and across items.

[6] Although students are much more likely to respond negatively to this item, it still demonstrates sound statistical properties overall and is highly correlated with other items and the total score.

Table 17. *Summary item descriptives, grades 3-5*

Item	Most Common Response	Mean[7]
1. The schoolwork we do helps me learn.	Most of the time	3.16
2. The schoolwork we do is interesting.	Most of the time	2.80
3. What I learn in this class is useful to me in my real life.	Always	2.99
4. My teacher knows what makes me excited about learning.	Some of the time	2.62
5. In this class, we learn a lot almost every day.	Always	3.22
6. My teacher makes sure that we think hard about things we read & write.	Always	3.32
7. When the work is too hard, my teacher helps me keep trying.	Always	3.14
8. In this class, it is more important to understand the lesson than to memorize the answers.	Always	3.10
9. My teacher uses a lot of different ways to explain things.	Always	3.03
10. My teacher knows when we understand the lesson and when we do not.	Most of the time	2.89
11. Our classroom materials/supplies have a special place & things are easy to find.	Always	3.27
12. In this class, we learn to correct our mistakes.	Always	3.33
13. My teacher tells us what we are learning and why.	Always	3.21
14. My teacher wants us to share what we think.	Most of the time	2.87
15. My teacher asks questions to be sure we are following along.	Always	3.12
16. Students feel comfortable sharing their ideas in this class.	Most of the time	2.79
17. My teacher talks to me about my work to help me understand my mistakes.	Always	3.00
18. My teacher writes notes on my work that help me do better next time.	Always	2.64
19. My teacher talks about things we learn in other classes, subjects, & years.	Some of the time	2.51
20. My teacher cares about me.	Always	3.39
21. If I am sad or angry, my teacher helps me feel better.	Always	2.88
22. My teacher would notice if something was bothering me.	Most of the time	2.77
23. Our class stays busy and does not waste time.	Most of the time	2.74
24. Students in my class are respectful to our teacher.	Most of the time	2.89
25. My classmates behave the way my teacher wants them to.	Most of the time	2.64
26. All of the kids in my class know what they are supposed to be doing & learning.	Most of the time	2.99
28. The people we learn and read about in this class are like me.	Some of the time	1.94
29. My teacher teaches us to respect people's differences.	Always	3.15
30. In this class, I feel like I fit in.	Always	2.97
31. I feel like an important part of my classroom community.	Always	2.86
32. My teacher knows what my life is like outside of school.	Never	2.05
33. My teacher knows what is important to me.	Most of the time	2.60
34. I ask for help when I need it.	Always	3.21
35. I feel like I do a good job in this class.	Always	3.23

[7] Consider this useful heuristic for interpreting the mean scores of items on Colorado's Student Perception Survey: A mean score of 2.5 can be thought of, generally, the point at which students are more likely to respond in the top two categories ("Always" or "Most of the time") than the bottom two categories ("Some of the time" or "Never"). Note: This does not mean that an item with a mean score under 2.5 will necessarily have the most common response – which is a simple plurality – in the bottom two categories, or vice versa.

Table 18. *Summary item descriptives, grades 6-12*

Item	Most Common Response	Mean
1. My teacher makes learning enjoyable.	Most of the time	2.87
2. What I learn in this class is useful to me in my real life.	Most of the time	2.91
3. My teacher teaches things that are important to me.	Most of the time	2.84
4. My teacher knows the things that make me excited about learning.	Some of the time	2.53
5. In this class, we learn a lot every day.	Most of the time	3.03
6. In this class, it is more important to understand the lesson than to memorize the answers.	Always	3.10
7. When the work is too hard, my teacher helps me keep trying.	Always	3.13
8. My teacher accepts nothing less than my best effort.	Always	3.31
9. My teacher knows when we understand the lesson and when we do not.	Most of the time	2.96
10. If I don't understand something, my teacher explains it a different way.	Always	2.97
11. My teacher explains difficult things clearly.	Most of the time	2.94
12. My classroom is organized and I know where to find what I need.	Always	3.38
13. Students feel comfortable sharing their ideas in this class.	Always	3.15
14. My teacher respects my opinions and suggestions.	Always	3.25
15. In this class, we have a say in what we learn and do.	Some of the time	2.44
16. My teacher talks to me about my work to help me understand my mistakes.	Most of the time	2.90
17. My teacher writes notes on my work that help me improve.	Always	2.69
18. When we study a topic, my teacher makes connections to other subjects/classes.	Some of the time	2.64
19. My teacher cares about me.	Always	3.18
20. My teacher pays attention to what all students are thinking and feeling.	Always	3.00
21. My teacher would notice if something was bothering me.	Most of the time	2.66
22. Our class stays busy and does not waste time.	Most of the time	2.88
23. Students in this class treat the teacher with respect.	Most of the time	3.06
24. The students behave the way my teacher wants them to.	Most of the time	2.81
26. Our classroom materials reflect my cultural background.	Some of the time	2.43
27. My teacher respects my cultural background.	Always	3.46
28. My teacher respects me as an individual.	Always	3.45
29. Students in this class respect each other's differences.	Most of the time	3.09
30. In this class, I feel like I fit in.	Always	3.11
31. I feel like an important part of this classroom community.	Most of the time	2.88
32. My teacher knows what my life is like outside of school.	Never	2.05
33. My teacher knows what is important to me.	Some of the time	2.53
34. I ask for help when I need it.	Always	3.18
35. I feel like I do a good job in this class.	Most of the time	3.18

Teacher-Level Results

Calculating teacher-level results. Teachers were only included in the teacher-level analyses if they had at least 8 student responses.[8] For teachers with greater than 8 responses, student responses were aggregated for each item, each of the four elements (Student Learning, Student-Centered Environment, Classroom Community, and Classroom Management), and for the overall instrument.

In general, there are two ways we recommend aggregating teacher-level results – by calculating (a) a mean score, or (b) a percent of responses in the top two categories, referred to as "percent favorable." In focus groups, teachers expressed a strong preference for the percent favorable method for its simplicity and for ease of interpretation (e.g., they argued that it is easier to understand the concept that 78 percent of students responded favorably to a given item than to interpret a mean score of 3.2).

Tables 19 and 20 present a summary of teacher-level mean score and percent favorable results for each of the four elements and an overall score for grades 3-5 and grades 6-12. In general, teacher-level results overall and in each element are quite positive.

Table 19. S*ummary of teacher-level results, grades 3-5*

	Mean	Median	SD	Minimum	Maximum
Mean Score					
Overall Score	3.0	3.0	0.3	1.8	3.5
Student Learning Element	3.1	3.2	0.3	1.8	3.7
Student-Centered Environment Element	2.7	2.7	0.3	1.6	3.3
Classroom Community Element	3.2	3.2	0.3	2.1	3.8
Classroom Management Element	2.8	2.9	0.3	1.9	3.5
Percent Favorable					
Overall Score	68.7	69.6	11.6	21.8	90.9
Student Learning Element	75.7	77.9	12.1	21.0	96.3
Student-Centered Environment Element	55.0	55.7	13.0	12.8	79.6
Classroom Community Element	75.9	77.7	11.4	31.7	100.0
Classroom Management Element	66.1	67.2	15.9	12.5	99.0

[8] For the purposes of generating teacher-level reports to share back with districts, schools, and teachers, we used a more conservative requirement of at least 10 student responses, although all participating teachers were given access to aggregate school- and district-data. For these analyses, however, which are presented only in aggregate, a less conservative value felt reasonable. In general, comparisons between the two populations (teachers with at least 10 responses and those with at least 8) suggest that there are no significant differences.

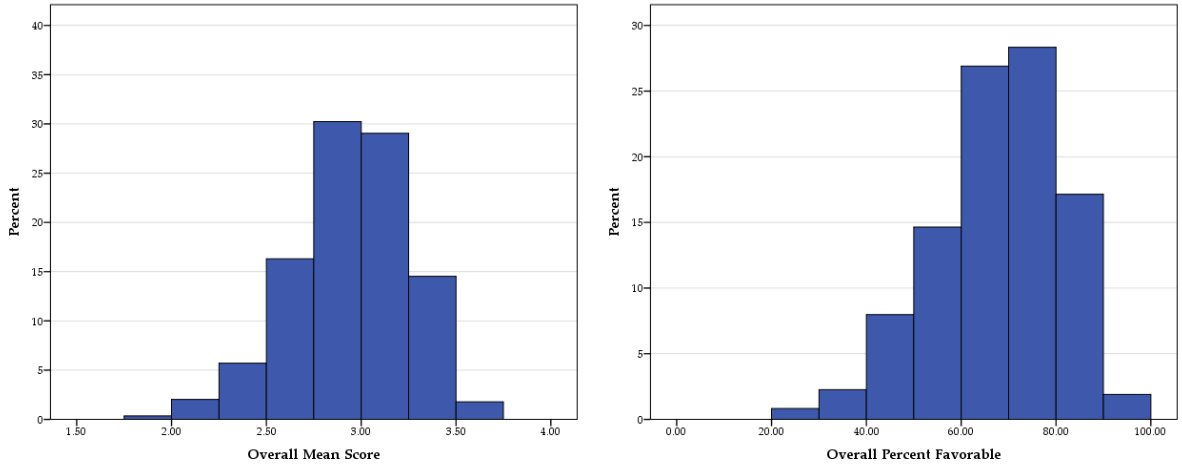
Interestingly, trends with regard to the four elements are reversed for students in grades 6-12 when compared to those in grades 3-5: For elementary teachers, the two highest categories are **Student Learning** and **Classroom Community**, whereas for secondary teachers **Student-Centered Environment** and **Classroom Management** were scored most positively. This may suggest that students experience these areas differently in elementary school than they do in middle and high school; on the other hand, it could also suggest that teachers in these two grade spans naturally have different strengths.

Table 20. S*ummary of teacher-level results, grades 6-12*

	Mean	Median	SD	Minimum	Maximum
Mean Score					
Overall Score	2.9	3.0	0.3	1.8	3.7
Student Learning Element	2.9	2.9	0.3	1.6	3.7
Student-Centered Environment Element	3.3	3.3	0.3	1.9	4.0
Classroom Community Element	2.8	2.8	0.3	1.9	3.5
Classroom Management Element	3.0	3.0	0.3	1.9	3.7
Percent Favorable					
Overall Score	67.8	69.1	13.5	20.5	95.0
Student Learning Element	65.4	66.7	15.5	11.7	97.1
Student-Centered Environment Element	79.3	82.1	13.4	23.4	100.0
Classroom Community Element	60.8	61.0	12.6	23.4	94.3
Classroom Management Element	72.5	74.8	15.4	15.6	100.0

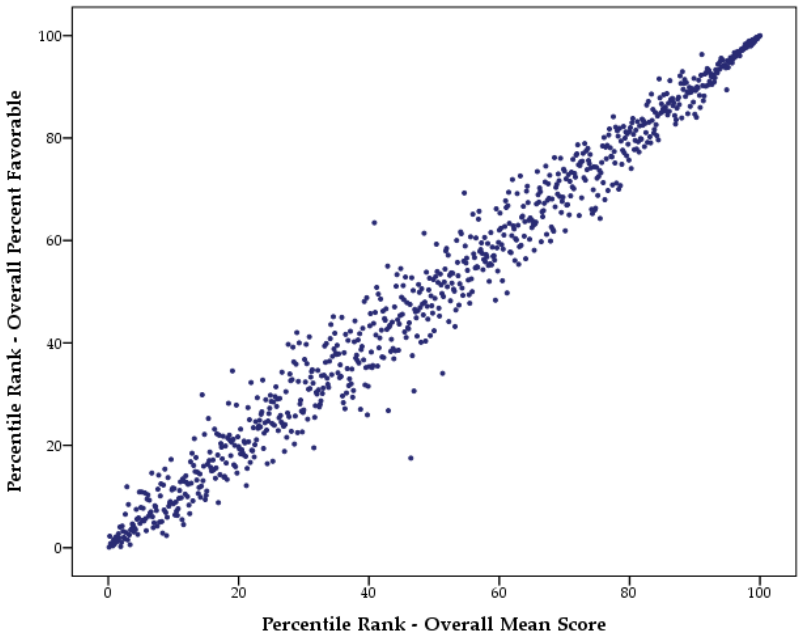
Nevertheless, in general, the two measures (mean score and percent favorable) produce nearly identical substantive results. Mean scores produce a slightly more normal distribution, while the percent favorable approach is slightly negatively skewed (Figure 3).

Figure 3. *Overall teacher mean score and percent favorable, grades 6-12*



When teachers are assigned a percentile rank based on each method, the results are almost perfectly correlated (r = 0.988). Figure 4 graphically compares teacher-level overall percent favorable and the overall mean score.

Figure 4. *Teacher-level overall percent favorable as a function of overall mean score*



Interpreting teacher-level pilot results. Overall, the instrument does differentiate between teachers. As Figure 3 demonstrates, distributions of teachers' overall scores (both mean scores and percent favorable) are relatively normal.

At the item-level, distributions of teachers' mean scores are also relatively normal, with mean scores ranging from 2.0 to 3.4 for grades 3-5 and 2.1 to 3.5 for grades 6-12. This suggests that items demonstrate differing levels of difficulty, with some yielding quite positive overall responses and others yielding more negative responses (a point that will be discussed in further detail in the Item Response Theory Analyses section below).

Furthermore, when teacher-level results are disaggregated by the four elements (Student Learning, Student-Centered Environment, Classroom Community, and Classroom Management), we once again see relatively normal distributions, both for mean score and overall percent favorable. Figures 5 and 6 presents the distributions of teacher overall mean scores by the four elements. On each of the four elements, the vast majority of teachers score above a 2.5, the threshold at which they are more likely to score in the top two categories ("Always" or "Most of the time") than the bottom two ("Some of the time" and "Never").

Figure 5. *Teacher distributions by survey element, grades 3-5*

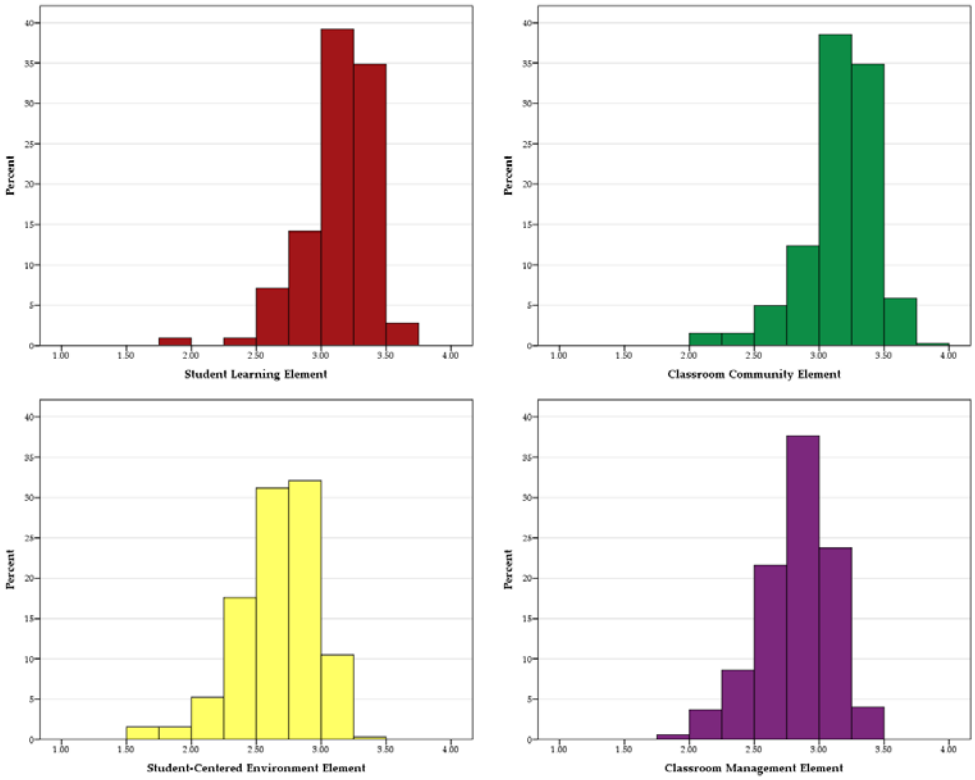
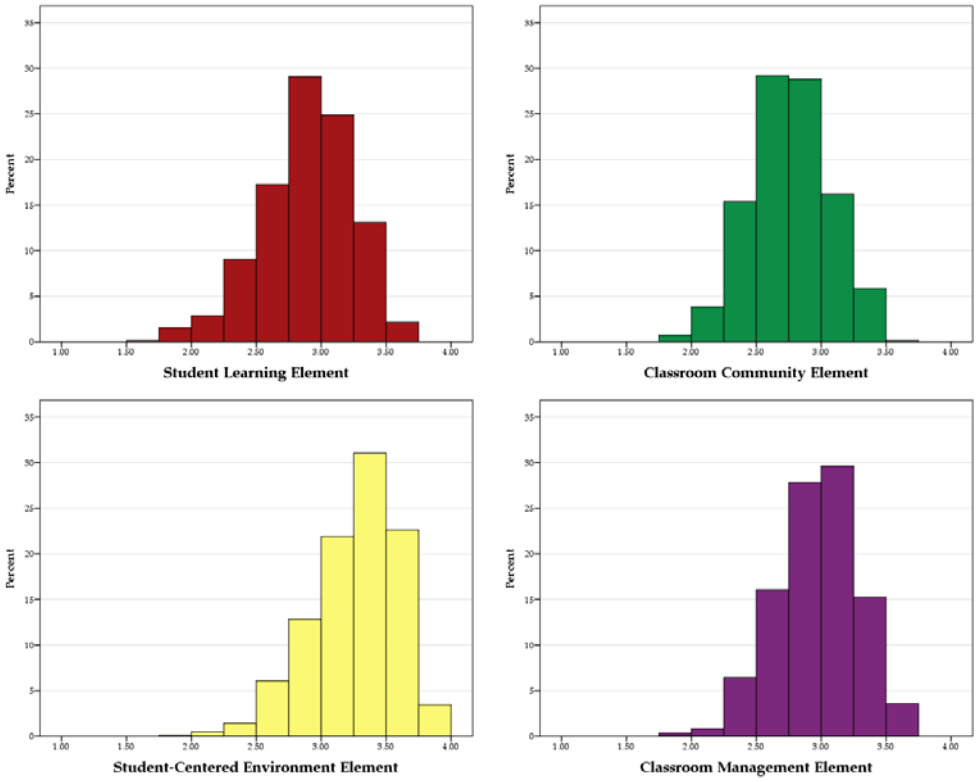


Figure 6. *Teacher distributions by survey element, grades 6-12*



4.3 Item Response Theory Analyses

Results from the pilot were analyzed using the Construct Map software developed at the BEAR research center at the University of California Berkeley. In the initial review of the data, it was determined necessary to eliminate an item – "*The way students behave in this class makes it hard to learn*" – from both survey instruments; not only did this item provide very little substantive information about teacher instruction, but it also performed poorly in classical and IRT analyses. For this reason, it was omitted from all analyses presented below and removed from the final Student Perception Survey instruments presented in Appendix A.

Overview of the IRT Model

Item-response theory (IRT) is a model-based approach to measuring latent traits. To further examine the properties of the student perception survey, we applied the Rasch Partial Credit Model (PCM; Masters, 1982) to estimate students perceptions of their teachers' instructional practices. IRT models are particularly useful because they model the probability of a given item response as a function of both a respondent's attitudes (represented by the common psychometric convention, θ) and the item's "difficulty." In PCM, this calculated with the following equation.

$$P_{ix}(\theta) = \frac{\exp\left[\sum_{j=1}^{x}(\theta - \delta_{ij})\right]}{\sum_{r=0}^{m_i}\left[\exp\sum_{j=1}^{r}(\theta - \delta_{ij})\right]}$$

where $P_{ix}(\theta)$ is the probability of a randomly chosen student, with overall perceptions θ, responding x on item i, and where

$$\sum_{j=1}^{x}(\theta - \delta_{ij}) \equiv 0$$

where δ_{ij} represents the relative difficulty δ for each score category j on item i.

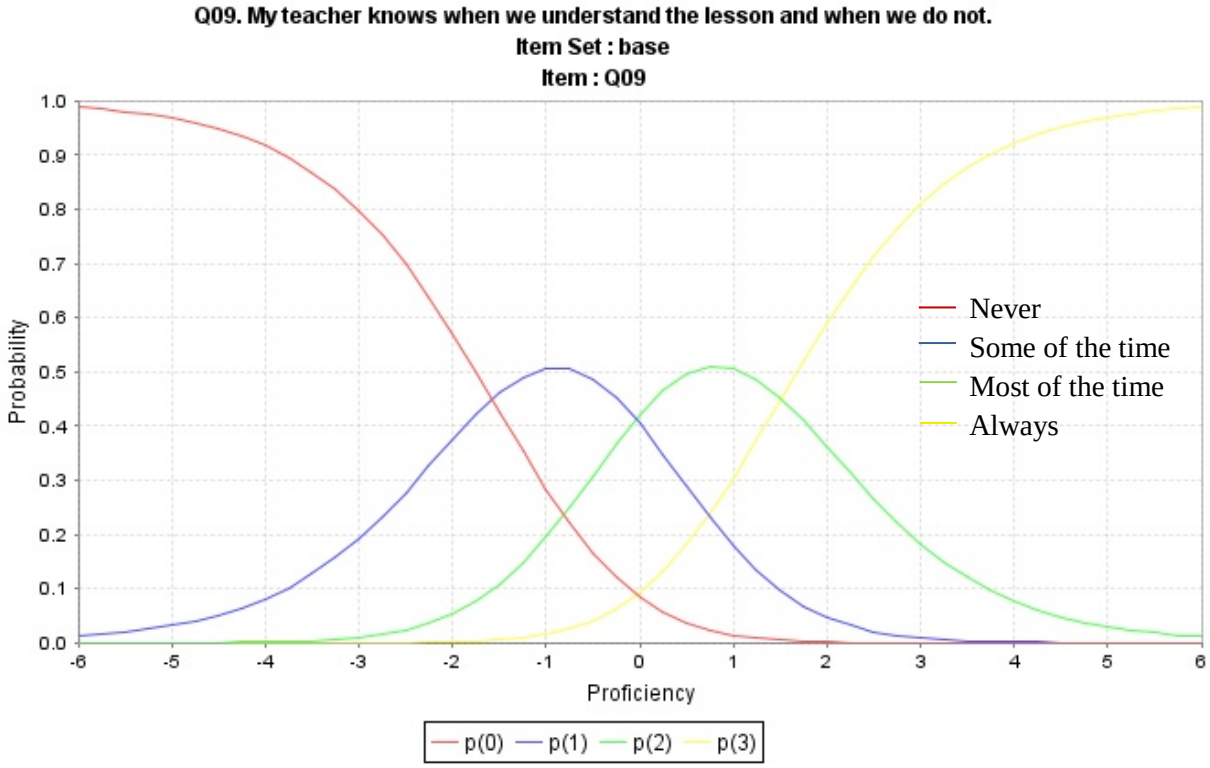
Item difficulty estimates are derived empirically, and in a traditional testing context, an item's difficulty estimate is relatively straightforward. Simply put, difficult items are those that very few people answer correctly while easy items are those that most people get right. Similar principles apply in a survey of perceptions. Again, item difficulty estimates are derived empirically, and difficult items are those where very few students respond in the highest categories ("Always" or "Most of the time"). Similarly, easy items are those where more students respond in those categories.

Item Analyses. Estimates of item difficulty from the Rasch model are useful, because they can help us assess the degree to which students in general report experiencing various instructional teaching behaviors. For example, the hardest item on the survey for grades 3-5 is "*The people we learn and read about in this class are like me*"[9] meaning that students were least likely to respond favorably to this item. For grades 6-12, the most difficult item was "*My teacher knows what my life is like outside of school.*"

Tables 21 and 22 (found on pages 42 and 43) present the item estimates for each item on both instruments. They also present error estimates and estimates of overall item fit (which will be addressed in the following section).

Item Characteristic Curves. The Item Characteristic Curve (ICC) maps the probability of each response on any given item as a function of the perceptions ($\hat{\theta}$) of students. A single ICC is presented below (the full set of ICCs can be found in Appendices C and D).

Figure 7. *ICC for "My teacher knows when we understand the lesson and when we do not"*



[9] An astute reader may recall that the item "*My teacher knows what my life is like outside of school*" was also the only item on the 3-5 instrument where the most common response was "Never," and may be surprised that this is not the most difficult item on the 3-5 survey (although it is the second most difficult item). In short, it is slightly less difficult than "*The people we learn and read about in this class are like me*" because, even though it had more students that responded "Never," students responded slightly more positively overall (e.g., across all students in grades 3-5, the mean score is slightly higher on this item).

Table 21. *Item difficulty and fit estimates, grades 3-5*

			Outfit		Infit	
Item[10]	**Estimate**	**Error**	**Mean Square**	**t**	**Mean Square**	**t**
Q01	-0.56	0.012	0.93	-4.0	0.97	-1.7
Q02	0.08	0.012	0.97	-1.9	0.98	-1.5
Q03	-0.13	0.011	1.12	6.5	1.10	5.6
Q04	0.51	0.011	0.97	-1.7	0.98	-1.0
Q05	-0.58	0.012	1.03	1.7	1.08	4.3
Q06	-0.46	0.012	1.20	10.6	1.21	10.0
Q07	-0.29	0.011	0.94	-3.8	0.98	-1.3
Q08	-0.26	0.012	1.16	8.7	1.18	9.6
Q09	-0.21	0.012	1.04	2.6	1.10	5.7
Q10	0.00	0.012	0.99	-0.4	1.02	1.4
Q11	-0.52	0.012	1.12	6.6	1.12	6.6
Q12	-0.68	0.012	1.01	0.5	1.06	3.3
Q13	-0.51	0.012	1.03	1.5	1.08	4.2
Q14	0.03	0.012	1.11	6.2	1.13	7.7
Q15	-0.33	0.012	1.09	5.3	1.15	8.0
Q16	0.11	0.012	1.03	1.5	1.04	2.5
Q17	-0.09	0.011	0.95	-2.7	1.01	0.5
Q18	0.53	0.011	1.20	10.6	1.20	11.0
Q19	0.65	0.011	1.16	8.9	1.18	10.1
Q20	-0.58	0.012	0.83	-10.3	0.87	-6.9
Q21	0.25	0.011	0.88	-7.2	0.92	-5.1
Q22	0.33	0.011	0.96	-2.6	0.99	-0.6
Q23	0.21	0.012	1.06	3.7	1.07	4.2
Q24	-0.10	0.012	1.05	2.7	1.06	3.2
Q25	0.42	0.012	1.05	2.8	1.05	3.1
Q26	-0.39	0.012	1.00	0.1	1.02	1.5
Q28	1.66	0.012	1.11	6.1	1.10	6.0
Q29	-0.30	0.011	1.05	2.6	1.10	5.7
Q30	0.08	0.011	0.83	-10.4	0.80	-12.0
Q31	0.17	0.011	0.73	-17.3	0.73	-17.0
Q32	1.40	0.011	1.11	6.3	1.10	6.2
Q33	0.58	0.011	0.90	-5.8	0.91	-5.3
Q34	-0.56	0.012	0.93	-3.9	0.95	-2.8
Q35	-0.44	-	-	-	-	-

[10] These item numbers refer to the order tested in the spring validation pilot; for reference, Tables 17 and 18 include the full list of items with relevant item numbers.

Table 22. *Item difficulty and fit estimates, grades 6-12*

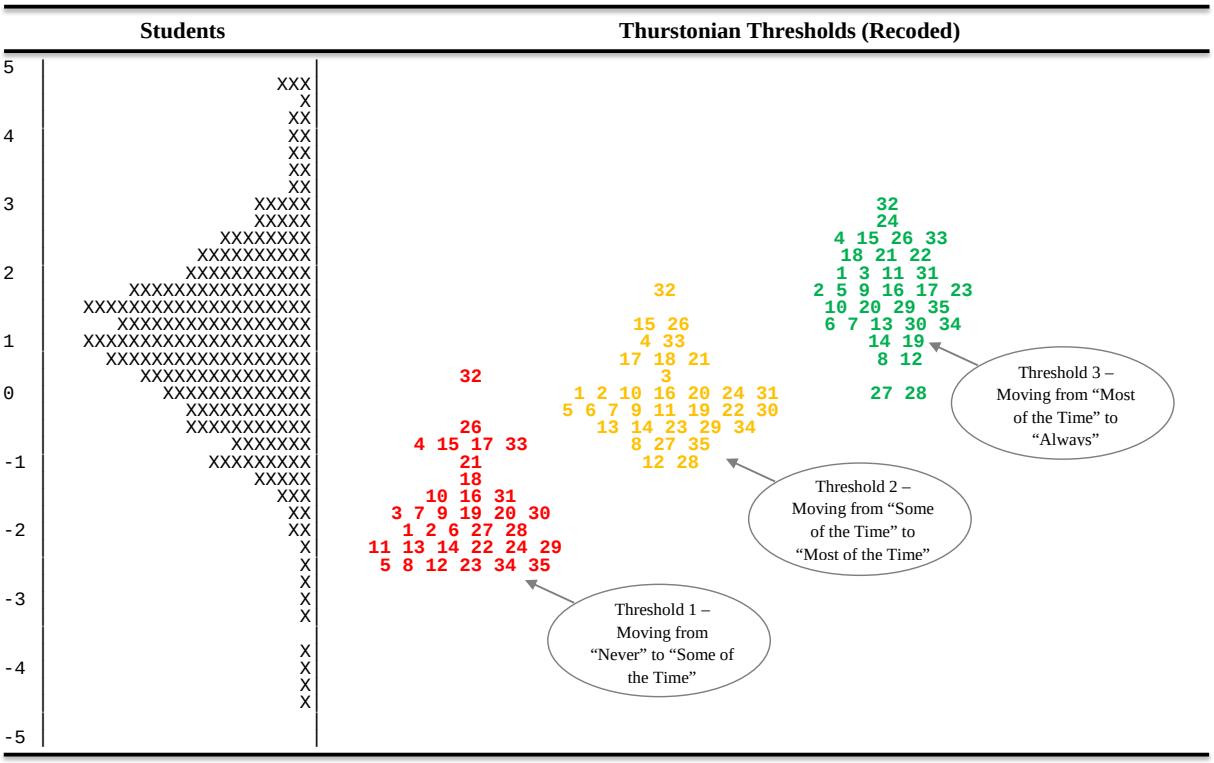
			Outfit		Infit	
Item	Estimate	Error	Mean Square	t	Mean Square	Estimate
Q01	0.087	0.007	0.92	-8.5	0.93	-7.6
Q02	0.007	0.007	1.07	7.4	1.08	8.5
Q03	0.164	0.007	1.01	1.3	1.03	2.7
Q04	0.847	0.007	0.97	-3.0	0.95	-4.9
Q05	-0.301	0.007	1.12	12.2	1.15	15.2
Q06	-0.303	0.007	1.29	27.5	1.25	23.7
Q07	-0.286	0.007	0.94	-6.6	0.97	-2.9
Q08	-0.78	0.008	1.08	8.6	1.11	10.3
Q09	-0.036	0.007	0.97	-2.8	1.01	1.1
Q10	0.012	0.007	1.00	-0.2	1.03	2.8
Q11	-0.073	0.007	0.92	-8.9	0.95	-5.5
Q12	-0.898	0.008	1.13	12.8	1.17	15.6
Q13	-0.444	0.007	1.05	4.7	1.08	7.8
Q14	-0.563	0.007	0.95	-5.4	1.00	-0.2
Q15	0.987	0.007	1.20	19.8	1.19	19.1
Q16	0.123	0.007	1.01	1.1	1.05	4.8
Q17	0.579	0.007	1.33	30.6	1.31	29.9
Q18	0.568	0.007	1.17	16.4	1.18	18.1
Q19	-0.332	0.007	0.90	-10.5	0.94	-6.3
Q20	-0.057	0.007	0.94	-6.4	0.98	-2.2
Q21	0.62	0.007	1.03	2.9	1.05	4.9
Q22	0.01	0.008	1.13	13.0	1.16	15.9
Q23	-0.38	0.008	1.15	15.0	1.18	17.3
Q24	0.151	0.008	1.11	10.7	1.13	13.3
Q26	1.028	0.007	1.45	40.1	1.40	37.7
Q27	-0.834	0.008	1.11	10.5	1.16	13.6
Q28	-0.883	0.008	0.92	-8.0	0.97	-2.5
Q29	-0.367	0.008	1.15	14.7	1.17	16.8
Q30	-0.231	0.007	0.94	-6.2	0.95	-5.4
Q31	0.162	0.007	0.81	-20.9	0.83	-19.0
Q32	1.691	0.007	1.23	22.3	1.21	20.4
Q33	0.843	0.007	0.96	-4.2	0.97	-3.4
Q34	-0.542	0.007	0.86	-14.9	0.87	-13.3
Q35	-0.569	-	-	-	-	-

The horizontal axis represents student overall perceptions ($\hat{\theta}$) regarding their teachers, ranging from most negative to most positive. The vertical axis represents the probability that a student with a given $\hat{\theta}$will select that response option. The four curves represent each of the four response options.

In general, ICCs allow researchers to evaluate the scale and difficulty of items. Each item should have clear categories (e.g., ideally, each of the four response options should be most likely, relative to the other response options, for at least some portion of the $\hat{\theta}$ distribution). Across items, varying difficulty levels are also desirable. Results from Colorado's Student Perception Survey are generally consistent with both of these criteria.

Wright Map. The Wright Map (Figure 8) provides a graphical way to map estimates of student perceptions, $\hat{\theta}$, on the same scale as the item estimates, for all items on the instrument. This is particularly helpful in an analyzing student perceptions and item characteristics in tandem.

Figure 8. *Wright Map of student estimates and response model parameter estimates, grades 6-12*



* Each X represents 103 students, each row is 0.265 logits

The Threshold values at level 2, represented as yellow in the figure, are especially important for Colorado's Student Perception Survey: They represent the place at which students have a 50

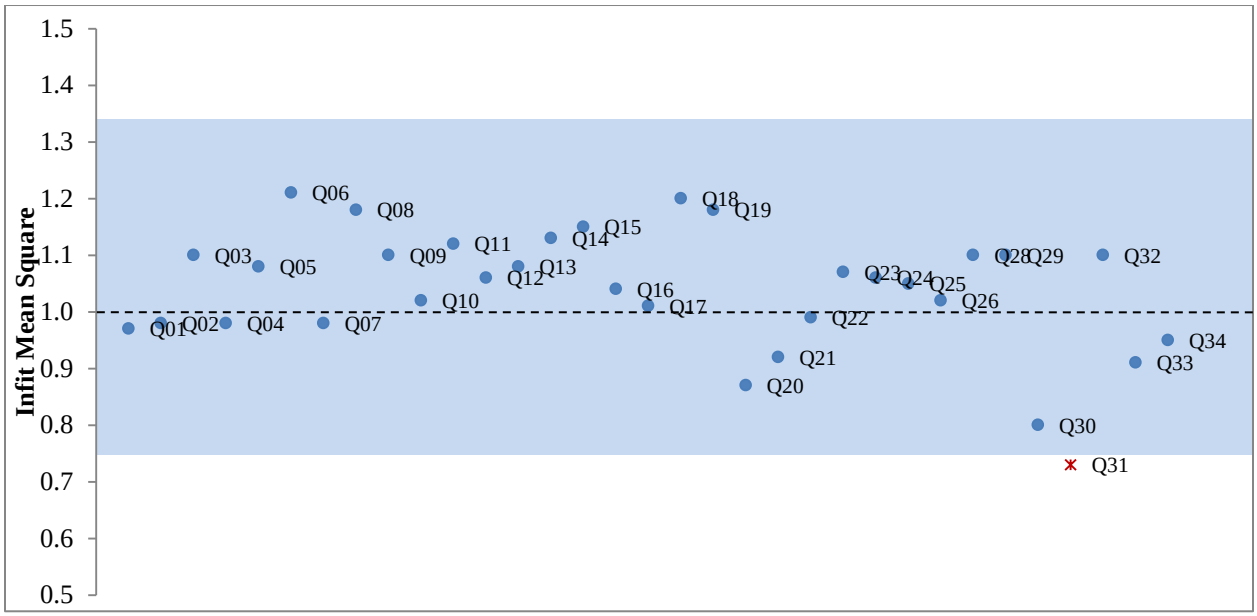
percent probability of responding in the top two categories.[11] The most important thing to note from the Wright Map is that the relative 'difficulties' of the items generally cover the range of student perceptions ($\hat{\theta}$).

Model & Item Fit Analyses

Analyses of model and item fit are useful diagnostic tools when evaluating an instrument, its utility, and the appropriateness of the measurement model. In what follows, mean square fit statistics (estimates of overall item fit) are presented for both grade spans, along with a graphical representation of overall model fit.

Item Fit Analyses. Generally, conservative interpretation guidelines regarding mean square fit statistics indicate that any estimate between 0.75 and 1.33 can be considered reasonably well-fitted (Adams & Khoo, 1996, as cited in Wilson, 2005), and provide evidence of overall model fit. Some researchers (e.g., Linacre, 2006) and states use the less conservative bounds of 0.5 and 1.5.

Figure 9. *Infit mean square item estimates, grades 3-5*



[11] In general, Thurstonian Thresholds provide another way to think about item difficulty, and in that way can be interpreted similarly to the item step difficulties (δ_{ij}) estimated in the model above. Unlike step difficulties, which are empirically derived during model estimation, Thurstonian Thresholds are calculated post hoc. Although their interpretion is quite similar, they also represent different concepts: Step difficulties represent the place at which the curves in the ICC intersect, whereas Thurstonian thresholds represent the $\hat{\theta}$ value where the probability of responding in that category (or higher) reaches 50 percent.

Figure 10. *Outfit mean square item estimates, grades 3-5*

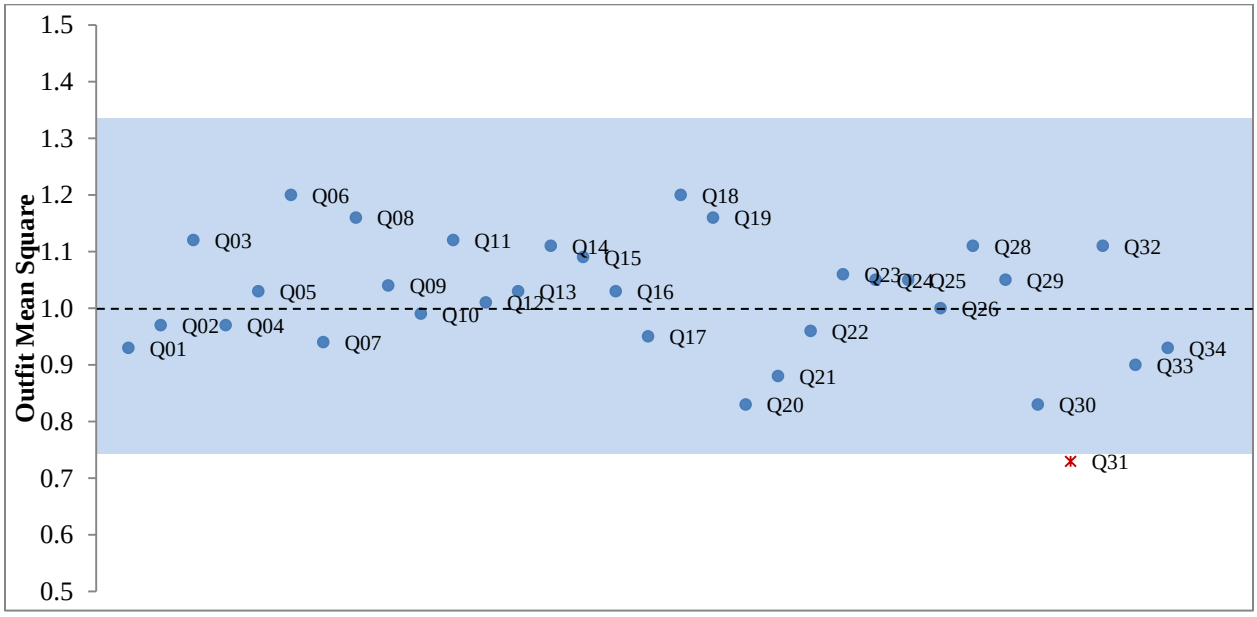


Figure 11. *Infit mean square item estimates, grades 6-12*

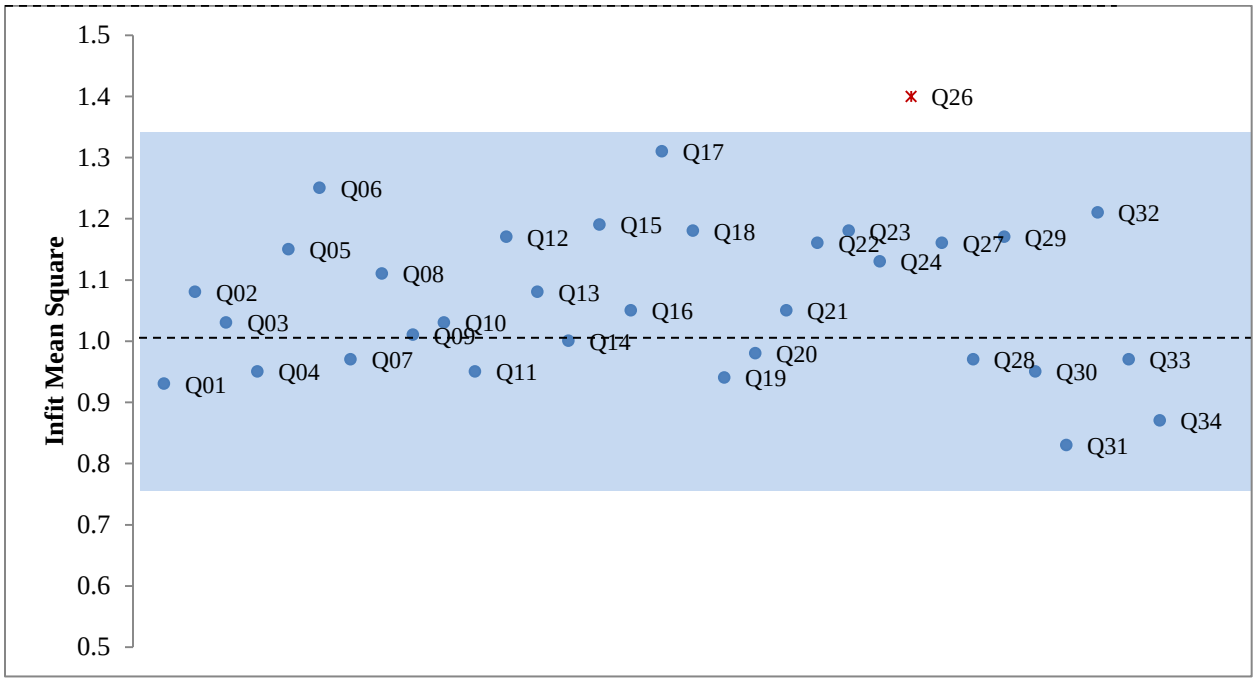
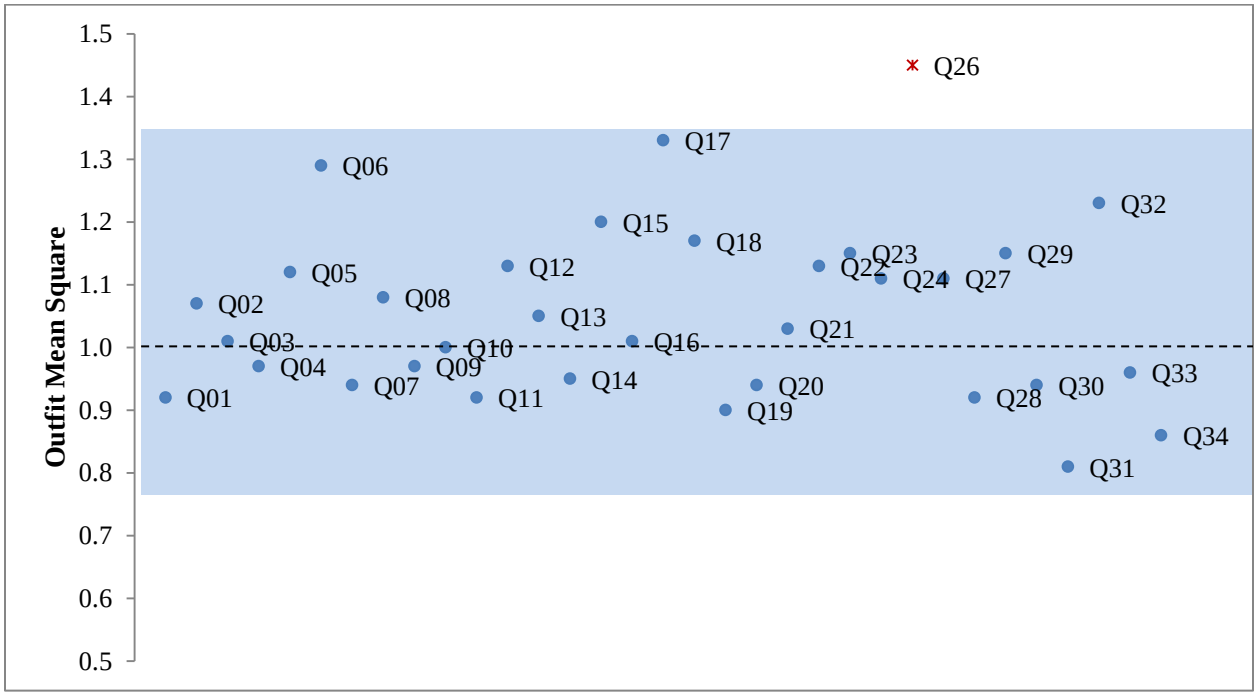


Figure 12. *Outfit mean square item estimates, grades 6-12*



Overall, the items on Colorado's Student Perception Survey demonstrate good fit; all but one item on both the 3-5 and 6-12 instrument fall within the most conservative guidelines; that item does meet the less conservative guidelines advanced by other scholars. Furthermore, items that demonstrate poor or questionable fit should not be immediately discarded; instead, they should be flagged for further examination. Other analyses of item properties suggest that the items on Colorado's Student Perception Survey that do not meet the most conservative fit guidelines ("*I feel like an important part of my classroom community*" for elementary students and "*Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background*" for secondary) are generally good items overall.

Overall Respondent Fit Patterns. Analyses of respondent fit patterns represent visually the relationship between item difficulty and student responses. In Figures 13-14 below, the matrix is purposefully ordered. Rows are sorted in ascending order according to students' raw scores, while columns are sorted left to right by item difficulty, as measured by p-values. Cells in the matrix present respondents' score levels by item, with positive responses in green and negative responses in red.

Figure 13. *Respondent fit patterns by item, grades 3-5*

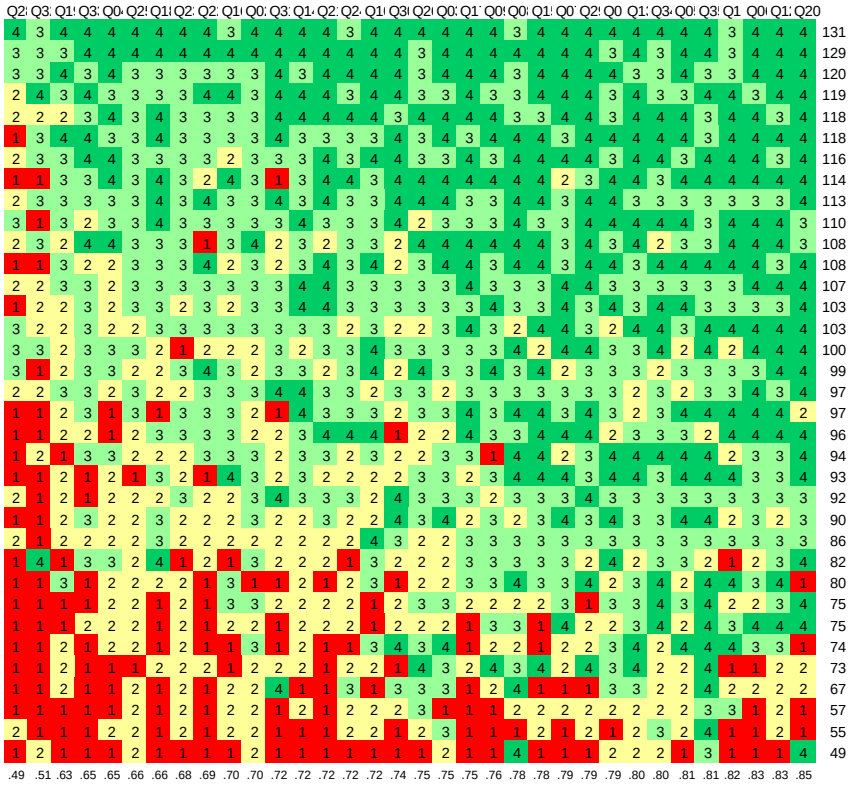
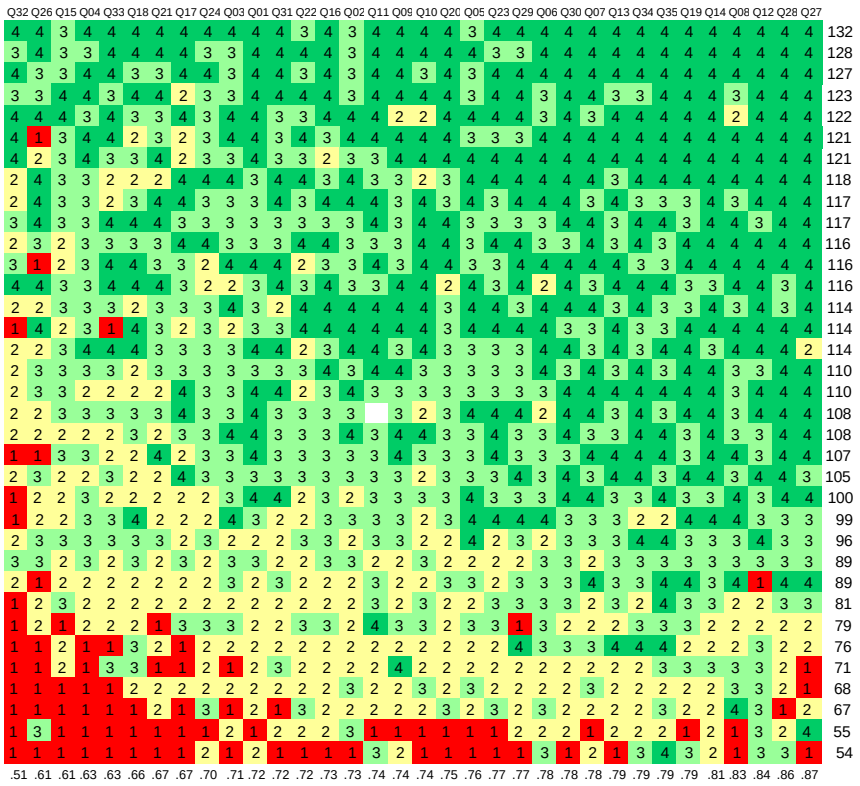


Figure 14. *Respondent fit patterns by item, grades 6-12*



In these matrices, the upper-right hand corner contains mostly green threes and the lower left-hand corner contains mostly red zeros. In general, that suggests students' responses to items of increasing difficulty are dependent upon their overall perceptions of their teachers. Students with more negative perceptions are less likely to respond positively than their peers with more positive perceptions. Furthermore, all students are more likely to respond positively to "easier" items. The scattering of ones and two throughout the table, particularly across the downward sloping diagonal line, indicates that there is a transition between scores of 0 and 3, and between students with more and less positive perceptions.

4.4.Analyses of Dimensionality

Principal components analyses (PCA) were conducted to examine the underlying structure of the survey. This helps to determine whether student responses can be summarized in more easily interpreted groupings, and whether the four elements underlying the Student Perception Survey are statistically defensible. PCA results suggest that a single factor can account for 44.7 percent of the total variance in responses – the strength of this one factor suggests that it is plausible to treat the instrument as one dimension and report overall scores. However, in total, four factors emerged with eigenvalues greater than 1.0 (consistent with the Kaiser (1960) criterion, suggesting that it is also defensible to present results from the Student Perception Survey in four categories.

In general, results from the PCA confirm the underlying theory described above, with four factors that largely map to the four elements underlying the SPS (Student Learning, Student-Centered Environment, Classroom Community, and Classroom Management).

Tables of factor loadings can be found in Tables 23 and 24. Squared factor loadings can be interpreted as the variance explained by each of the four components. For example, in item 1 in Table 23, the first component – which is most closely mapped to **Student Learning** – explains about 0.612^2, or about 37.5 percent of the variance of that item. Each of the other components explains less than 12 percent of the variance.

Table 23. *Factor loadings, grades 3-5*

	Component, Grades 3-5			
	1 (mapped most closely to Student Learning)	**2** (mapped most closely to Student-Centered Env.)	**3** (mapped most closely to Classroom Community)	**4** (mapped most closely to Classroom Management)
1. The schoolwork we do helps me learn.	.612	.080	.345	.144
3. What I learn in this class is useful to me in my real life.	.524	.064	.310	.101
5. In this class, we learn a lot almost every day.	.672	.097	.210	.138

	Component, Grades 3-5			
	1 (mapped most closely to Student Learning)	**2** (mapped most closely to Student-Centered Env.)	**3** (mapped most closely to Classroom Community)	**4** (mapped most closely to Classroom Management)
6. My teacher makes sure that we think hard about things we read and write.	.673	.181	.060	.038
7. When the work is too hard, my teacher helps me keep trying.	.542	.326	.359	.117
8. In this class, it is more important to understand the lesson than to memorize the answers.	.508	.122	.117	.148
9. My teacher uses a lot of different ways to explain things.	.557	.292	.181	.120
10. My teacher knows when we understand the lesson and when we do not.	.473	.385	.229	.160
11. Our classroom materials and supplies have a special place and things are easy to find.	.361	.098	.264	.272
12. In this class, we learn to correct our mistakes.	.633	.174	.178	.140
13. My teacher tells us what we are learning and why.	.537	.231	.186	.170
15. My teacher asks questions to be sure we are following along.	.561	.351	.030	.085
17. My teacher talks to me about my work to help me understand my mistakes.	.545	.406	.270	.138
14. My teacher wants us to share what we think.	.412	.428	.036	.147
18. My teacher writes notes on my work that help me do better next time.	.443	.481	.043	.051
29. My teacher teaches us to respect people's differences.	.421	.396	.201	.142
4. My teacher knows what makes me excited about learning.	.351	.484	.361	.134
19. My teacher talks about things we learn in other classes, subjects, and years.	.380	.520	-.007	.156
21. If I am sad or angry, my teacher helps me feel better.	.343	.496	.471	.127
22. My teacher would notice if something was bothering me.	.309	.532	.403	.101
28. The people we learn and read about in this class are like me.	.089	.575	.150	.190
32. My teacher knows what my life is like outside of school.	.079	.668	.223	.125
33. My teacher knows what is important to me.	.225	.620	.419	.118
16. Students feel comfortable sharing their ideas in this class.	.265	.343	.255	.321
2. The schoolwork we do is interesting.	.331	.193	.393	.217
20. My teacher cares about me.	.411	.337	.524	.115

	Component, Grades 3-5			
	1 (mapped most closely to Student Learning)	**2** (mapped most closely to Student-Centered Env.)	**3** (mapped most closely to Classroom Community)	**4** (mapped most closely to Classroom Management)
30. In this class, I feel like I fit in.	.129	.209	.715	.238
31. I feel like an important part of my classroom community.	.137	.300	.715	.200
34. I ask for help when I need it.	.343	.078	.443	.139
35. I feel like I do a good job in this class.	.179	.108	.678	.167
23. Our class stays busy and does not waste time.	.171	.131	.106	.661
24. Students in my class are respectful to our teacher.	.107	.107	.219	.720
25. My classmates behave the way my teacher wants them to.	.095	.201	.121	.771
26. All of the kids in my class know what they are supposed to be doing and learning.	.227	.125	.220	.628

Table 24. *Factor loadings, grades 6-12*

	Component, Grades 6-12			
	1 (mapped most closely to Student Learning)	**2** (mapped most closely to Student-Centered Env.)	**3** (mapped most closely to Classroom Community)	**4** (mapped most closely to Classroom Management)
1. My teacher makes learning enjoyable.	.630	.370	.295	.169
2. What I learn in this class is useful to me in my real life.	.635	.306	.084	.119
3. My teacher teaches things that are important to me.	.648	.370	.117	.114
4. My teacher knows the things that make me excited about learning.	.608	.494	.204	.116
5. In this class, we learn a lot every day.	.656	.115	.233	.262
6. In this class, it is more important to understand the lesson than to memorize the answers.	.500	.067	.302	.150
7. When the work is too hard, my teacher helps me keep trying.	.595	.225		.094
8. My teacher accepts nothing less than my best effort.	.496	.097	.400	.159
9. My teacher knows when we understand the lesson and when we do not.	.624	.236	.386	.146
10. If I don't understand something, my teacher explains it a different way.	.641	.233	.412	.108
11. My teacher explains difficult things clearly.	.650	.261	.384	.172
15. In this class, we have a say in what we learn and do.	.501	.430	.159	.056

	Component, Grades 6-12			
	1 (mapped most closely to Student Learning)	**2** (mapped most closely to Student-Centered Env.)	**3** (mapped most closely to Classroom Community)	**4** (mapped most closely to Classroom Management)
16. My teacher talks to me about my work to help me understand my mistakes.	.627	.293	.358	.081
17. My teacher writes notes on my work that help me improve.	.527	.272	.202	.082
18. When we study a topic, my teacher makes connections to other subjects or classes.	.581	.352	.142	.074
21. My teacher would notice if something was bothering me.	.441	.528	.368	.042
30. In this class, I feel like I fit in.	.063	.580	.395	.362
31. I feel like an important part of this classroom community.	.214	.647	.366	.278
32. My teacher knows what my life is like outside of school.	.287	.719	.073	.019
33. My teacher knows what is important to me.	.413	.686	.231	.052
34. I ask for help when I need it.	.233	.409	.334	.155
35. I feel like I do a good job in this class.	.225	.473	.304	.193
26. Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural...	.243	.491	.125	.062
19. My teacher cares about me.	.424	.381	.621	.070
20. My teacher pays attention to what all students are thinking and feeling.	.518	.364	.514	.130
12. My classroom is organized and I know where to find what I need.	.386	.047	.499	.252
13. Students feel comfortable sharing their ideas in this class.	.373	.297	.469	.232
14. My teacher respects my opinions and suggestions.	.437	.254	.645	.141
27. My teacher respects my cultural background.	.151	.257	.695	.156
28. My teacher respects me as an individual.	.297	.273	.749	.145
22. Our class stays busy and does not waste time.	.470	.115	.157	.544
23. Students in this class treat the teacher with respect.	.323	.190	.148	.716
24. The students behave the way my teacher wants them to.	.342	.243	.083	.710
25. The way students behave in this class makes it hard to learn.*	-.059	-.085	.084	.661
29. Students in this class respect each other's differences.	.030	.396	.276	.576

Section 5: Reliability, Validity, & Fairness

5.1. Analyses of Reliability & Errors of Measurement

Overview

Student-Level Reliability

Reliability analyses consider the internal structure and consistency of the items in each of the Student Perception Surveys – and for the four component elements. Cronbach's Alpha (α) is a measure of internal consistency, designed to estimate the extent to which the items in an instrument measure a similar construct. Generally, for high-stakes assessments (like TCAP), researchers recommend $\alpha > 0.9$; for other purposes, $\alpha > 0.7$ is considered defensible.

Cronbach's Alpha for both the 3-5 instrument and 6-12 instrument is exceptionally high – 0.94 and 0.96 respectively. Cronbach's Alpha for each of the four elements is also high – ranging from 0.75 to 0.94 – suggesting that each element also meets the generally accepted standards of reliability.

Table 25. *Student-level reliability*

	Reliability (α)	
	Grades 3-5	**Grades 6-12**
Overall Reliability (all items)	0.94	0.96
Student Learning	0.90	0.94
Student-Centered Environment	0.86	0.90
Classroom Community	0.80	0.86
Classroom Management	0.75	0.80

Results for Student Learning Element

Grades 3-5. For the grade 3-5 Student Perception Survey, the **Student Learning** element is comprised of the following 15 items:

- Q1. "*The schoolwork we do helps me learn.*"
- Q2. "*The schoolwork we do is interesting.*"
- Q3. "*What I learn in this class is useful to me in my real life.*"
- Q5. "*In this class, we learn a lot almost every day.*"
- Q6. "*My teacher makes sure that we think hard about things we read and write*".
- Q7. "*When the work is too hard, my teacher helps me keep trying.*"
- Q8. "*In this class, it is more important to understand the lesson than to memorize the answers.*"

- Q9. "*My teacher uses a lot of different ways to explain things.*"
- Q10. "*My teacher knows when we understand the lesson and when we do not.*"
- Q11. "*Our classroom materials and supplies have a special place and things are easy to find.*"
- Q12. "*In this class, we learn to correct our mistakes.*"
- Q13. "*My teacher tells us what we are learning and why.*"
- Q15. "*My teacher asks questions to be sure we are following along.*"
- Q17. "*My teacher talks to me about my work to help me understand my mistakes.*"
- Q18. "*My teacher writes notes on my work that help me do better next time.*"

As is shown in Table 26 and 27 below, these items are all correlated ($r > 0.4$) with the total element score (item-total correlation), and with the other items comprising this element. For ease of interpretation, especially weak inter-item correlations ($r > 0.25$) are marked in Table 27; in general, this occurs quite infrequently, suggesting that the items are related to one another, and are related to the total element subscore.

Table 26. *Grades 3-5 Student Learning item properties*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q1	3.18	.80	.62	.89
Q2	2.78	.84	.49	.89
Q3	3.01	.94	.52	.89
Q5	3.25	.85	.62	.89
Q6	3.34	.93	.57	.89
Q7	3.15	.95	.67	.88
Q8	3.12	.91	.48	.89
Q9	3.05	.92	.60	.89
Q10	2.90	.90	.61	.89
Q11	3.27	.89	.44	.89
Q12	3.35	.85	.61	.89
Q13	3.22	.87	.57	.89
Q15	3.15	.91	.56	.89
Q17	3.02	.96	.68	.88
Q18	2.67	1.12	.52	.89

Table 27. *Student Learning item correlations, grades 3-5*

	Q1	Q2	Q3	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q15	Q17	Q18
Q1	1.00	0.42	0.45	0.52	0.40	0.44	0.34	0.39	0.40	0.31	0.41	0.38	0.37	0.43	0.32
Q2	0.42	1.00	0.33	0.34	0.21	0.39	0.28	0.35	0.37	0.28	0.30	0.34	0.27	0.35	0.22
Q3	0.45	0.33	1.00	0.41	0.34	0.38	0.28	0.36	0.33	0.25	0.34	0.32	0.28	0.37	0.27
Q5	0.52	0.34	0.41	1.00	0.45	0.43	0.34	0.43	0.37	0.28	0.43	0.39	0.39	0.43	0.34
Q6	0.40	0.21	0.34	0.45	1.00	0.41	0.34	0.39	0.35	0.25	0.43	0.35	0.39	0.41	0.39
Q7	0.44	0.39	0.38	0.43	0.41	1.00	0.37	0.45	0.50	0.32	0.46	0.43	0.41	0.57	0.39
Q8	0.34	0.28	0.28	0.34	0.34	0.37	1.00	0.29	0.33	0.25	0.33	0.30	0.29	0.34	0.26
Q9	0.39	0.35	0.36	0.43	0.39	0.45	0.29	1.00	0.43	0.28	0.41	0.41	0.42	0.45	0.34
Q10	0.40	0.37	0.33	0.37	0.35	0.50	0.33	0.43	1.00	0.31	0.41	0.41	0.39	0.49	0.36
Q11	0.31	0.28	0.25	0.28	0.25	0.32	0.25	0.28	0.31	1.00	0.33	0.29	0.25	0.32	0.25
Q12	0.41	0.30	0.34	0.43	0.43	0.46	0.33	0.41	0.41	0.33	1.00	0.38	0.38	0.49	0.36
Q13	0.38	0.34	0.32	0.39	0.35	0.43	0.30	0.41	0.41	0.29	0.38	1.00	0.38	0.44	0.32
Q15	0.37	0.27	0.28	0.39	0.39	0.41	0.29	0.42	0.39	0.25	0.38	0.38	1.00	0.44	0.36
Q17	0.43	0.35	0.37	0.43	0.41	0.57	0.34	0.45	0.49	0.32	0.49	0.44	0.44	1.00	0.46
Q18	0.32	0.22	0.27	0.34	0.39	0.39	0.26	0.34	0.36	0.25	0.36	0.32	0.36	0.46	1.00

* Items with correlations under 0.25 are marked in red

Grades 6-12. For the grade 6-12 Student Perception Survey, the **Student Learning** element is comprised of the following 15 items:

- Q1. "*My teacher makes learning enjoyable*."
- Q2. "*What I learn in this class is useful to me in my real life.*"
- Q3. *"My teacher teaches things that are important to me."*
- Q4. *"My teacher knows the things that make me excited about learning."*
- Q5. *"In this class, we learn a lot every day."*
- Q6. *"In this class, it is more important to understand the lesson than to memorize the answers."*
- Q7. *"When the work is too hard, my teacher helps me keep trying."*
- Q8. *"My teacher accepts nothing less than my best effort."*
- Q9. *"My teacher knows when we understand the lesson and when we do not."*
- Q10. *"If I don't understand something, my teacher explains it a different way."*
- Q11. *"My teacher explains difficult things clearly."*
- Q15. *"In this class, we have a say in what we learn and do."*
- Q16. *"My teacher talks to me about my work to help me understand my mistakes."*
- Q17. *"My teacher writes notes on my work that help me improve."*
- Q18. "*When we study a topic, my teacher makes connections to other subjects or classes.*"

As is shown in Table 28 and 29 below, these items are all correlated ($r > 0.5$) with the total element score (item-total correlation), and with the other items comprising this element. Furthermore, all inter-item correlations are $r > 0.25$, suggesting that the items are related to one another, and are related to the total element subscore.

Table 28. *Student Learning item properties, grades 6-12*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q1	2.87	.91	.77	.93
Q2	2.91	.92	.64	.93
Q3	2.84	.93	.69	.93
Q4	2.53	1.02	.75	.93
Q5	3.03	.88	.66	.93
Q6	3.11	.92	.53	.94
Q7	3.13	.95	.76	.93
Q8	3.31	.84	.58	.93
Q9	2.97	.92	.74	.93
Q10	2.97	.98	.76	.93
Q11	2.94	.87	.77	.93
Q15	2.43	.99	.60	.93
Q16	2.90	.96	.74	.93
Q17	2.69	1.08	.58	.93
Q18	2.64	.96	.63	.93

Table 29. *Student Learning item correlations, grades 6-12*

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q15	Q16	Q17	Q18
Q1	1.00	0.54	0.61	0.71	0.54	0.42	0.62	0.45	0.60	0.61	0.65	0.52	0.57	0.42	0.49
Q2	0.54	1.00	0.67	0.53	0.49	0.38	0.47	0.38	0.47	0.47	0.50	0.41	0.46	0.37	0.42
Q3	0.61	0.67	1.00	0.61	0.51	0.40	0.51	0.40	0.51	0.51	0.54	0.47	0.51	0.38	0.44
Q4	0.71	0.53	0.61	1.00	0.52	0.39	0.59	0.43	0.58	0.58	0.61	0.53	0.57	0.43	0.50
Q5	0.54	0.49	0.51	0.52	1.00	0.42	0.53	0.45	0.52	0.52	0.55	0.36	0.50	0.41	0.43
Q6	0.42	0.38	0.40	0.39	0.42	1.00	0.46	0.37	0.42	0.42	0.43	0.31	0.41	0.32	0.34
Q7	0.62	0.47	0.51	0.59	0.53	0.46	1.00	0.52	0.63	0.67	0.65	0.46	0.65	0.47	0.47
Q8	0.45	0.38	0.40	0.43	0.45	0.37	0.52	1.00	0.49	0.45	0.48	0.33	0.47	0.36	0.36
Q9	0.60	0.47	0.51	0.58	0.52	0.42	0.63	0.49	1.00	0.66	0.65	0.46	0.59	0.44	0.48
Q10	0.61	0.47	0.51	0.58	0.52	0.42	0.67	0.45	0.66	1.00	0.69	0.48	0.61	0.47	0.51
Q11	0.65	0.50	0.54	0.61	0.55	0.43	0.65	0.48	0.65	0.69	1.00	0.49	0.61	0.46	0.51

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q15	Q16	Q17	Q18
Q15	0.52	0.41	0.47	0.53	0.36	0.31	0.46	0.33	0.46	0.48	0.49	1.00	0.49	0.38	0.46
Q16	0.57	0.46	0.51	0.57	0.50	0.41	0.65	0.47	0.59	0.61	0.61	0.49	1.00	0.56	0.50
Q17	0.42	0.37	0.38	0.43	0.41	0.32	0.47	0.36	0.44	0.47	0.46	0.38	0.56	1.00	0.48
Q18	0.49	0.42	0.44	0.50	0.43	0.34	0.47	0.36	0.48	0.51	0.51	0.46	0.50	0.48	1.00

Results for Student-Centered Environment Element

Grades 3-5. For the grade 3-5 Student Perception Survey, the **Student-Centered Environment** element is comprised of the following 10 items:

- Q4. *"My teacher knows what makes me excited about learning."*
- Q14. *"My teacher wants us to share what we think."*
- Q16. *"Students feel comfortable sharing their ideas in this class."*
- Q19. *"My teacher talks about things we learn in other classes, subjects, and years."*
- Q21. *"If I am sad or angry, my teacher helps me feel better."*
- Q22. *"My teacher would notice if something was bothering me."*
- Q28. *"The people we learn and read about in this class are like me."*
- Q29. *"My teacher teaches us to respect people's differences."*
- Q32. *"My teacher knows what my life is like outside of school."*
- Q33. *"My teacher knows what is important to me."*

As is shown in Table 30 and 31 below, these items are all correlated ($r > 0.4$) with the total element score (item-total correlation), and with the other items comprising this element. Furthermore, all inter-item correlations are $r > 0.25$, suggesting that the items are related to one another, and are related to the total element subscore.

Table 30. *Student-Centered Environment item properties, grades 3-5*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q4	2.62	.97	.63	.84
Q14	2.88	.91	.48	.85
Q16	2.79	.85	.49	.85
Q19	2.51	.96	.50	.85
Q21	2.88	1.12	.67	.84
Q22	2.78	1.01	.65	.84
Q28	1.94	.84	.47	.85
Q29	3.16	.97	.53	.85
Q32	2.07	.99	.55	.85
Q33	2.61	1.05	.69	.84

Table 31. *Student-Centered Environment item correlations – grades 3-5*

	Q4	Q14	Q16	Q19	Q21	Q22	Q28	Q29	Q32	Q33
Q4	1.00	0.33	0.37	0.36	0.50	0.49	0.36	0.36	0.42	0.53
Q14	0.33	1.00	0.37	0.33	0.36	0.34	0.25	0.34	0.29	0.34
Q16	0.37	0.37	1.00	0.30	0.37	0.36	0.28	0.30	0.30	0.35
Q19	0.36	0.33	0.30	1.00	0.36	0.37	0.31	0.33	0.32	0.36
Q21	0.50	0.36	0.37	0.36	1.00	0.62	0.32	0.45	0.41	0.56
Q22	0.49	0.34	0.36	0.37	0.62	1.00	0.32	0.41	0.39	0.55
Q28	0.36	0.25	0.28	0.31	0.32	0.32	1.00	0.28	0.35	0.37
Q29	0.36	0.34	0.30	0.33	0.45	0.41	0.28	1.00	0.30	0.42
Q32	0.42	0.29	0.30	0.32	0.41	0.39	0.35	0.30	1.00	0.54
Q33	0.53	0.34	0.35	0.36	0.56	0.55	0.37	0.42	0.54	1.00

Grades 6-12. For the grade 6-12 Student Perception Survey, the **Student-Centered Environment** element is comprised of the following 7 items:

- Q12. *"My classroom is organized and I know where to find what I need."*
- Q13. *"Students feel comfortable sharing their ideas in this class."*
- Q14. *"My teacher respects my opinions and suggestions."*
- Q19. *"My teacher cares about me."*
- Q20. *"My teacher pays attention to what all students are thinking and feeling."*
- Q27. *"My teacher respects my cultural background."*
- Q28. *"My teacher respects me as an individual."*

As is shown in Table 32 and 33 below, these items are all correlated ($r > 0.5$) with the total element score (item-total correlation), and with the other items comprising this element. Furthermore, all inter-item correlations are $r > 0.25$, suggesting that the items are related to one another, and are related to the total element subscore.

Table 32. *Student-Centered Environment item properties, grades 6-12*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q12	3.38	.82	.56	.90
Q13	3.15	.88	.64	.89
Q14	3.26	.90	.78	.88
Q19	3.18	.96	.78	.88
Q20	3.00	.95	.75	.88
Q27	3.47	.87	.64	.89
Q28	3.46	.85	.78	.88

Table 33. *Student-Centered Environment item correlations, grades 6-12*

	Q12	Q13	Q14	Q19	Q20	Q27	Q28
Q12	1.00	0.46	0.49	0.46	0.47	0.41	0.47
Q13	0.46	1.00	0.59	0.52	0.56	0.42	0.51
Q14	0.49	0.59	1.00	0.68	0.67	0.55	0.69
Q19	0.46	0.52	0.68	1.00	0.72	0.55	0.71
Q20	0.47	0.56	0.67	0.72	1.00	0.50	0.63
Q27	0.41	0.42	0.55	0.55	0.50	1.00	0.67
Q28	0.47	0.51	0.69	0.71	0.63	0.67	1.00

Results for Classroom Community Element

Grades 3-5. For the grade 3-5 Student Perception Survey, the **Classroom Community** element is comprised of the following 5 items:

- Q20. *"My teacher cares about me."*
- Q30. *"In this class, I feel like I fit in."*
- Q31. *"I feel like an important part of my classroom community."*
- Q34. *"I ask for help when I need it."*
- Q35. *"I feel like I do a good job in this class."*

As is shown in Table 34 and 35 below, these items are all correlated ($r > 0.4$) with the total element score (item-total correlation), and with the other items comprising this element. Furthermore, all inter-item correlations are $r > 0.25$, suggesting that the items are related to one another, and are related to the total element subscore.

Table 34. *Classroom Community item properties, grades 3-5*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q20	3.39	.94	.57	.77
Q30	2.98	1.04	.66	.74
Q31	2.87	.99	.68	.73
Q34	3.22	.87	.44	.81
Q35	3.23	.86	.59	.77

Table 35. *Classroom Community item correlations, grades 3-5*

	Q20	Q30	Q31	Q34	Q35
Q20	1.00	0.33	0.37	0.36	0.50
Q30	0.33	1.00	0.37	0.33	0.36
Q31	0.37	0.37	1.00	0.30	0.37
Q34	0.36	0.33	0.30	1.00	0.36
Q35	0.50	0.36	0.37	0.36	1.00

Grades 6-12. For the grade 6-12 Student Perception Survey, the **Classroom Community** element is comprised of the following 8 items:

- Q21. *"My teacher would notice if something was bothering me."*
- Q26. *"Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background."*
- Q30. *"In this class, I feel like I fit in."*
- Q31. *"I feel like an important part of this classroom community."*
- Q32. *"My teacher knows what my life is like outside of school."*
- Q33. *"My teacher knows what is important to me."*
- Q34. *"I ask for help when I need it."*
- Q35. *"I feel like I do a good job in this class."*

As is shown in Table 36 and 37 below, these items are all correlated ($r > 0.4$) with the total element score (item-total correlation), and with the other items comprising this element. For ease of interpretation, especially weak inter-item correlations ($r > 0.25$) are marked in Table 37; in general, this occurs quite infrequently, suggesting that the items are related to one another, and are related to the total element subscore.

Table 36. *Classroom Community item properties, grades 6-12*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q21	2.66	1.03	.67	.84
Q26	2.43	1.05	.44	.87
Q30	3.11	.95	.62	.84
Q31	2.87	.95	.73	.83
Q32	2.05	1.02	.62	.84
Q33	2.53	1.02	.74	.83
Q34	3.18	.88	.53	.85
Q35	3.17	.82	.56	.85

Table 37. *Classroom Community item correlations, grades 6-12*

	Q21	Q26	Q30	Q31	Q32	Q33	Q34	Q35
Q21	1.00	0.36	0.44	0.53	0.55	0.65	0.41	0.39
Q26	0.36	1.00	0.29	0.37	0.36	0.39	0.24	0.26
Q30	0.44	0.29	1.00	0.70	0.38	0.46	0.40	0.47
Q31	0.53	0.37	0.70	1.00	0.47	0.58	0.44	0.52
Q32	0.55	0.36	0.38	0.47	1.00	0.68	0.32	0.32
Q33	0.65	0.39	0.46	0.58	0.68	1.00	0.42	0.43
Q34	0.41	0.24	0.40	0.44	0.32	0.42	1.00	0.50
Q35	0.39	0.26	0.47	0.52	0.32	0.43	0.50	1.00

* Items with correlations under 0.25 are marked in red

Results for Classroom Management Element

Grade 3-5. For the grade 3-5 Student Perception Survey, the **Classroom Management** element is comprised of the following 4 items:

- Q23. *"Our class stays busy and does not waste time."*
- Q24. *"Students in my class are respectful to our teacher."*
- Q25. *"My classmates behave the way my teacher wants them to."*
- Q26. *"All of the kids in my class know what they are supposed to be doing and learning."*

As is shown in Table 38 and 39 below, these items are all correlated ($r > 0.4$) with the total element score (item-total correlation), and with the other items comprising this element. Furthermore, all inter-item correlations are $r > 0.25$, suggesting that the items are related to one another, and are related to the total element subscore.

Table 38. *Classroom Management item properties, grades 3-5*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q23	2.74	.83	.48	.73
Q24	2.89	.83	.57	.67
Q25	2.64	.77	.61	.65
Q26	2.99	.78	.52	.70

Table 39. *Classroom Management item correlations, grades 3-5*

	Q23	Q30	Q31	Q34
Q23	1.00	0.38	0.41	0.36
Q24	0.38	1.00	0.53	0.42
Q25	0.41	0.53	1.00	0.46
Q26	0.36	0.42	0.46	1.00

***Grade 6-12*.** For the grade 6-12 Student Perception Survey, the **Classroom Management** element is comprised of the following 4 items:

- Q22. *"Our class stays busy and does not waste time."*
- Q23. *"Students in this class treat the teacher with respect."*
- Q24. *"The students behave the way my teacher wants them to."*
- Q29. *"Students in this class respect each other's differences."*

As is shown in Table 40 and 41 below, these items are all correlated ($r > 0.5$) with the total element score (item-total correlation), and with the other items comprising this element. Furthermore, all inter-item correlations are $\underline{r} > 0.25$, suggesting that the items are related to one another, and are related to the total element subscore.

Table 40. *Grades 6-12 Classroom Management item properties*

	Mean	SD	Item-Total Correlation	α if Item Deleted
Q22	2.88	.83	.56	.78
Q23	3.06	.83	.69	.71
Q24	2.81	.78	.71	.71
Q29	3.09	.85	.51	.80

Table 41. *Classroom Management item correlations, grades 6-12*

	Q22	Q23	Q24	Q29
Q22	1.00	0.51	0.54	0.36
Q23	0.51	1.00	0.69	0.47
Q24	0.54	0.69	1.00	0.46
Q29	0.36	0.47	0.46	1.00

Teacher-Level Reliability

Because the unit of analysis for the Student Perception Survey is actually the teacher, not the student, it is also important to consider teacher-level reliabilities. As such, we calculated Cronbach's α (overall and for the four elements) using teacher mean scores on each item. In general, teacher-level reliability analyses yield results quite similar to student-level analyses. The estimates of reliability are slightly higher at the teacher-level, which suggests that the aggregate values for teachers may eliminate some error and/or variation. Furthermore, for each item, we see slightly higher item-total correlations at the teacher-level, and strong inter-item correlations (both for the overall instrument and for each element). Because they are duplicative of the results outline in the section above, full results are not presented here but can instead be found in Appendices E and F.

Table 42. *Teacher-level reliability*

	Reliability (α)	
	Grades 3-5	**Grades 6-12**
Overall Reliability (all items)	0.97	0.98
Student Learning	0.95	0.97
Students-Centered Environment	0.94	0.96
Classroom Community	0.90	0.94
Classroom Management	0.90	0.91

Standard Error of Measurement

The constant Standard Error of Measurement (SEM) provides another way to estimate the accuracy and reliability of the Student Perception Survey. The SEM is calculated using the following equation:

$$SEM = \sigma_x\sqrt{1 - \hat{\alpha}}$$

where α is the calculated Cronbach's α, and σ_x is the standard deviation. The SEM can be interpreted as an estimate of measurement error associated with the survey, and represents the amount we might expect a given teachers' score to vary due to measurement error. In general, smaller SEMs are associated with less variability and suggest that the survey produces more precise scores than surveys with larger SEMs.

The overall estimates of SEM for each instrument – for teacher mean score and for percent favorable – are quite small, and are presented in Table 43.

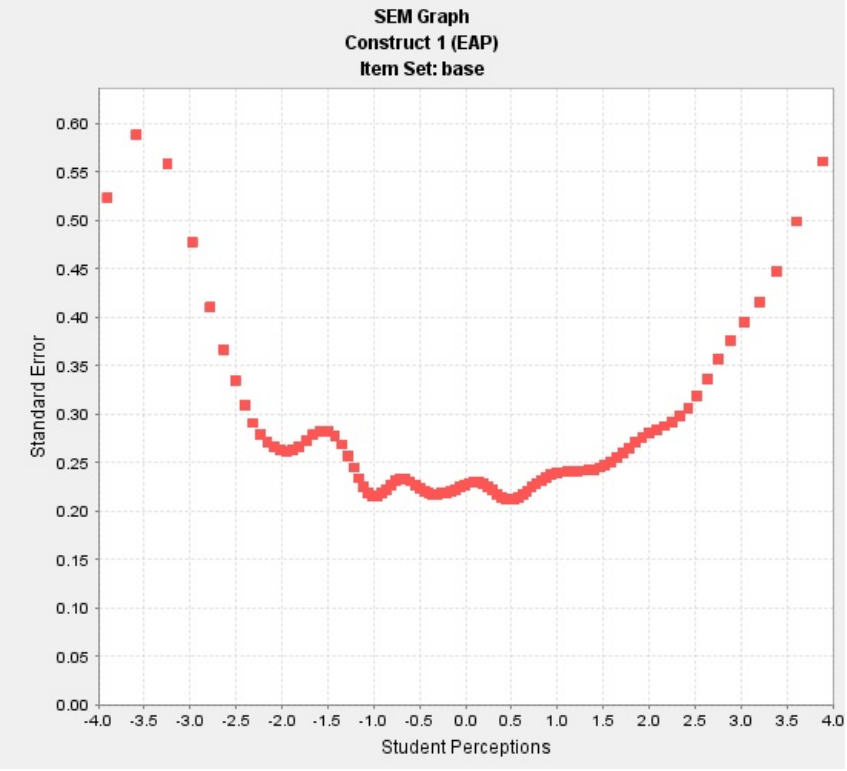
Table 43. S*tandard error of measurement (SEM) for overall teacher-level mean and percent favorable scores*

Standard Error of Measurement (SEM)	Grades 3-5	Grades 6-12
Teacher Mean Score	0.04	0.04
Teacher Percent Favorable	1.6%	1.9%

The SEM calculations are also useful in a practical sense: They allow us to calculate confidence intervals (±1.96 SEM) around teacher-level scores. For example, if a teacher's observed mean score on the survey was 3.0, we can be 95 percent confident that teacher's true score lies between 2.92 and 3.08 [or 3.0 ± (0.04*1.96)]. Similarly, if a grade 4 teacher's percent favorable observed score was 75%, we can be 95 percent confident that teacher's true score lies between 71.9% and 78.1%.

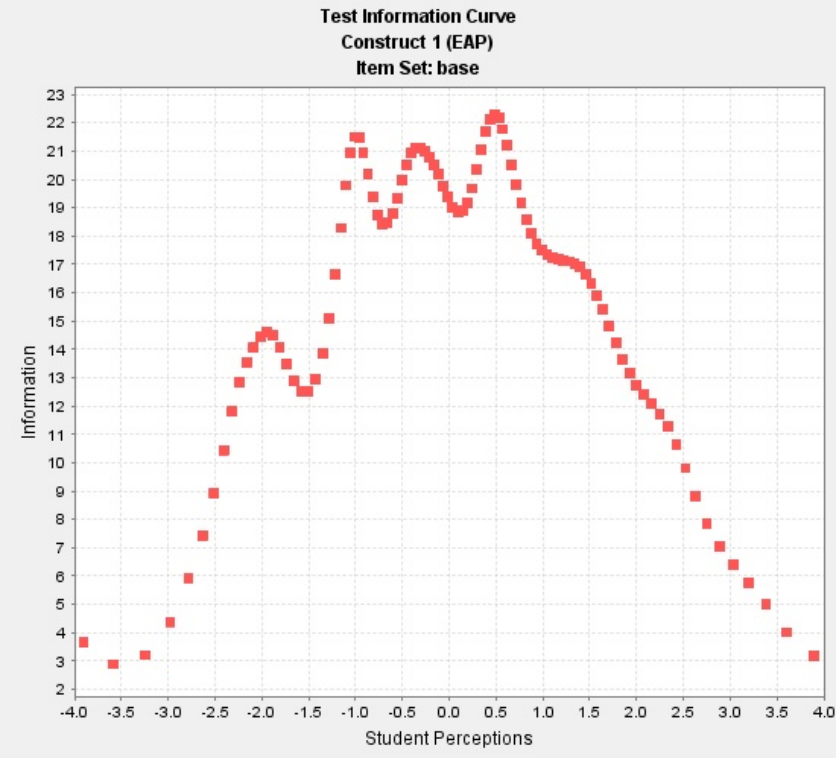
Standard error of measurement associated with student perceptions. Having modeled the Student Perception Survey data via the partial credit model, we can also estimate the level of error associated with estimating student perceptions ($\hat{\theta}$). These estimates vary based on student's estimated $\hat{\theta}$, with larger SEMs – and therefore less precise estimates – for students on the very top and bottom of the distribution (although this represents a small portion of the student population). As Figure 15 shows, the SEM for $\hat{\theta}$ range from roughly 0.2 to 0.6 logits.

Figure 15. *Standard error of measurement, grades 6-12*



Furthermore, the test information curve presents the range over $\hat{\theta}$ where the instrument can provide the most information, and best discriminate among respondents. Generally, higher information indicates more precision. The function is peaked toward the middle of the scale (where the vast majority of students lie). This means that the instruments provide the most information for students with perceptions centered around average, and that it provides less precise information for students with extremely negative or extremely positive perceptions.

Figure 16. *Test information curve, grades 6-12*



5.2. Validity Analyses

Correlations with Other Measures

To assess the extent to which Colorado's Student Perception Survey is measuring some underlying aspect of teacher effectiveness, we examined the relationship between scores on the SPS and (a) ratings from the state model evaluation system, (b) measures of student growth and achievement, and (c) student demographics. For both instruments, we find correlations that are generally consistent with – or even stronger than – the correlations observed in the MET study.[12]

[12] To contextualize these results, we remind readers that the MET study found moderate correlations between student survey results and teacher value-added estimates from state mathematics and reading assessments (ranging from 0.07 to 0.22 for the total student survey score, and from 0.03 to 0.24 on the various student survey subscores).

State model evaluation system ratings. To assess the extent to which the results from the Student Perception Survey are related to results from the teacher evaluation system, we tested the correlations between overall teacher-level survey results (overall and for the four elements) and teacher Performance Ratings on the Colorado State Model Evaluation System.[13] In 2012-13, teachers across 27 pilot districts participated in a pilot of the state model system and, as part of that pilot, were rated on the five Teacher Quality Standards that measure professional practice:

- Standard I. Teachers demonstrate mastery of and pedagogical expertise in the content they teach;
- Standard II. Teachers establish a safe, inclusive and respectful learning environment for a diverse population of students;
- Standard III. Teachers plan and deliver effective instruction and create an environment that facilitates learning for their students;
- Standard IV. Teachers reflect on their practice; and
- Standard V. Teachers demonstrate leadership.

Performance on these Quality Standards was measured using the state-developed rubric that identifies the practices necessary to achieve the standards. From these rubrics, teachers are assigned to one of five performance rating levels for each Standard, which are then aggregated to assign an overall performance rating.

In general, there is a significant positive correlation between Colorado's Student Perception Survey and teacher performance ratings on the state model evaluation system. Tables 44 and 45 present results of the correlational analyses for both instruments.

Table 44. *Correlations of the SPS and teacher ratings on the state model evaluation system, grades 3-5*

	Standard I Rating	Standard II Rating	Standard III Rating	Standard IV Rating	Standard V Rating	Overall Rating	Summative Score
Overall Percent Favorable	.269***	.337**	.249***	.380***	.190***	.312***	.357***
Student Learning	.233***	.288***	.203**	.319***	.154*	.253***	.300***
Student-Centered Envir.	.232***	.309***	.205**	.360***	.152*	.284***	.317***
Classroom Community	.268***	.334***	.287***	.374***	.224***	.324***	.369***
Classroom Management	.325***	.352***	.320***	.383***	.247***	.362***	.402***

* $p < 0.05$ (one-tailed) ** $p < 0.01$ (one-tailed) *** $p < 0.001$ (one-tailed)

[13] For more information about the Colorado State Model Evaluation System, visit the official Colorado Department of Education Educator Effectiveness website.

Table 45. *Correlations of the SPS and teacher ratings on the state model evaluation system, grades 6-12*

	Standard I Rating	Standard II Rating	Standard III Rating	Standard IV Rating	Standard V Rating	Overall Rating	Summative Score
Overall Score	.171***	.203***	.089*	.098*	.120**	.144***	.161***
Student Learning	.182***	.211***	.112**	.121**	.123**	.159***	.177***
Student-Centered Envir.	.132**	.168***	.055	.067	.091*	.108**	.122**
Classroom Community	.128**	.141***	.014	.038	.083*	.081*	.097*
Classroom Management	.186***	.253***	.148***	.116**	.172***	.194***	.207***

* $p < 0.05$ (one-tailed) ** $p < 0.01$ (one-tailed) *** $p < 0.001$ (one-tailed)

In reviewing these results, several interesting themes emerge. First, the grade 3-5 instrument is generally more strongly correlated to both teacher evaluation ratings (and to subsequent analyses of measures of student growth and achievement). This indicates that the ratings of elementary students are more positively associated with those from principals/evaluators and with overall student growth and achievement than ratings from students in grades 6-12. This is an interesting finding in that it seems to contradict fears expressed by many teachers and some scholars that student surveys may be developmentally inappropriate for elementary students.

Second, overall student survey results seem to be most strongly correlated with Teacher Quality Standard #2 (which loosely measures classroom climate) and Standard #4 (which focuses on teacher reflection). Furthermore, among the four elements on the Student Perception Survey, the **Classroom Management** element is consistently more positively correlated to teacher evaluation ratings than any other element (and often more than the overall score).

Measures of student learning. To assess the extent to which the results from the Student Perception Survey are related to measures of student learning, we tested the correlations between overall teacher-level survey results (overall and for the four elements) and a measure of student growth (teacher median growth percentiles)[14] and achievement (percent of students proficient or advanced within a given teachers' course load). Tables 46 and 47 present results of the correlational analyses for both instruments.

Overall, Colorado's Student Perception Survey results are more strongly correlated to teacher evaluation ratings than to measures of student growth, although significant correlations exist

[14] The teacher median growth percentile summarizes student growth rates across all students taught by a teacher, within a given subject area. It is calculated by taking the *median* student growth percentile, across all the students assigned to a given teacher. The median growth percentile tells us how much growth a teacher's students make in a year, in comparison with other teachers. For example, a teacher median growth percentile of 70 indicates that the median student in that teachers' class grew as well or better than 70 percent of her academic peers; another way to say this is that half of the class had student growth percentiles above 70.

with both. With regard to measures of student growth, we see slightly stronger correlations to reading and writing than to mathematics, although the differences are marginal. This represents a departure from results from prior research, most notably the MET study which found stronger correlations between student surveys and mathematics than English language arts. Finally, although both growth and achievement measures are positively related to the student survey results, we find slightly stronger correlations with the measures of student achievement than the measures of student growth, across subject areas. This suggests that teachers of students who demonstrate high growth and high achievement are also slightly more likely to garner positive perceptions from their students on the Student Perception Survey.

Table 46. *Correlations of the SPS and measures of student growth and achievement, grades 3-5*

	Median Growth Percentiles			Percent of Students Proficient or Advanced		
	TCAP Math	TCAP Reading	TCAP Writing	TCAP Math	TCAP Reading	TCAP Writing
Overall Percent Favorable	.198**	.227**	.177*	.229**	.148*	.258***
Student Learning	.236**	.206**	.161*	.167*	.071	.178
Student-Centered Environ.	.117	.164*	.132*	.169*	.103	.214**
Classroom Community	.127*	.230**	.157*	.288***	.201**	.277***
Classroom Management	.262***	.324***	.273***	.378***	.356***	.453***

* p < 0.05 (one-tailed) ** p < 0.01 (one-tailed) *** p < 0.001 (one-tailed)

Table 47. *Correlations of the SPS and measures of student growth and achievement, grades 6-12*

	Median Growth Percentiles			Percent of Students Proficient or Advanced		
	TCAP Math	TCAP Reading	TCAP Writing	TCAP Math	TCAP Reading	TCAP Writing
Overall Percent Favorable	.064	.134*	.165*	.123	.079	.074
Student Learning	.032	.105	.142*	.063	.019	.015
Student-Centered Environ.	.097	.127*	.172**	.129	.108	.092
Classroom Community	-.011	.126*	.137*	.099	.060	.070
Classroom Management	.274***	.232***	.243***	.381***	.311***	.294***

* p < 0.05 (one-tailed) ** p < 0.01 (one-tailed) *** p < 0.001 (one-tailed)

Student demographics. The correlation between the Student Perception Survey Results and student demographics within a given teacher's course load was also tested. In general, there is no strong evidence that results are significantly related to student demographics. For grades 6-12, there is a weak positive, but statistically significant, relationship between the percent of ELL students enrolled in a teachers' course and the overall results from the student survey; this means

that teachers with higher proportions of ELL students are slightly more likely to perform well on the student survey. Furthermore, for elementary teachers, there is weak negative relationship between the percentage of students of color and teachers' scores on the **Classroom Management** element.

Table 48. *Correlations of the SPS and classroom demographics, grades 3-5*

	Percent Minority Students	Percent of ELL Students
Overall Score	.022	.066
Student Learning	.092	.095
Student-Centered Environment	.041	.068
Classroom Community	-.076	.021
Classroom Management	-.155*	-.027

* p < 0.05 (one-tailed) ** p < 0.01 (one-tailed) *** p < 0.001 (one-tailed)

Table 49. *Correlations of the SPS and classroom demographics, grades 6-12*

	Percent Minority Students	Percent of ELL Students
Overall Score	-.033	.094*
Student Learning Element	.027	.103*
Student-Centered Environment	-.077	.062
Classroom Community	-.044	.123*
Classroom Management	-.189*	-.009

* p < 0.05 (one-tailed) ** p < 0.01 (one-tailed) *** p < 0.001 (one-tailed)

Analyses of Open-Ended Responses[15]

The spring validation pilot of Colorado's Student Perception Survey also included space for students to respond freely to the question, *"Do you have any other thoughts or feedback for your teacher?"* To evaluate the open question about the appropriateness of asking students to assess their teachers' instructional behaviors, we conducted in-depth qualitative analyses of the open-ended responses. In particular, we were interested in whether students took the survey seriously and in what ways responses were substantive in nature. This analysis serves as a critical tool to address teacher and district concerns about the Student Perception Survey as well as questions emerging as to how student feedback will be used in conjunction with other measures of teacher effectiveness.

[15] For the full analyses of the open-ended responses, please visit the CEI Student Perception Survey website.

Methods. Amongst the almost 30,000 student responses collected via the Student Perception Survey in the spring pilot, there were 14,539 responses to the question *"Do you have any other thoughts or feedback for your teacher?"*[16] These responses were then coded as substantive (meaning that the student feedback was on-topic and provided at least a general statement about their teacher and/or classroom environment) or off-topic (meaning that students provided written feedback that did not address the question). The substantive responses were then coded as actionable or not, with an actionable response considered feedback specific enough for teachers to take action to alter or maintain their current classroom practices. For instance, the response, "My teacher is great!" would be deemed substantive but not actionable, whereas "My teacher is great because he provides us with multiple ways of understanding the material" would be considered actionable.

Next, actionable responses were categorized thematically to get a sense of the overall trends that existed across student feedback. This step enabled high-level takeaways concerning the nature of the actionable student responses in order to provide an overall sense of the feedback students had for their teachers.

Key Findings. In analyzing the 14,539 open-ended responses from both elementary and secondary students, we find that not only were the majority of students taking the survey seriously, but that many of the responses were specific and actionable in nature (i.e., teachers could likely alter or maintain their practices given student feedback).

Of the 14,539 student responses, 98.6 percent (N=14,341) were considered *substantive* (meaning they were on-topic and provided general information about what the student thought about their teacher). This finding was consistent across grades and subject areas (see Tables 50 and 51).

Table 50. *Percent of open-ended responses coded as substantive or actionable, by subject area*

Subject Area	Percent Substantive	Percent Actionable
Art	98.8%	55.6%
Elementary (Homeroom)	99.1%	60.2%
Language Arts	98.4%	68.8%
Mathematics	98.5%	72.7%
Music	98.1%	68.6%
Physical Education & Health	98.0%	59.4%
Science	98.9%	72.3%
Social Studies	98.6%	70.1%
World Languages	98.6%	74.9%

[16] Not included in this number are the approximately 1,500 students that wrote "No" or the equivalent thereof to the open-ended question. These responses were placed in the "non-response" group.

Table 51. *Percent of open-ended responses coded as substantive or actionable, by grade*

Grade	Percent Substantive	Percent Actionable
Grade 3	98.6%	45.8%
Grade 4	98.6%	60.3%
Grade 5	99.5%	69.7%
Grade 6	99.4%	69.6%
Grade 7	98.7%	68.1%
Grade 8	97.3%	69.2%
Grade 9	98.0%	66.9%
Grade 10	98.6%	70.0%
Grade 11	98.8%	73.9%
Grade 12	98.6%	70.7%

An additional 66.3 percent of these responses (N=9,646) were coded as *actionable* (meaning that they contained specific feedback that could likely be acted upon by teachers). Moreover, although some subjects and grades were slightly more likely to garner actionable feedback (e.g., students in higher grades and in core academic subjects), in general actionable responses came from students in **all** grades (3-12) and from a variety of different types of classrooms (including from music, art and physical education classes).

In general, the actionable responses from students can be categorized into the following 11 themes, described below in greater detail:

- **Help For Understanding** – Students referred to the additional help they received from their teachers when they needed to be more successful in school.
- **Personal Relationships** – Students referred to the ways their teachers connected with them and got to know them beyond the classroom context.
- **Care** – Students referred to the care they felt from their teachers.
- **Content Knowledge** – Students referred to their teacher's knowledge of the subject or discipline.
- **Preparation For The Future** – Students referred to their teachers preparing them for future endeavors, including more advanced classes/grades, college, jobs and assessments.
- **Instruction** – Students referred to the instruction they experienced or did not experience in their classroom. These responses include suggestions for future instructional practices and comments about the nature of this instruction.
- **Classroom Management** – Students referred to their teacher's ability and willingness to manage student behavior in the classroom.
- **Respect** – Students referred to the respect they had for their teacher and/or the respect they got from their teacher.
- **Grading** – Students referred to the grading policies/practices in their classroom.

- **Justice/Fairness** – Students referred to fairness in their classroom, particularly in terms of how they or others are treated in relation to other students or groups.
- **Student Voice And Choice** – Students referred to the degree to which they felt listened to in their class and/or the amount of choice they had in the curriculum.

Specific findings related to instruction. In the next phase of analysis, we took a deeper look into students' views on instructional practices. Overall, we found that among those responses coded as "actionable," 26 percent referred to specific instructional practices. Several themes emerged from these student responses, including (a) the clarity of explanation provided by teachers, (b) the provision of tools for learning (physical and experiential), (c) the pace of instruction, (d) assessment design, (e) allotted time on task, (f) connections to "real-life" experience, (g) differentiation, and (h) opportunities for critical thinking. These themes suggest that students consider not only their relationships with teachers and the quality of the classroom environment, but that they also reflect seriously on *how* their teachers are teaching.

In sum, results from the open-ended analyses suggest that students are in tune with effective teaching practices, and have a good sense of what effective instruction looks and feels like in practice. Taken together, these results suggest that students are well-poised to respond substantively – and even actionably – to survey items about the teaching practices they experience.

References

Johnstone, C. J., Bottsford-Miller, N. A., & Thompson, S. J. (2006). *Using the think aloud method (cognitive labs) to evaulate test design for students with disabilities and English language learners* (Technical Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from http://education.umn.edu/NCEO/OnlinePubs/Tech44/

Linacre, J. M. (2006). Misfit diagnosis: infit outfit mean-square standardized. Retrieved from http://www.winsteps.com/winman/diagnosingmisfit.htm

Nielson, J. (1994). Estimating the number of subjects needed for a thinking aloud test. *International Journal of Human-Computer Studies, 41*(3), 385–397.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149—174.

Willis, G. (1994). *Cognitive Interviewing and Questionnaire Design: A Training Manual*. National Center for Health Statistics, Cognitive Methods Staff, Working Paper No. 7.

Wilson, M. (2003). On choosing a model for measuring. *Methods of Psychological Research, 8*(3), 1-22.

Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Mahwah, NJ: Lawrence Erlbaum Associates.

SPS Technical Report: Appendices

Appendix A: *Full Instrument (Grades 3-5 & Grades 6-12)*

Table A1. *Grades 3-5 full instrument*

Element	Items (Response options: "Always," "Most of the time," "Some of the time," "Never")
Student Learning: *How teachers use content and pedagogical knowledge to help students learn, understand, and improve.*	The schoolwork we do helps me learn. What I learn in this class is useful to me in my real life. In this class, we learn a lot almost every day. My teacher makes sure that we think hard about things we read and write. When the work is too hard, my teacher helps me keep trying. In this class, it is more important to understand the lesson than to memorize the answers. My teacher uses a lot of different ways to explain things. My teacher knows when we understand the lesson and when we do not. Our classroom materials and supplies have a special place and things are easy to find. In this class, we learn to correct our mistakes. My teacher tells us what we are learning and why. My teacher asks questions to be sure we are following along. My teacher talks to me about my work to help me understand my mistakes. My teacher writes notes on my work that help me do better next time. The schoolwork we do is interesting.
Student-Centered Environment: *How teachers create an environment that responds to individual students' backgrounds, strengths, and interests.*	My teacher wants us to share what we think. My teacher teaches us to respect people's differences. My teacher knows what makes me excited about learning. My teacher talks about things we learn in other classes, subjects, and years. If I am sad or angry, my teacher helps me feel better. My teacher would notice if something was bothering me. The people we learn and read about in this class are like me. My teacher knows what my life is like outside of school. My teacher knows what is important to me. Students feel comfortable sharing their ideas in this class.
Classroom Community: *How teachers cultivate a classroom learning community where student differences are valued.*	My teacher cares about me. In this class, I feel like I fit in. I feel like an important part of my classroom community. I ask for help when I need it. I feel like I do a good job in this class. Our class stays busy and does not waste time. Students in my class are respectful to our teacher. My classmates behave the way my teacher wants them to. All of the kids in my class know what they are supposed to be doing and learning.
Classroom Management: *How teachers foster a respectful and predictable learning environment.*	Our class stays busy and does not waste time. Students in my class are respectful to our teacher. My classmates behave the way my teacher wants them to. All of the kids in my class know what they are supposed to be doing and learning.

Table A2. *Grades 6-12 full instrument*

Element	Items (Response options: "Always," "Most of the time," "Some of the time," "Never")
Student Learning: *How teachers use content and pedagogical knowledge to help students learn, understand, and improve.*	My teacher makes learning enjoyable. What I learn in this class is useful to me in my real life. My teacher teaches things that are important to me. My teacher knows the things that make me excited about learning. In this class, we learn a lot every day. In this class, it is more important to understand the lesson than to memorize the answers. When the work is too hard, my teacher helps me keep trying. My teacher accepts nothing less than my best effort. My teacher knows when we understand the lesson and when we do not. If I don't understand something, my teacher explains it a different way. My teacher explains difficult things clearly. In this class, we have a say in what we learn and do. My teacher talks to me about my work to help me understand my mistakes. My teacher writes notes on my work that help me improve. When we study a topic, my teacher makes connections to other subjects or classes.
Student-Centered Environment: *How teachers create an environment that responds to individual students' backgrounds, strengths, and interests.*	My classroom is organized and I know where to find what I need. Students feel comfortable sharing their ideas in this class. My teacher respects my opinions and suggestions. My teacher cares about me. My teacher pays attention to what all students are thinking and feeling. My teacher respects my cultural background. My teacher respects me as an individual.
Classroom Community: *How teachers cultivate a classroom learning community where student differences are valued.*	My teacher would notice if something was bothering me. Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background. In this class, I feel like I fit in. I feel like an important part of this classroom community. My teacher knows what my life is like outside of school. My teacher knows what is important to me. I ask for help when I need it. I feel like I do a good job in this class.
Classroom Management: *How teachers foster a respectful and predictable learning environment.*	Our class stays busy and does not waste time. Students in this class treat the teacher with respect. The students behave the way my teacher wants them to. Students in this class respect each other's differences.

Appendix B: *Comparison of Pre-Pilot, Fall Pilot and Final Items*

Table B1. *Grades 3-5 item development*

Grades 3-5 Pre-Pilot	Grades 3-5 Fall Pilot Items	Grades 3-5 Final Items
The schoolwork we do helps me learn.	The schoolwork we do helps me learn.	The schoolwork we do helps me learn.
The schoolwork we do is interesting.	The schoolwork we do is interesting.	The schoolwork we do is interesting.
---	What I learn in this class is useful to me in my real life.	What I learn in this class is useful to me in my real life.
I get bored in this class.	I get bored in this class.	My teacher knows what makes me excited about learning.
In this class, we learn a lot almost every day.	In this class, we learn a lot almost every day.	In this class, we learn a lot almost every day.
My teacher makes sure we think hard about the things we read and write.	My teacher makes sure that we think hard about things we read and write.	My teacher makes sure that we think hard about things we read and write.
My teacher doesn't let me give up when the work is hard.	When the work is too hard, my teacher helps me keep trying.	When the work is too hard, my teacher helps me keep trying.
My teacher wants us to understand what we learn, not just memorize facts.	In this class, it is more important to understand the lesson than to memorize the answers.	In this class, it is more important to understand the lesson than to memorize the answers.
My teacher uses a lot of different ways to explain things.	My teacher uses a lot of different ways to explain things.	My teacher uses a lot of different ways to explain things.
My teacher knows when we understand the lesson and when we do not.	My teacher knows when we understand the lesson and when we do not.	My teacher knows when we understand the lesson and when we do not.
Our classroom materials and supplies have a special place and things are easy to find.	Our classroom materials and supplies have a special place and things are easy to find.	Our classroom materials and supplies have a special place and things are easy to find.
---	---	In this class, we learn to correct our mistakes.
My teacher tells us what we are learning and why.	My teacher tells us what we are learning and why.	My teacher tells us what we are learning and why.
My teacher wants us to share what we think.	My teacher wants us to share what we think.	My teacher wants us to share what we think.
Students speak up and share their ideas about class work.	Students feel comfortable sharing their ideas in this class.	Students feel comfortable sharing their ideas in this class.
---	---	My teacher asks questions to be sure we are following along.
My teacher talks to me about my work to help me understand my mistakes.	My teacher talks to me about my work to help me understand my mistakes.	My teacher talks to me about my work to help me understand my mistakes.
My teacher writes notes on my work that help me do better next time.	My teacher writes notes on my work that help me do better next time.	My teacher writes notes on my work that help me do better next time.
My teacher builds on things we are learning in other classes.	My teacher builds on things we learn in other classes, subjects, and years.	My teacher talks about things we learn in other classes, subjects, and years.
My teacher cares about me.	My teacher cares about me.	My teacher cares about me.
If I am sad or angry, my teacher helps me feel better.	If I am sad or angry, my teacher helps me feel better.	If I am sad or angry, my teacher helps me feel better.
My teacher knows if something is bothering me.	My teacher would notice if something was bothering me.	My teacher would notice if something was bothering me.
We waste time in this class.	We waste time in this class.	Our class stays busy and does not waste time.
Students in my class are respectful to our teacher.	Students in my class are respectful to our teacher.	Students in my class are respectful to our teacher.

My classmates behave the way my teacher wants them to.	My classmates behave the way my teacher wants them to.	My classmates behave the way my teacher wants them to.
Everybody knows what they should be doing and learning.	All of the kids in my class know what they are supposed to be doing and learning.	All of the kids in my class know what they are supposed to be doing and learning.
Students behave so badly in this class that it slows down our learning.	Students behave so badly in this class that it slows down our learning.	---
The people in our classroom pictures, books, and art look like me.	The people we learn and read about in this class are like me.	The people we learn and read about in this class are like me.
---	My teacher teaches us to respect people's differences.	My teacher teaches us to respect people's differences.
---	In this class, I feel like I fit in.	In this class, I feel like I fit in.
I feel like an important part of my classroom community.	I feel like an important part of my classroom community.	I feel like an important part of my classroom community.
My teacher knows what my life is like outside of school.	My teacher knows what my life is like outside of school.	My teacher knows what my life is like outside of school.
---	My teacher knows what is important to me.	My teacher knows what is important to me.
School work in this class is too easy.	School work in this class is too easy.	---
I ask for help when I need it.	I ask for help when I need it.	I ask for help when I need it.
I feel like I do a good job in this class.	I feel like I do a good job in this class.	I feel like I do a good job in this class.

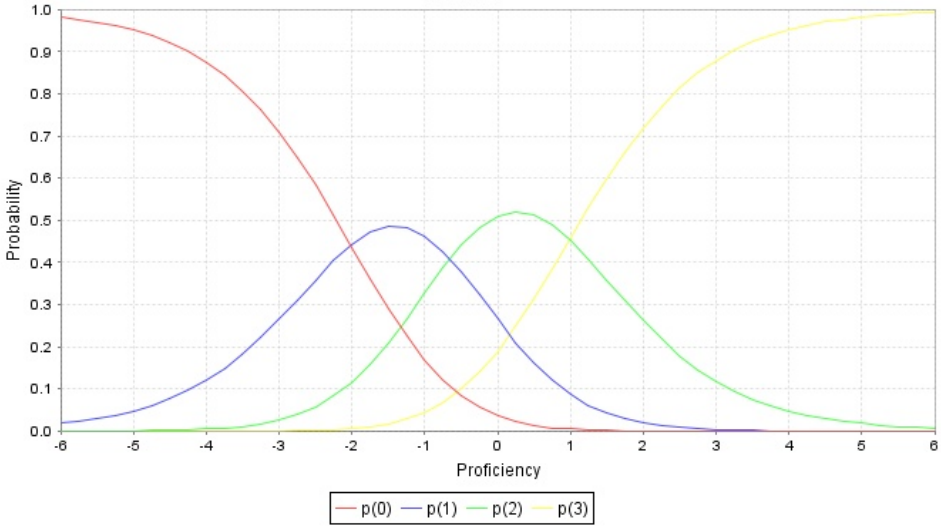
Table B2. *Grades 6-12 item development*

Grades 6-12 Pre-Pilot	Grades 6-12 Fall Pilot Items	Grades 6-12 Final Items
My teacher makes learning enjoyable.	My teacher makes learning enjoyable.	My teacher makes learning enjoyable.
My teacher makes lessons relevant to me.	What I learn in this class is useful to me in my real life.	What I learn in this class is useful to me in my real life.
---	My teacher teaches things that are important to me.	My teacher teaches things that are important to me.
---	My teacher knows the things that make me excited about learning	My teacher knows the things that make me excited about learning.
I get bored in this class.	I get bored in this class.	---
In this class, we learn a lot every day.	In this class, we learn a lot every day.	In this class, we learn a lot every day.
My teacher wants us to understand what we learn, not just memorize facts.	In this class, it is more important to understand the lesson than to memorize the answers.	In this class, it is more important to understand the lesson than to memorize the answers.
My teacher doesn’t let me give up when the work gets hard.	When the work is too hard, my teacher helps me keep trying.	When the work is too hard, my teacher helps me keep trying.
My teacher accepts nothing less than my best effort.	My teacher accepts nothing less than my best effort.	My teacher accepts nothing less than my best effort.
My teacher knows when we understand the lesson and when we do not.	My teacher knows when we understand the lesson and when we do not.	My teacher knows when we understand the lesson and when we do not.
If I don’t understand something, my teacher explains it a different way.	If I don't understand something, my teacher explains it a different way.	If I don't understand something, my teacher explains it a different way.
My teacher explains difficult things clearly.	My teacher explains difficult things clearly.	My teacher explains difficult things clearly.

My classroom is organized and I know where to find what I need.	My classroom is organized and I know where to find what I need.	My classroom is organized and I know where to find what I need.
Students speak up and share their ideas in this class.	Students feel comfortable sharing their ideas in this class.	Students feel comfortable sharing their ideas in this class.
My teacher respects my opinions and suggestions.	My teacher respects my opinions and suggestions.	My teacher respects my opinions and suggestions.
In this class, we have input on what we learn and do.	In this class, we have a say in what we learn and do.	In this class, we have a say in what we learn and do.
My teacher talks to me about my work to help me understand my mistakes.	My teacher talks to me about my work to help me understand my mistakes.	My teacher talks to me about my work to help me understand my mistakes.
My teacher writes notes on my work that help me improve.	My teacher writes notes on my work that help me improve.	My teacher writes notes on my work that help me improve.
When we study a topic, my teacher makes connections to other subjects or classes.	When we study a topic, my teacher makes connections to other subjects or classes.	When we study a topic, my teacher makes connections to other subjects or classes.
My teacher cares about me.	My teacher cares about me.	My teacher cares about me.
My teacher pays attention to what all students are thinking and feeling.	My teacher pays attention to what all students are thinking and feeling.	My teacher pays attention to what all students are thinking and feeling.
My teacher knows if something is bothering me.	My teacher would notice if something was bothering me.	My teacher would notice if something was bothering me.
We waste time in this class.	We waste time in this class.	Our class stays busy and does not waste time.
Students in this class treat the teacher with respect.	Students in this class treat the teacher with respect.	Students in this class treat the teacher with respect.
The students behave the way my teacher wants them to.	The students behave the way my teacher wants them to.	---
Student behavior in this class makes the teacher angry.	Student behavior in this class makes the teacher angry.	The way students behave in this class makes it hard to learn.
The classroom materials, pictures, words, books, and art reflect my cultural background.	The classroom materials, pictures, words, books, and art reflect my cultural background.	Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background.
My teacher respects my cultural background.	My teacher respects my cultural background.	My teacher respects my cultural background.
---	My teacher respects me as an individual.	My teacher respects me as an individual.
---	Students in this class respect each other's differences.	Students in this class respect each other's differences.
---	In this class, I feel like I fit in.	In this class, I feel like I fit in.
I feel like an important part of the classroom community.	I feel like an important part of this classroom community.	I feel like an important part of this classroom community.
My teacher knows what my life is like outside of school	My teacher knows what my life is like outside of school.	My teacher knows what my life is like outside of school.
---	My teacher knows what is important to me.	My teacher knows what is important to me.
I ask for help when I need it.	I ask for help when I need it.	I ask for help when I need it.
I feel like I do a good job in this class.	I feel like I do a good job in this class.	I feel like I do a good job in this class.
School work in this class is too easy.	School work in this class is too easy.	---

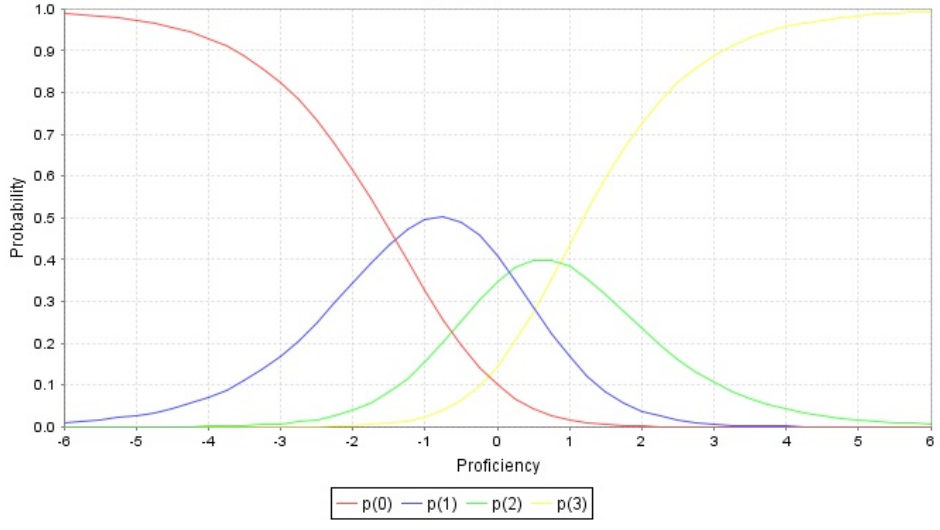
Appendix C: *Item Characteristic Curves, Grades 3-5*

1. The schoolwork we do helps me learn.
Item Set : base
Item : Q01

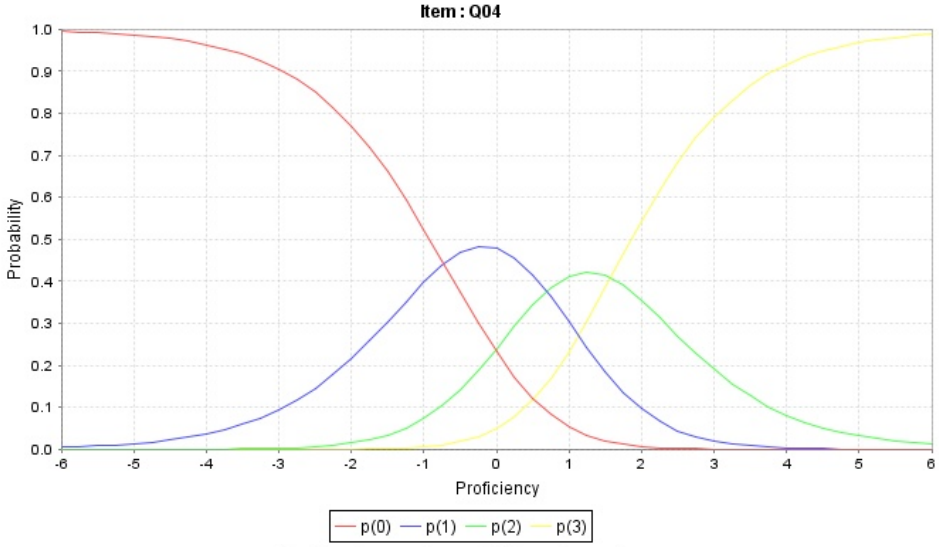


2. The schoolwork we do is interesting.
Item Set : base
Item : Q02

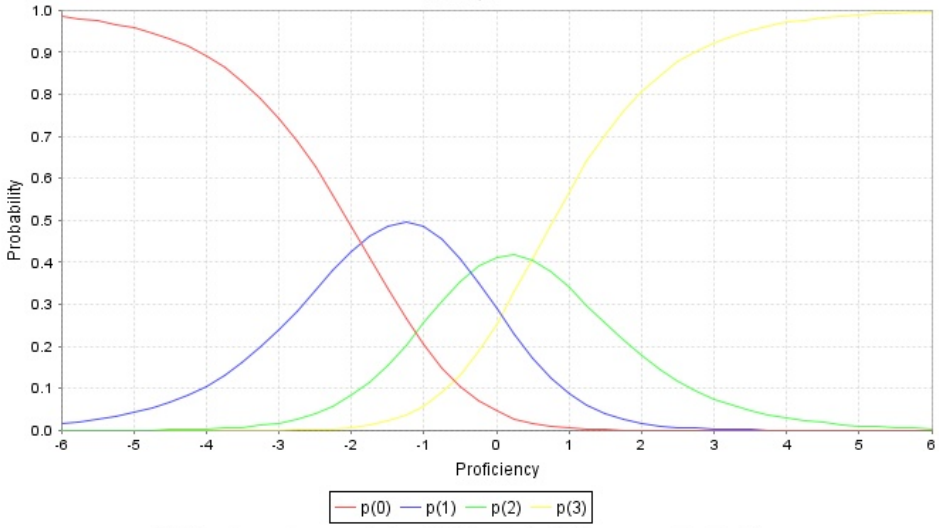
3. What I learn in this class is useful to me in my real life.
Item Set : base
Item : Q03



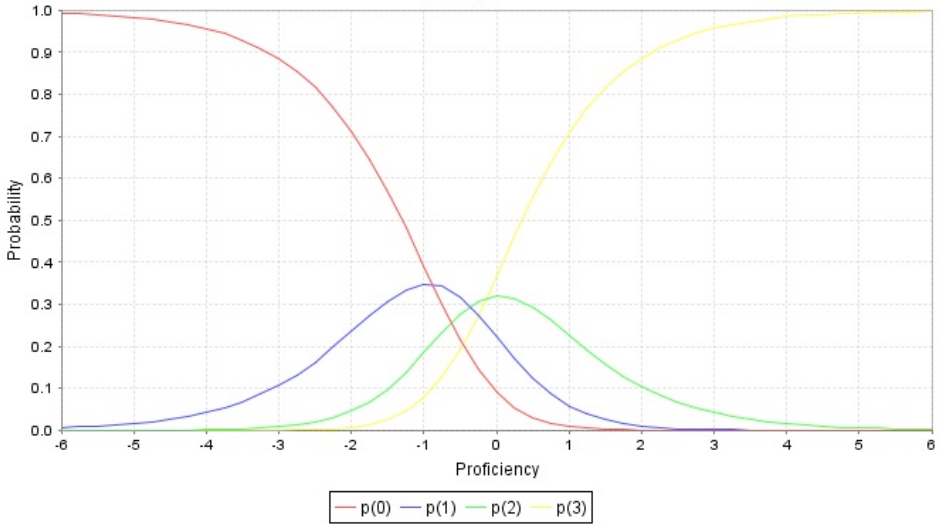
4. My teacher knows what makes me excited about learning.
Item Set : base
Item : Q04



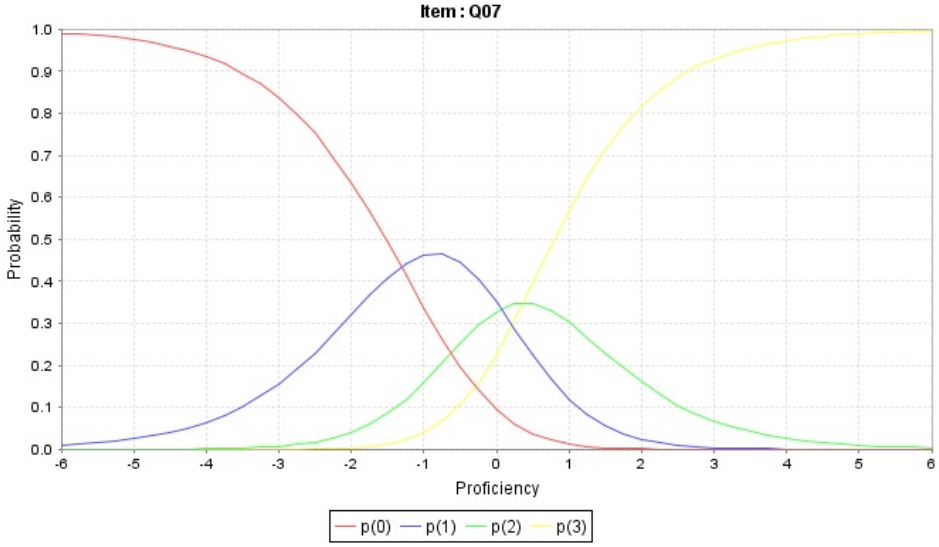
5. In this class, we learn a lot almost every day.
Item Set : base
Item : Q05



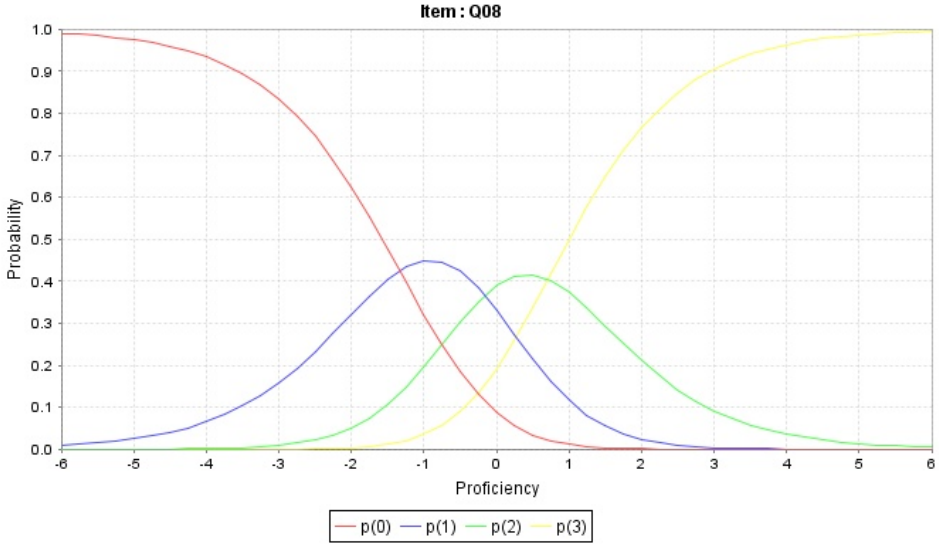
6. My teacher makes sure that we think hard about things we read and write.
Item Set : base
Item : Q06



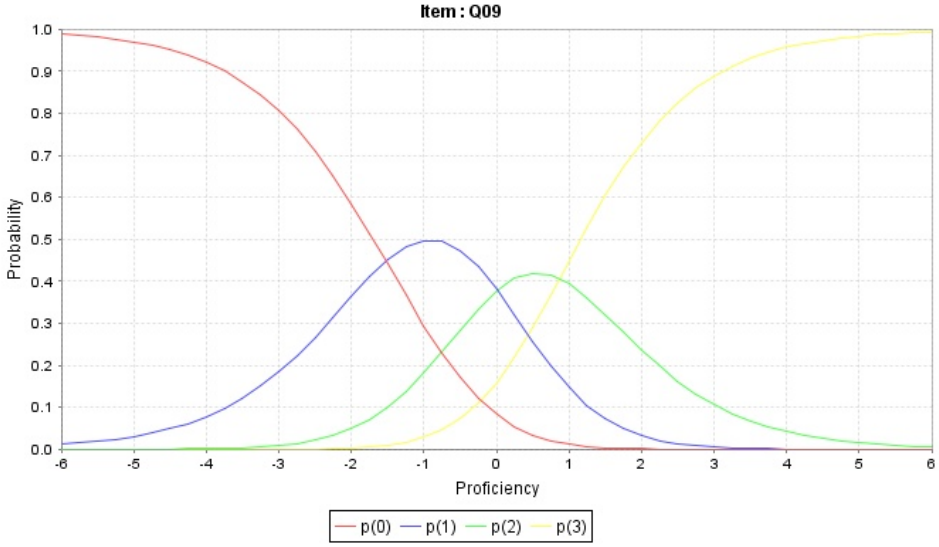
7. When the work is too hard, my teacher helps me keep trying.
Item Set : base
Item : Q07



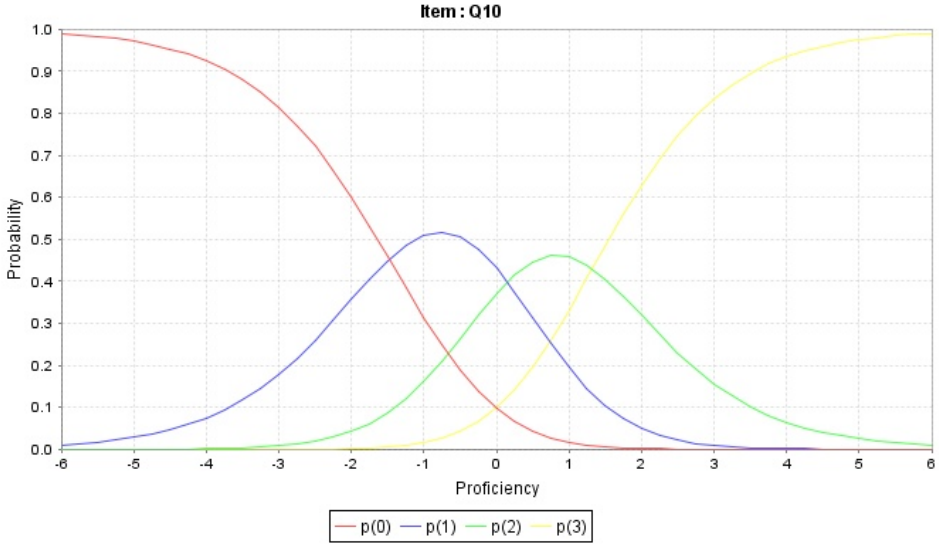
8. In this class, it is more important to understand the lesson than to memorize the answers.
Item Set : base
Item : Q08



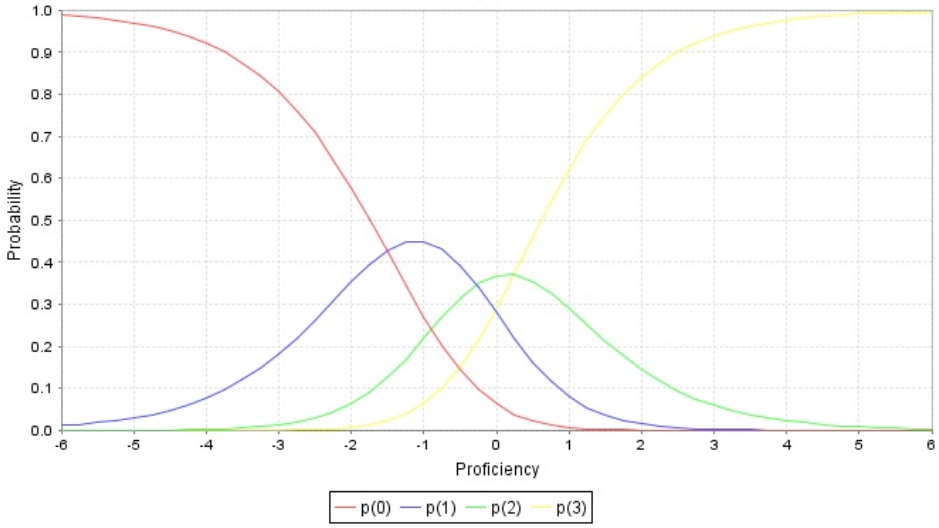
9. My teacher uses a lot of different ways to explain things.
Item Set : base
Item : Q09



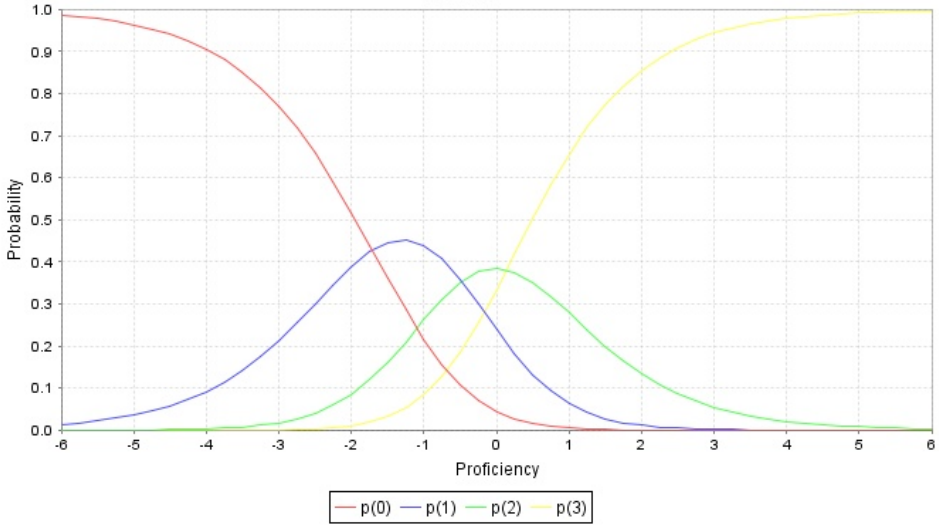
10. My teacher knows when we understand the lesson and when we do not.
Item Set : base
Item : Q10



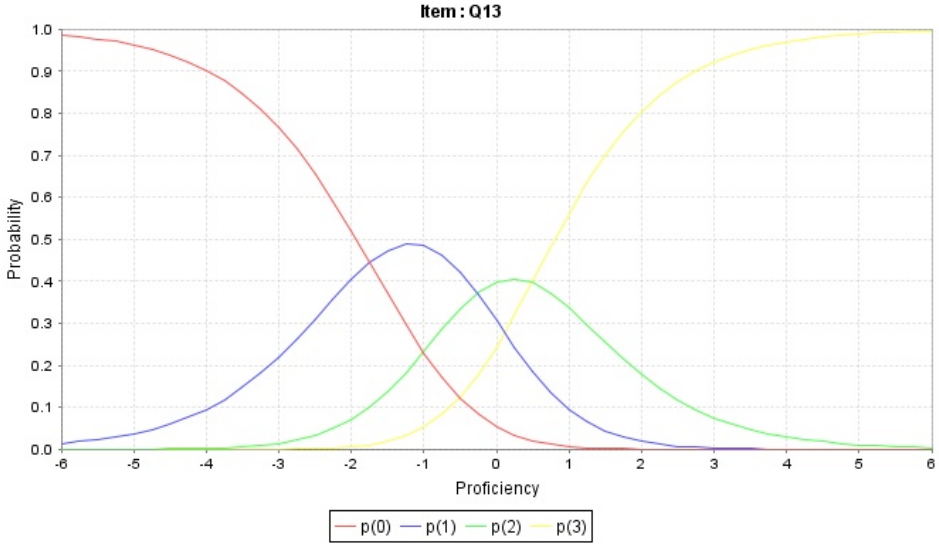
11. Our classroom materials and supplies have a special place and things are easy to find.
Item Set : base
Item : Q11



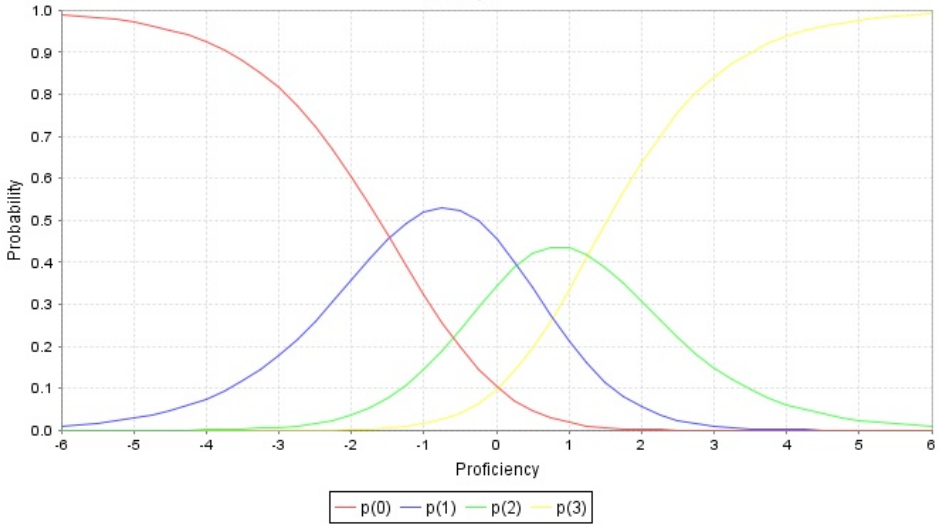
12. In this class, we learn to correct our mistakes.
Item Set : base
Item : Q12



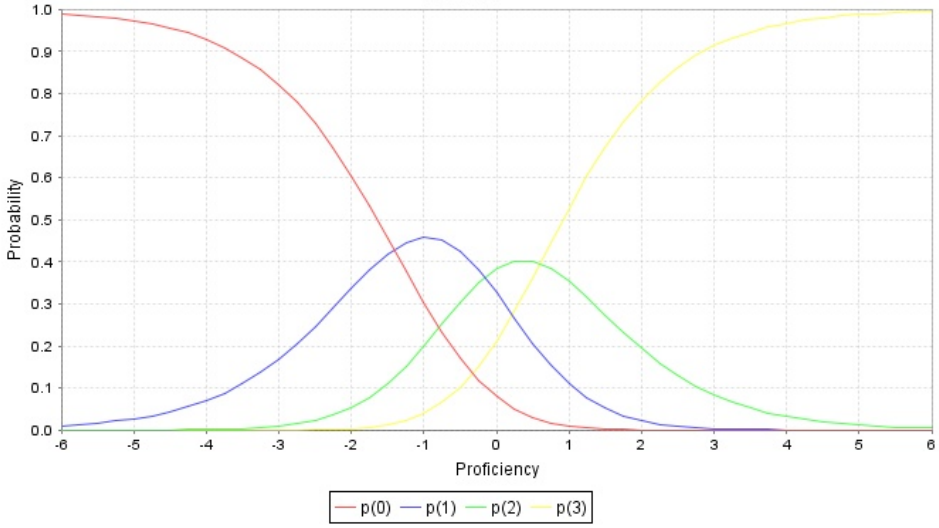
13. My teacher tells us what we are learning and why.
Item Set : base
Item : Q13



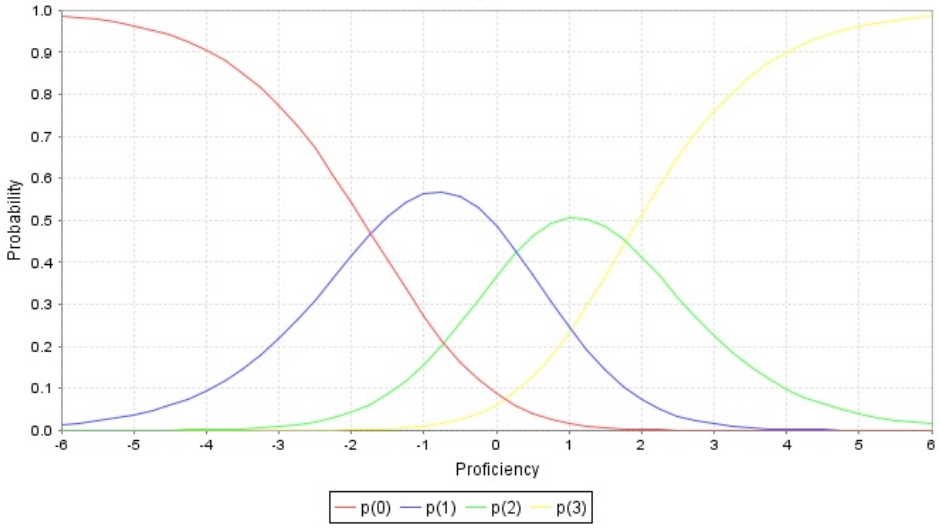
14. My teacher wants us to share what we think.
Item Set : base
Item : Q14



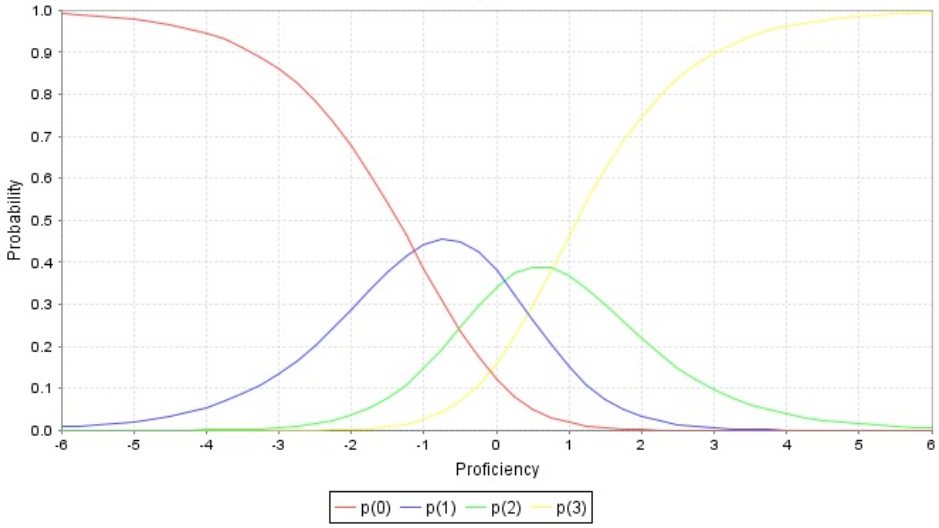
15. My teacher asks questions to be sure we are following along.
Item Set : base
Item : Q15



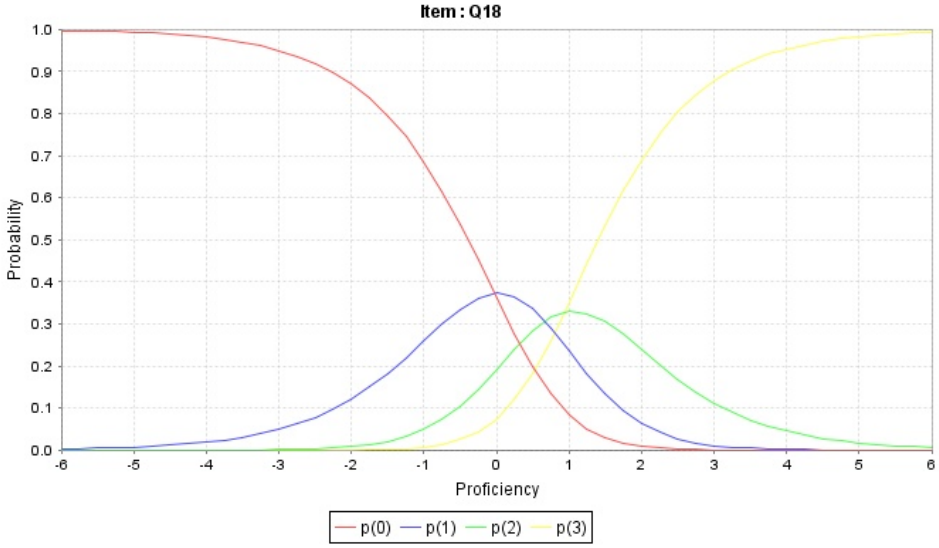
16. Students feel comfortable sharing their ideas in this class.
Item Set : base
Item : Q16



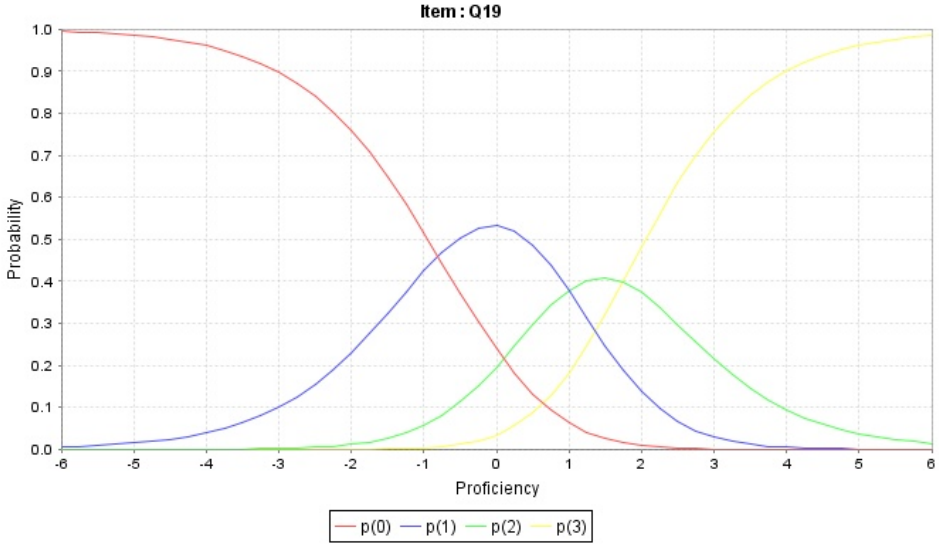
17. My teacher talks to me about my work to help me understand my mistakes.
Item Set : base
Item : Q17



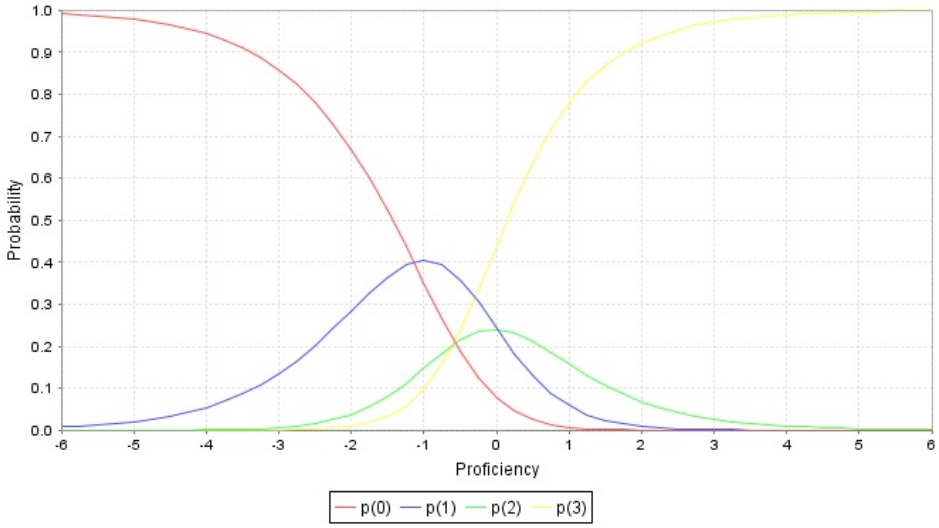
18. My teacher writes notes on my work that help me do better next time.
Item Set : base
Item : Q18



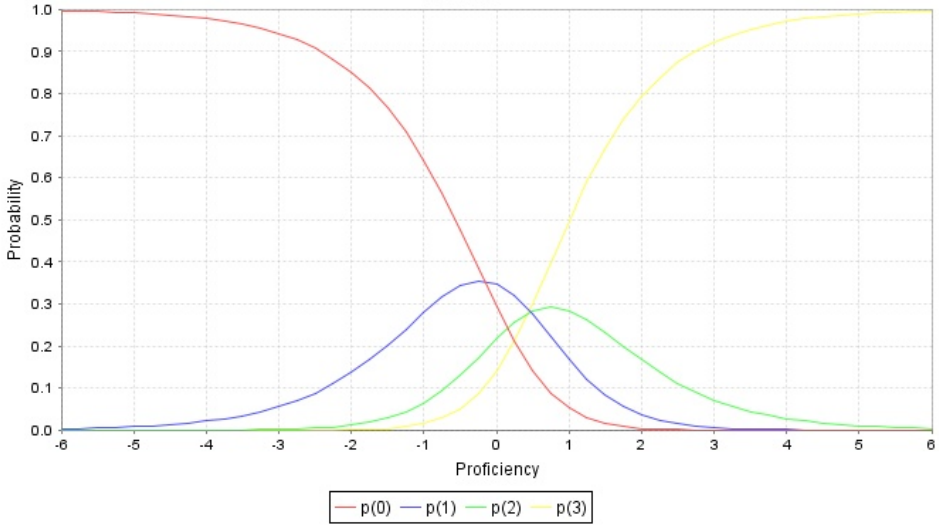
19. My teacher talks about things we learn in other classes, subjects, and years.
Item Set : base
Item : Q19



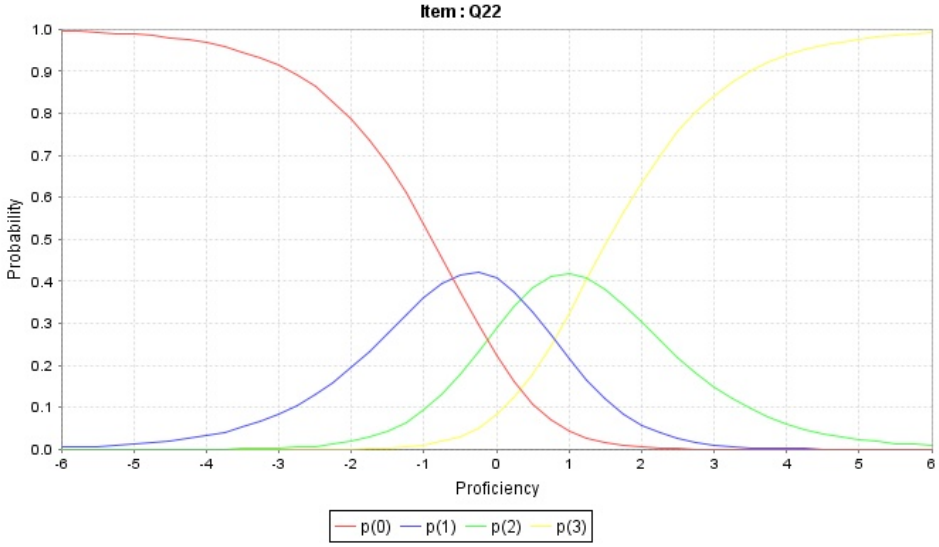
20. My teacher cares about me.
Item Set : base
Item : Q20



21. If I am sad or angry, my teacher helps me feel better.
Item Set : base
Item : Q21

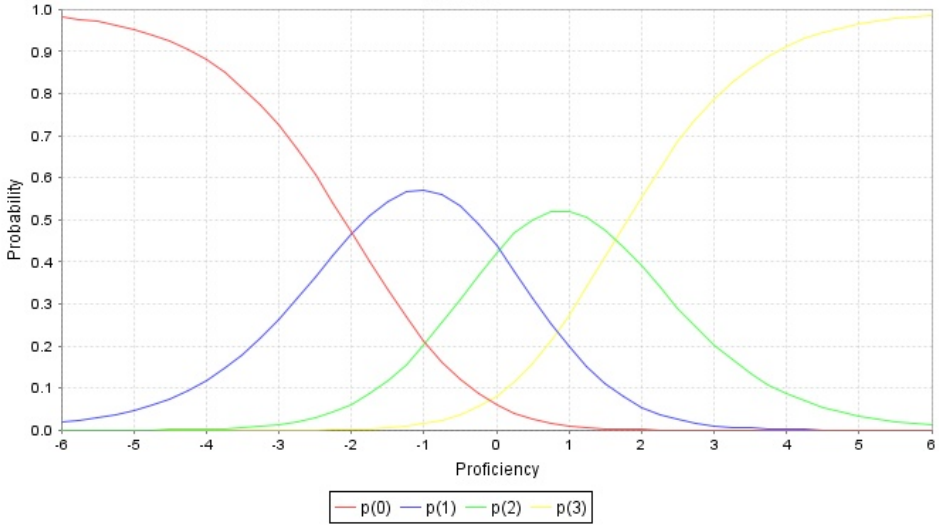


22. My teacher would notice if something was bothering me.
Item Set : base
Item : Q22

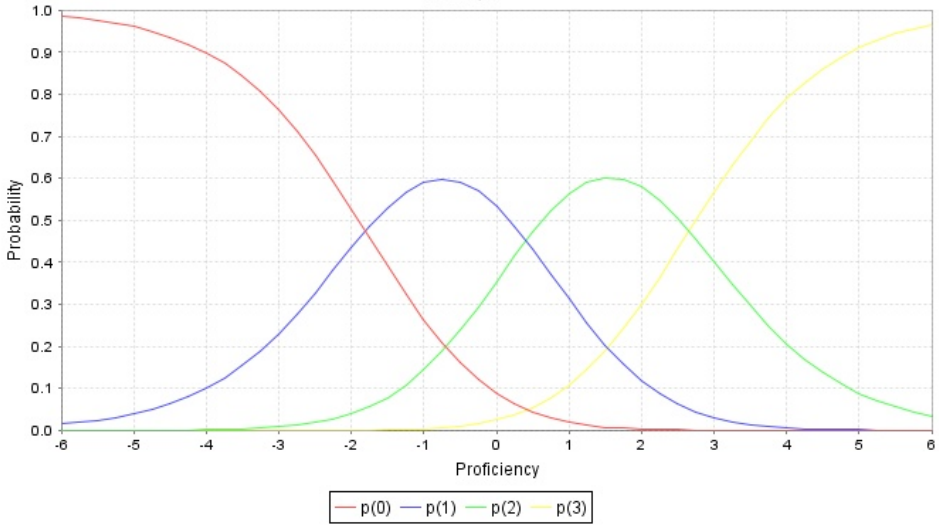


23. Our class stays busy and does not waste time.
Item Set : base
Item : Q23

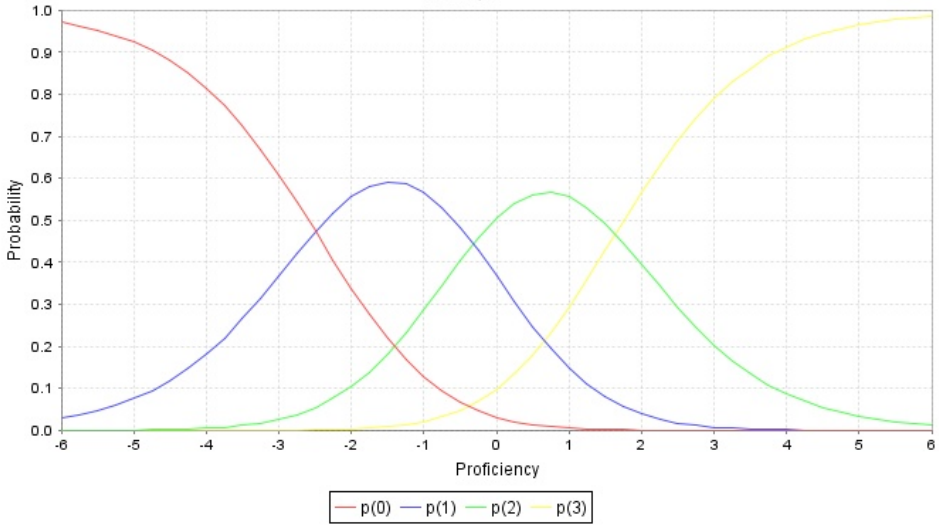
24. Students in my class are respectful to our teacher.
Item Set : base
Item : Q24



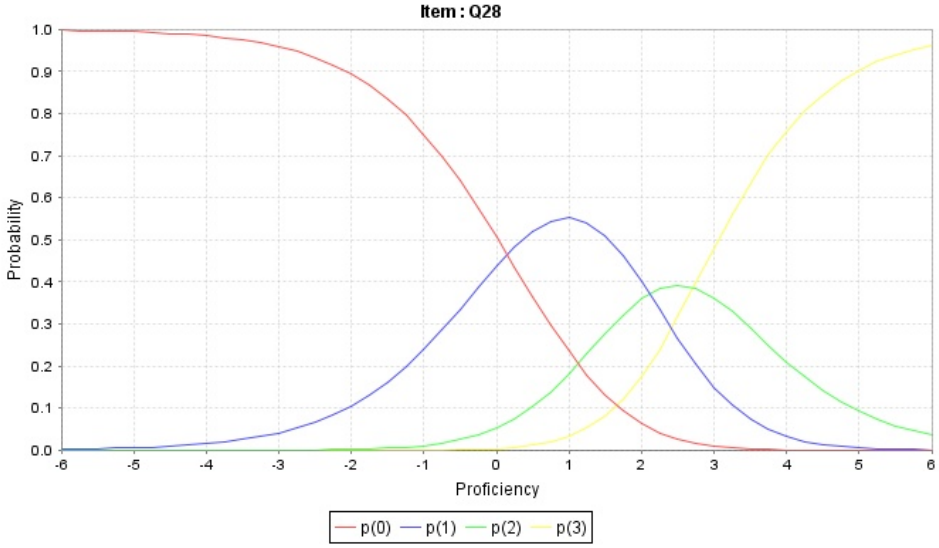
25. My classmates behave the way my teacher wants them to.
Item Set : base
Item : Q25



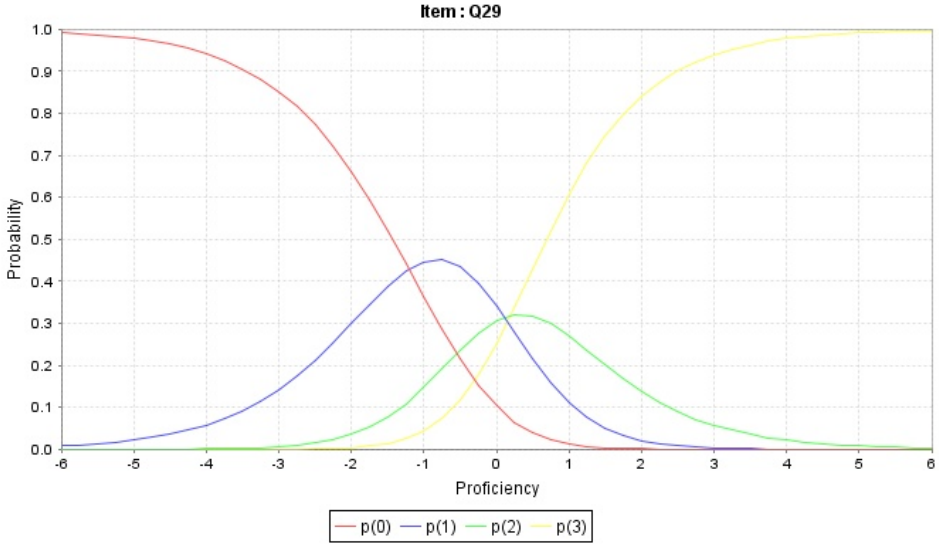
26. All of the kids in my class know what they are supposed to be doing and learning.
Item Set : base
Item : Q26



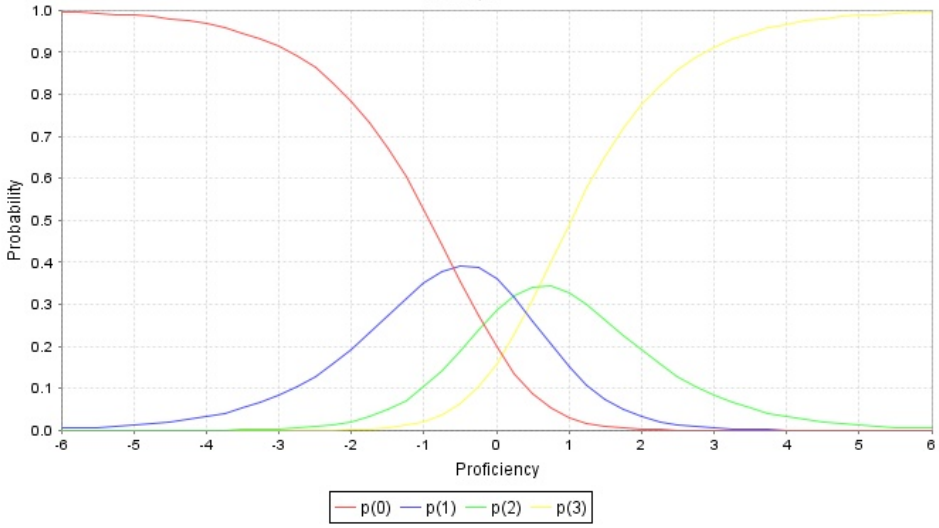
28. The people we learn and read about in this class are like me.
Item Set : base
Item : Q28



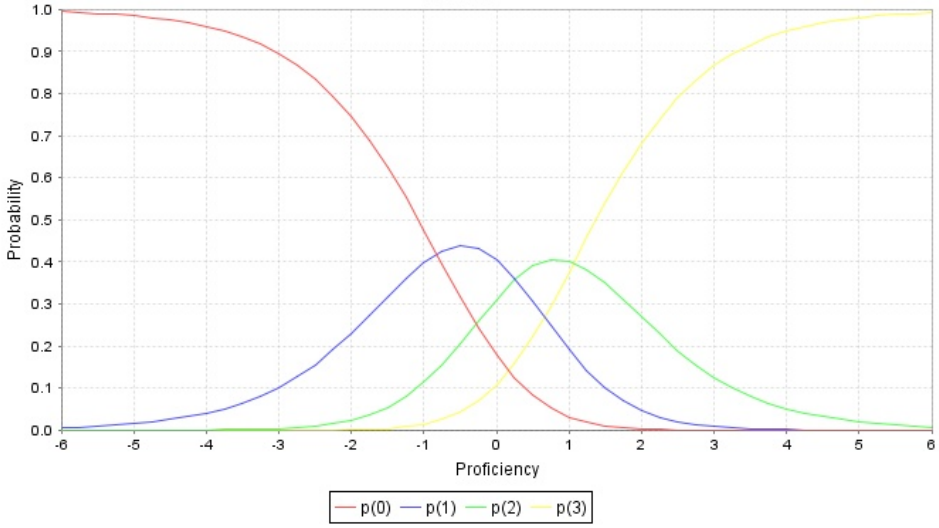
29. My teacher teaches us to respect people's differences.
Item Set : base
Item : Q29



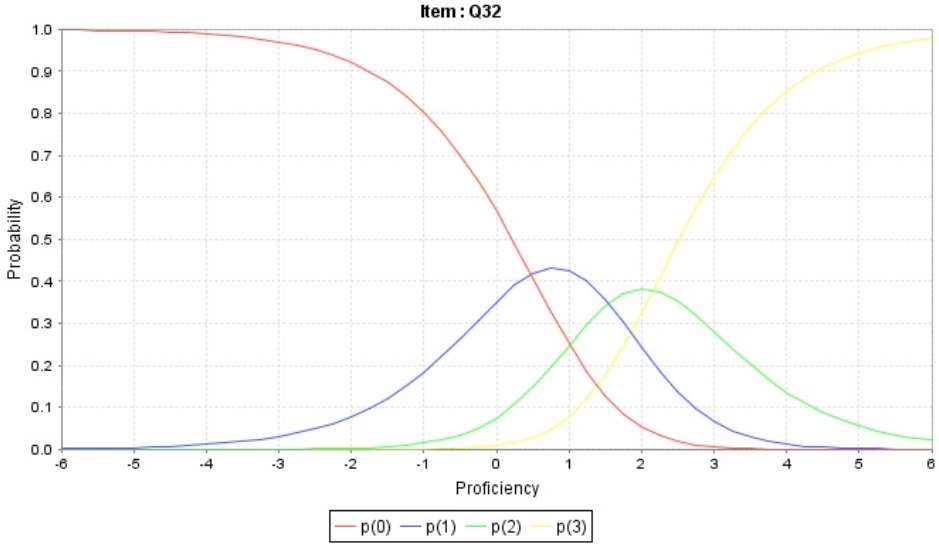
30. In this class, I feel like I fit in.
Item Set : base
Item : Q30



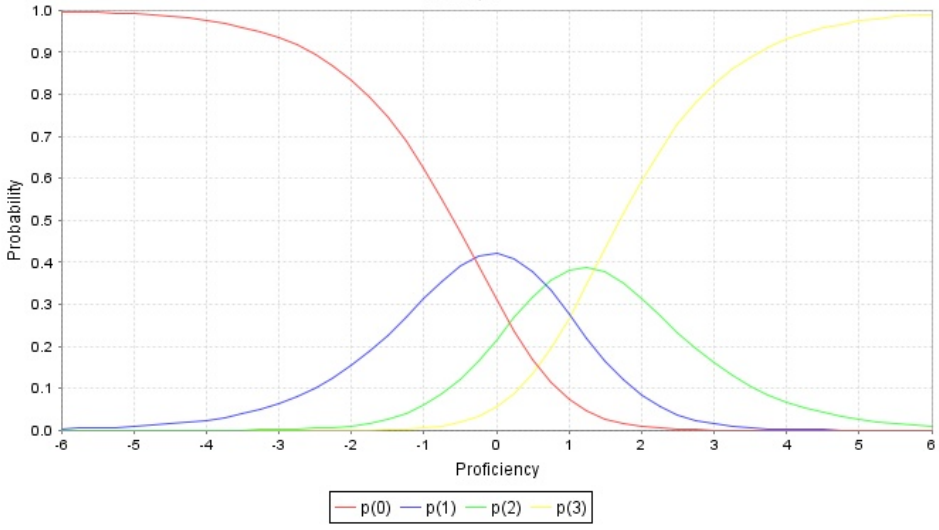
31. I feel like an important part of my classroom community.
Item Set : base
Item : Q31



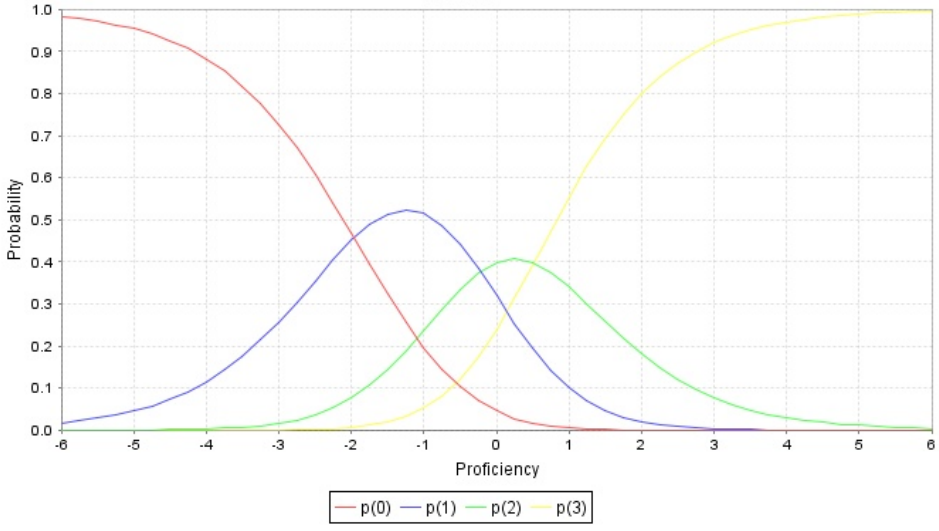
32. My teacher knows what my life is like outside of school.
Item Set : base
Item : Q32

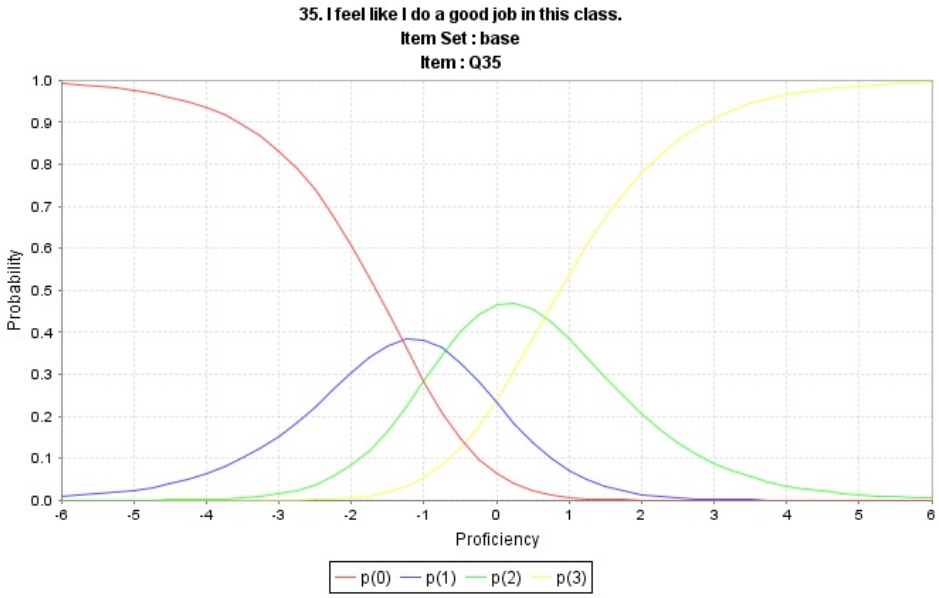


33. My teacher knows what is important to me.
Item Set : base
Item : Q33



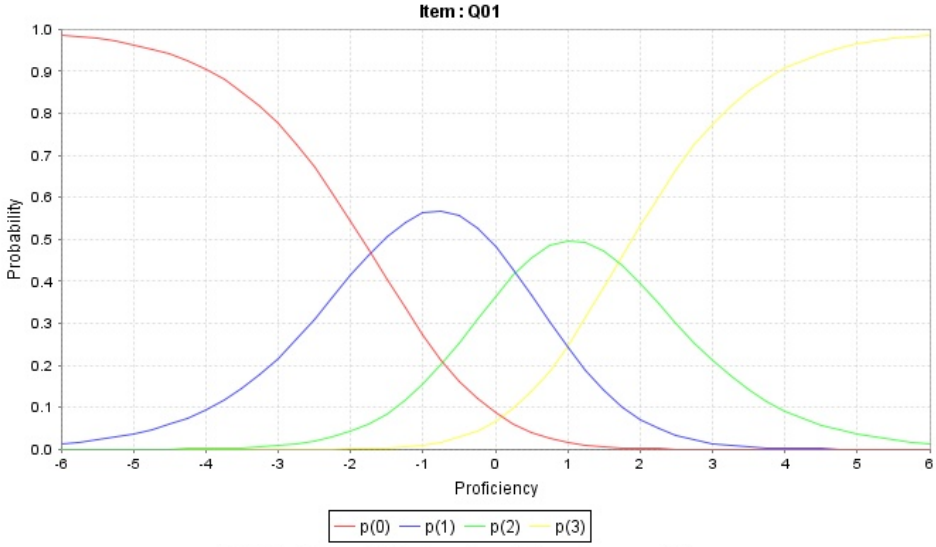
34. I ask for help when I need it.
Item Set : base
Item : Q34



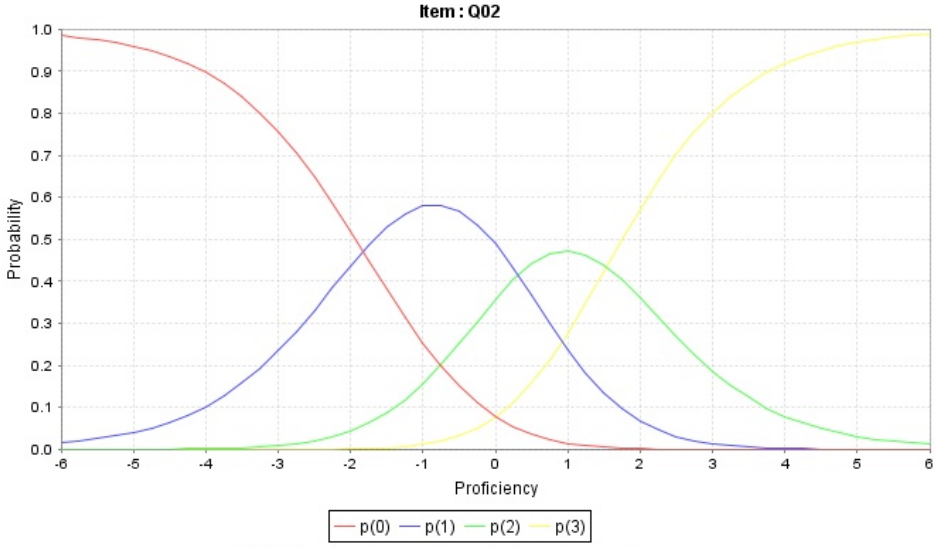

35. I feel like I do a good job in this class.
Item Set : base
Item : Q35
Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
Proficiency
p(0)
p(1)
p(2)
p(3)

Appendix D: *Item Characteristic Curves, Grades 6-12*

Q01. My teacher makes learning enjoyable.
Item Set : base
Item : Q01

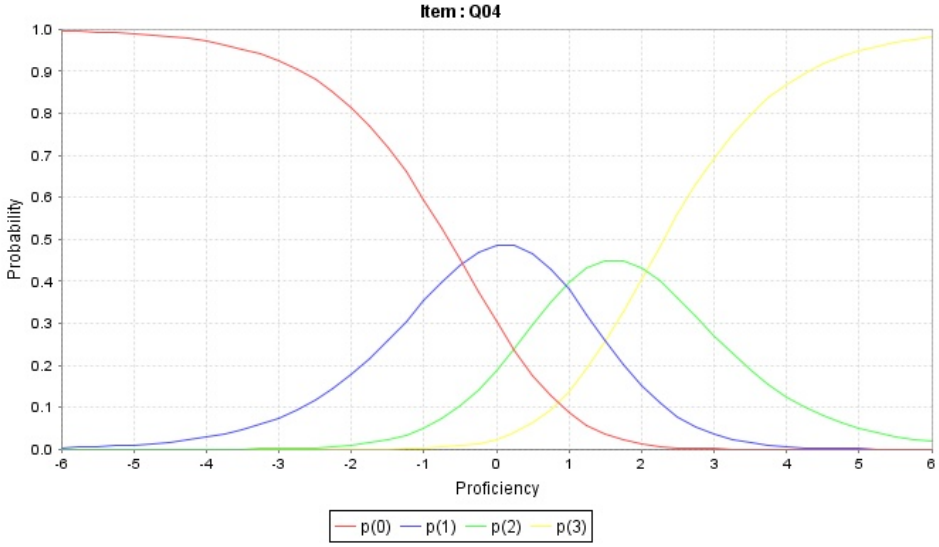


Q02. What I learn in this class is useful to me in my real life.
Item Set : base
Item : Q02

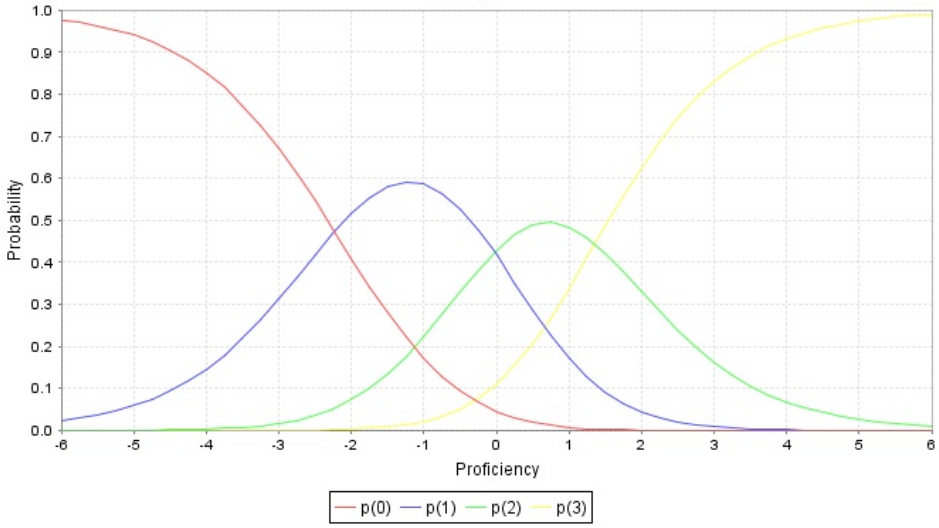


Q03. My teacher teaches things that are important to me.
Item Set : base
Item : Q03

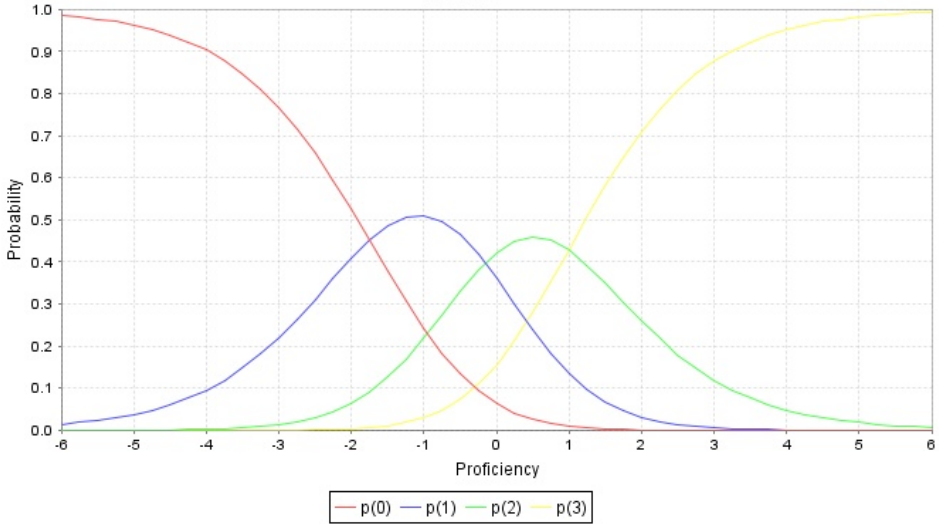
Q04. My teacher knows the things that make me excited about learning.
Item Set : base
Item : Q04



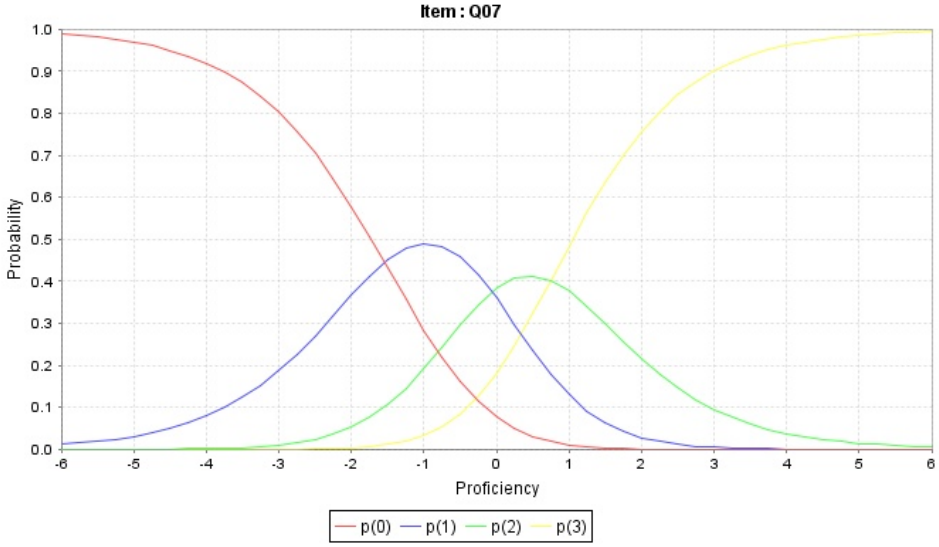
Q05. In this class, we learn a lot every day.
Item Set : base
Item : Q05



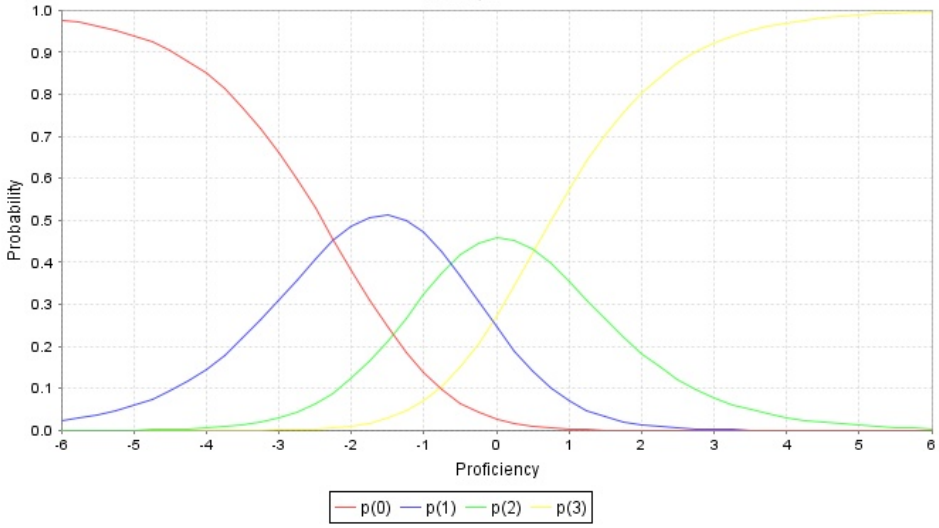
Q06. In this class, it is more important to understand the lesson than to memorize the answers.
Item Set : base
Item : Q06



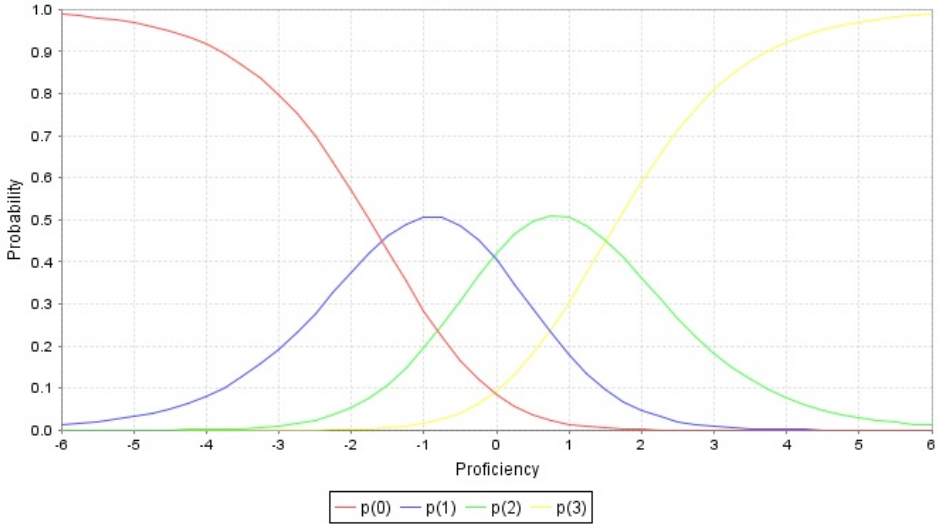
Q07. When the work is too hard, my teacher helps me keep trying.
Item Set : base
Item : Q07



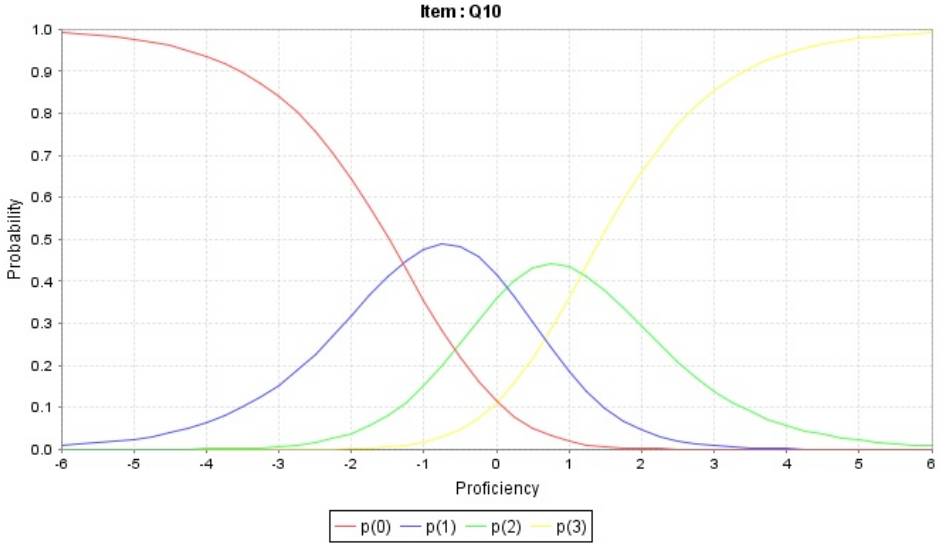
Q08. My teacher accepts nothing less than my best effort.
Item Set : base
Item : Q08



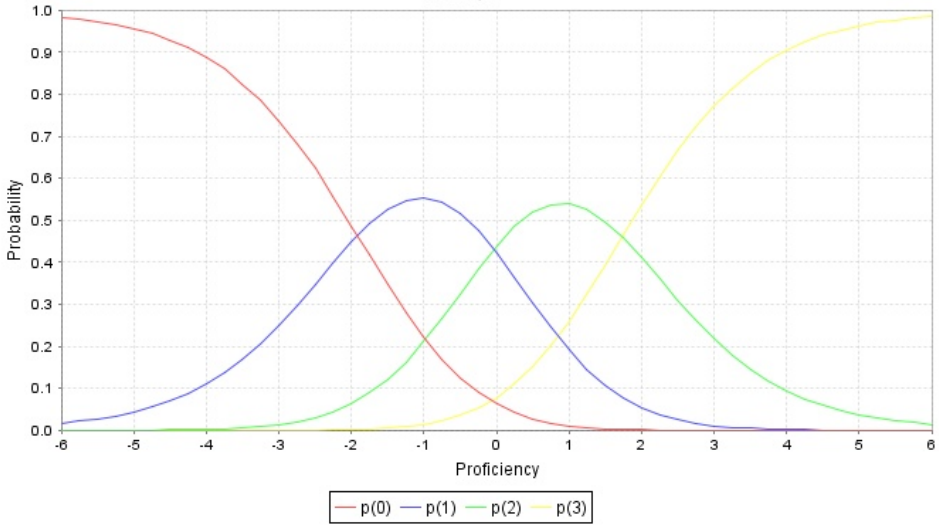
Q09. My teacher knows when we understand the lesson and when we do not.
Item Set : base
Item : Q09



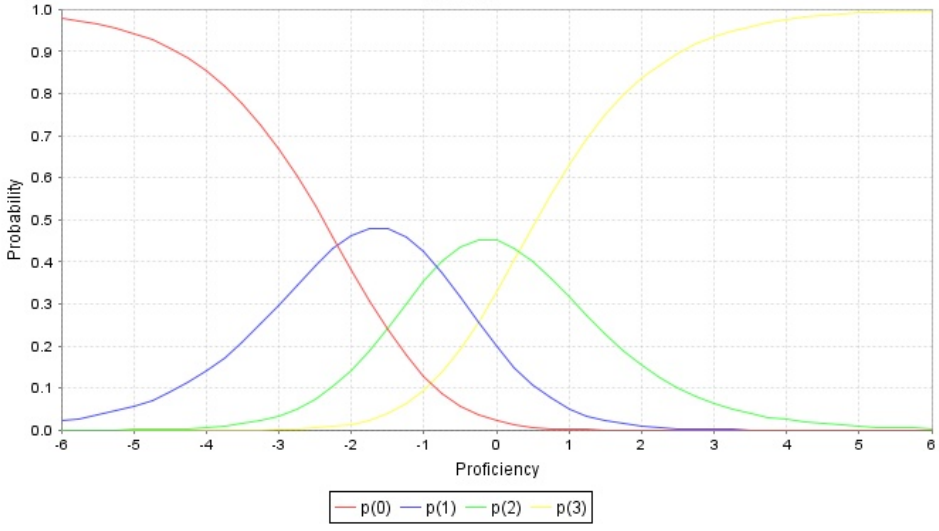
Q10. If I don't understand something, my teacher explains it a different way.
Item Set : base
Item : Q10



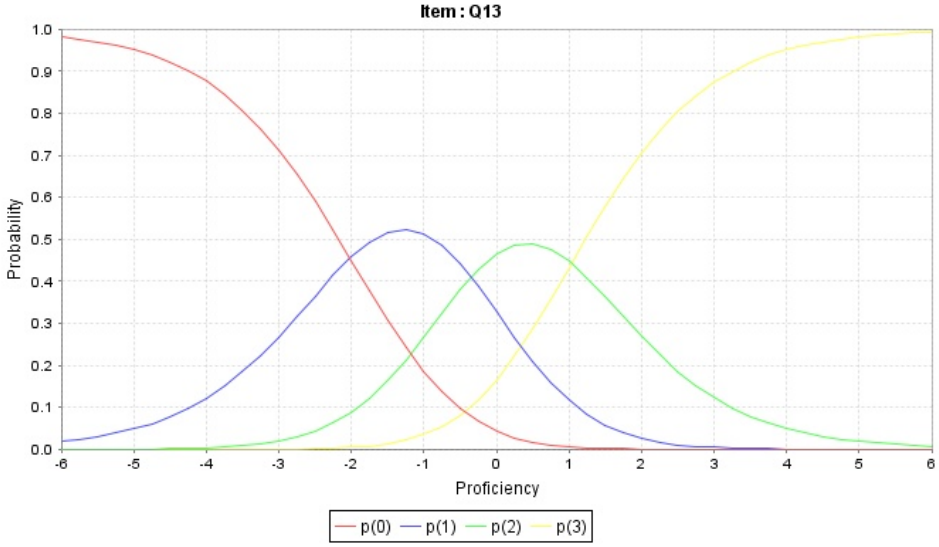
Q11. My teacher explains difficult things clearly.
Item Set : base
Item : Q11



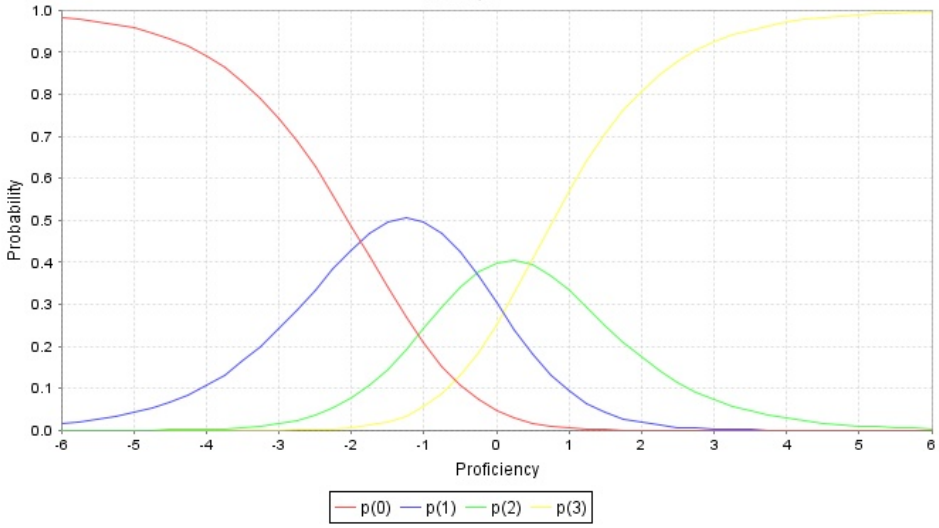
Q12. My classroom is organized and I know where to find what I need.
Item Set : base
Item : Q12



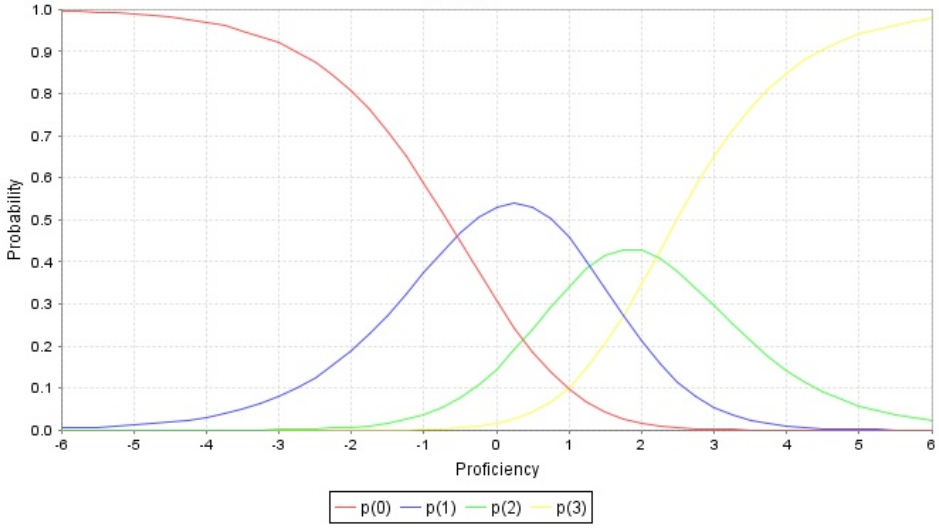
Q13. Students feel comfortable sharing their ideas in this class.
Item Set : base
Item : Q13



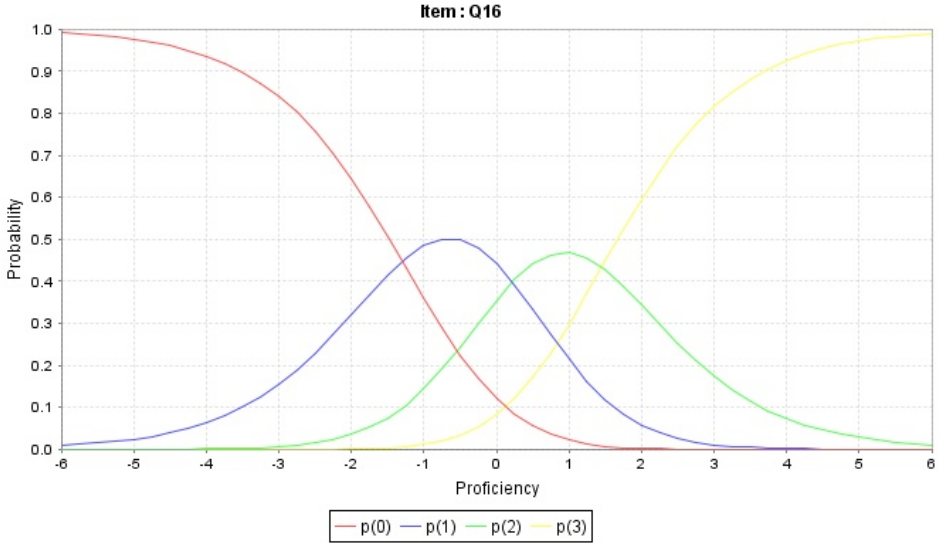
Q14. My teacher respects my opinions and suggestions.
Item Set : base
Item : Q14



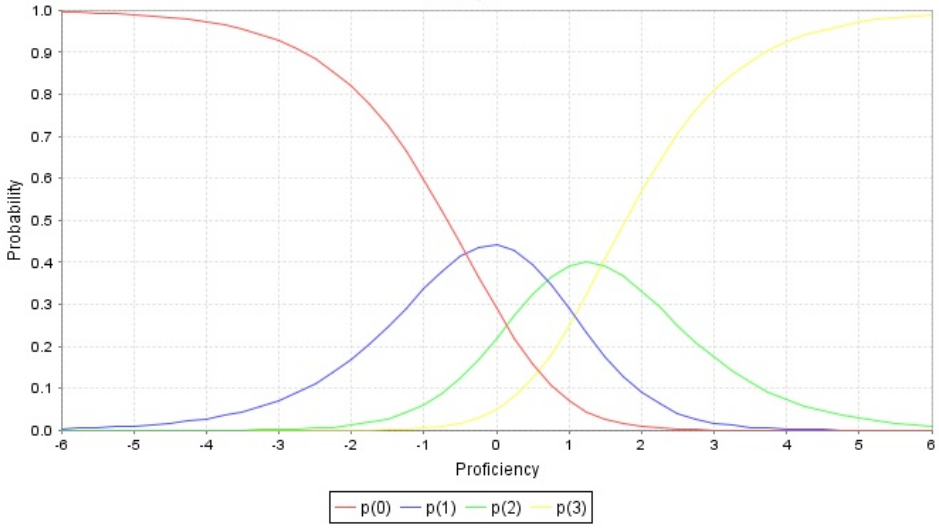
Q15. In this class, we have a say in what we learn and do.
Item Set : base
Item : Q15



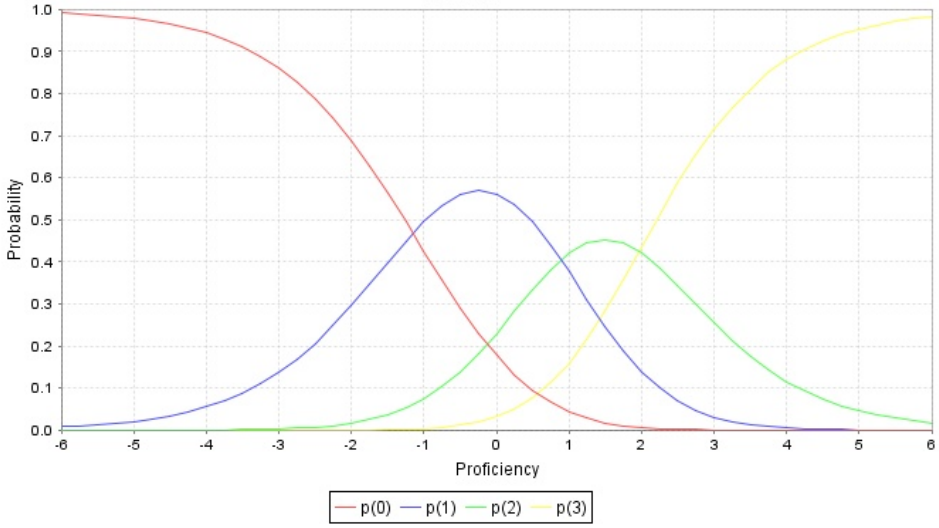
Q16. My teacher talks to me about my work to help me understand my mistakes.
Item Set : base
Item : Q16



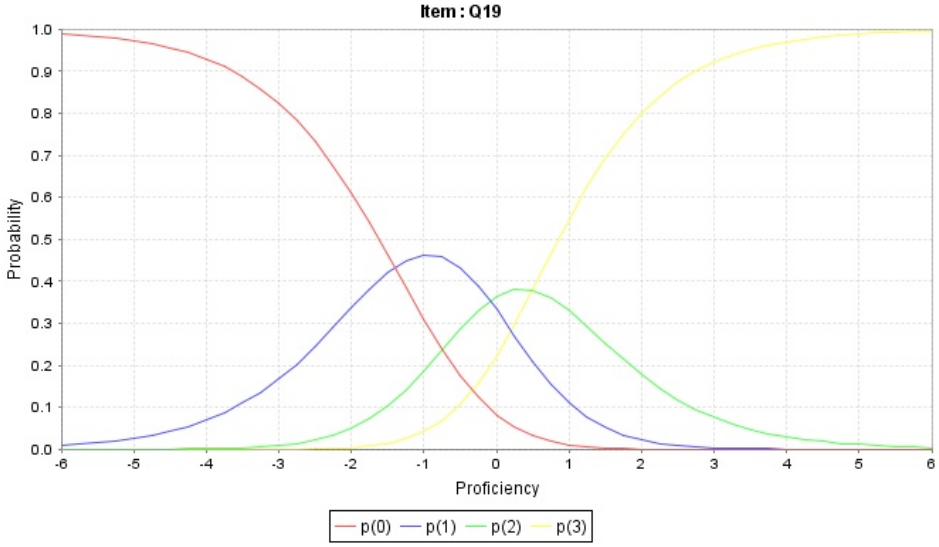
Q17. My teacher writes notes on my work that help me improve.
Item Set : base
Item : Q17



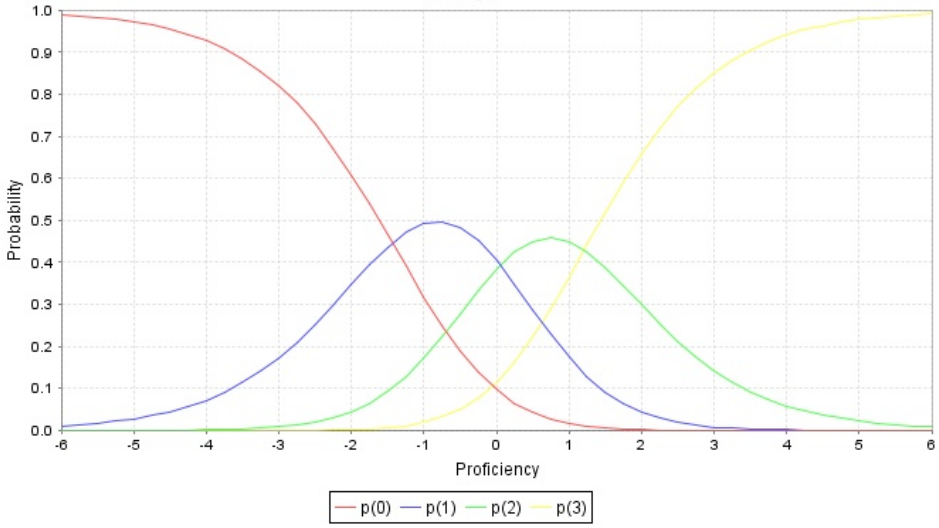
Q18. When we study a topic, my teacher makes connections to other subjects or classes.
Item Set : base
Item : Q18



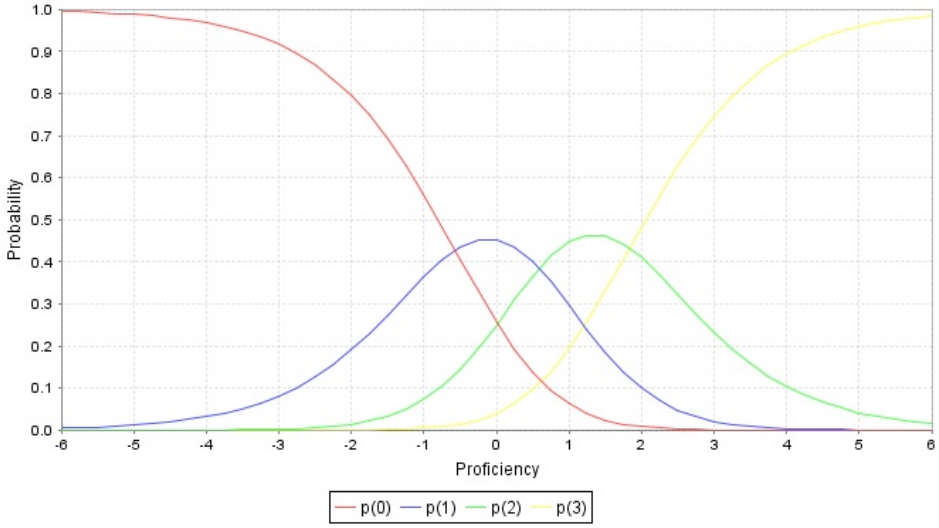
Q19. My teacher cares about me.
Item Set : base
Item : Q19



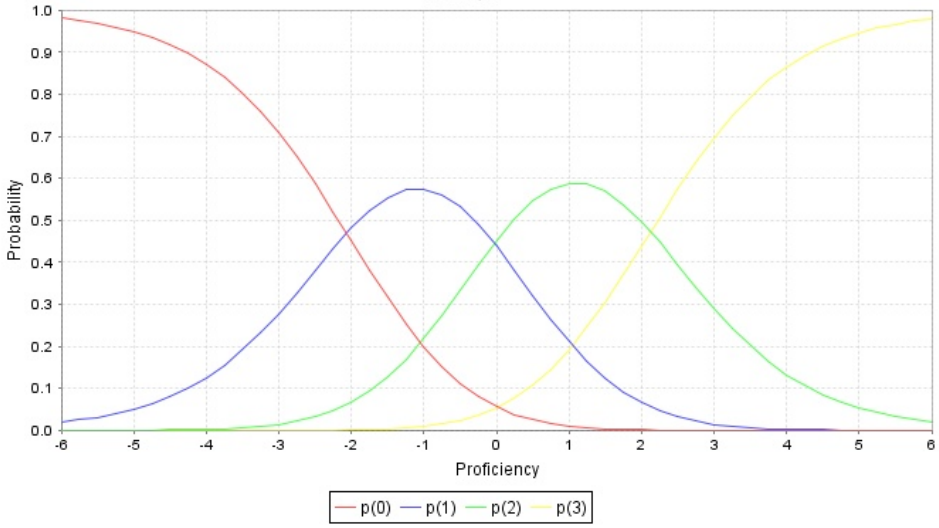
Q20. My teacher pays attention to what all students are thinking and feeling.
Item Set : base
Item : Q20



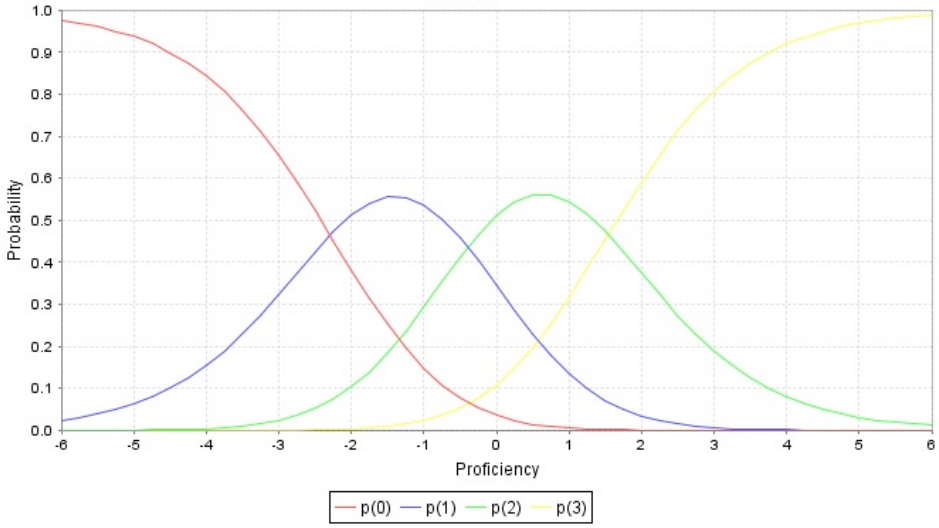
Q21. My teacher would notice if something was bothering me.
Item Set : base
Item : Q21



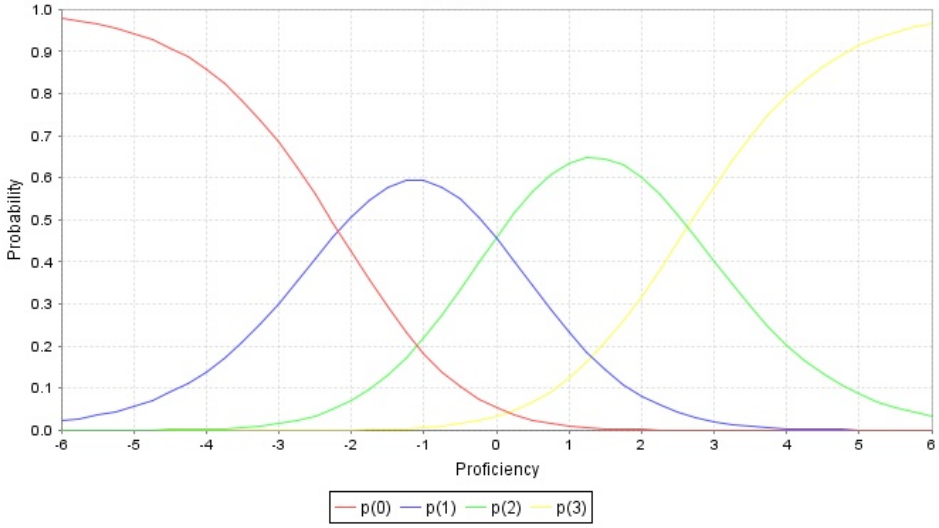
Q22. Our class stays busy and does not waste time.
Item Set : base
Item : Q22



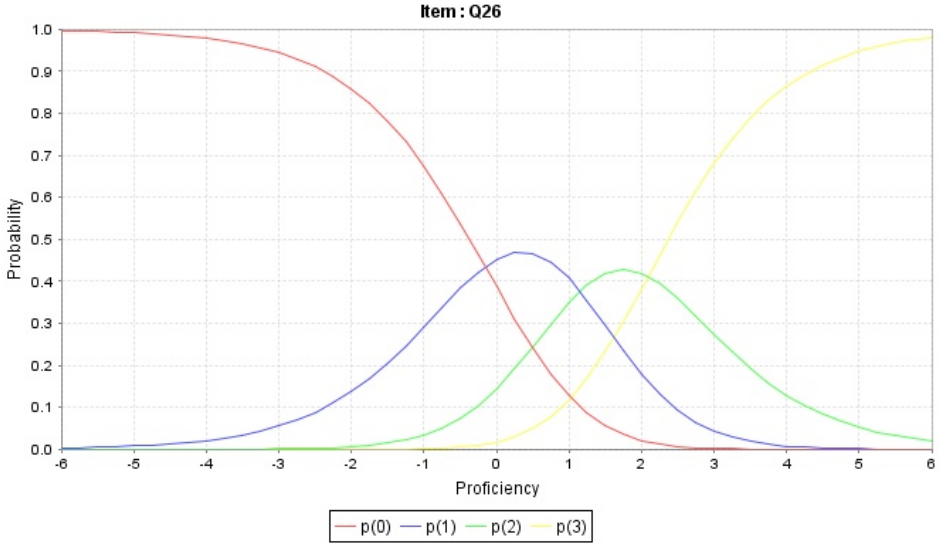
Q23. Students in this class treat the teacher with respect.
Item Set : base
Item : Q23



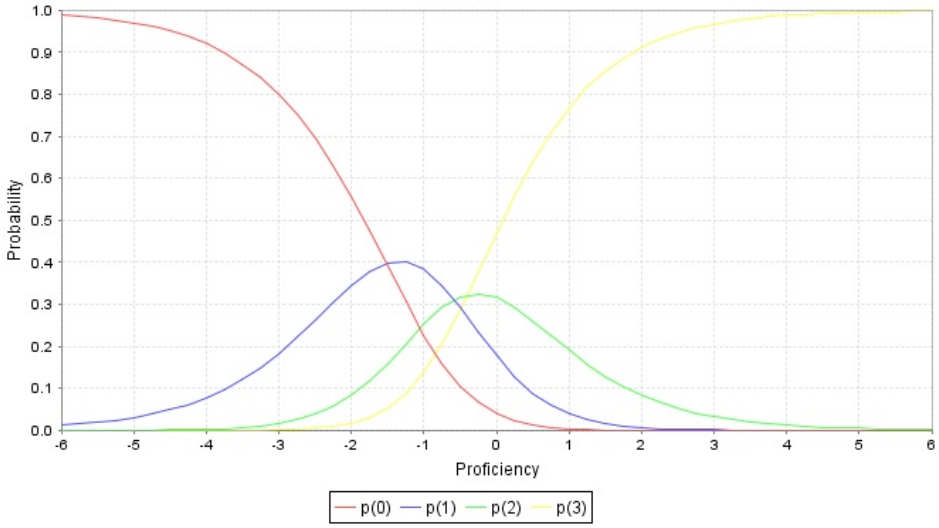
Q24. The students behave the way my teacher wants them to.
Item Set : base
Item : Q24



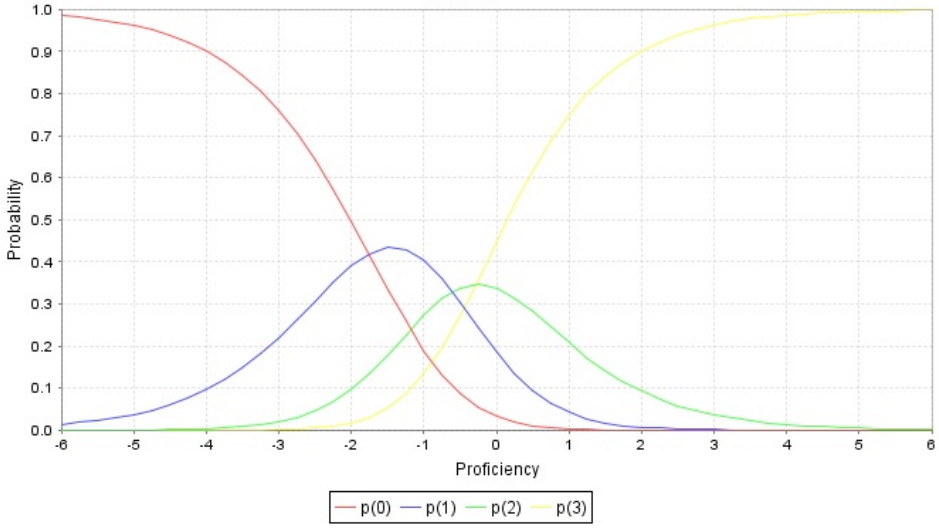
Q26. Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background.
Item Set : base
Item : Q26



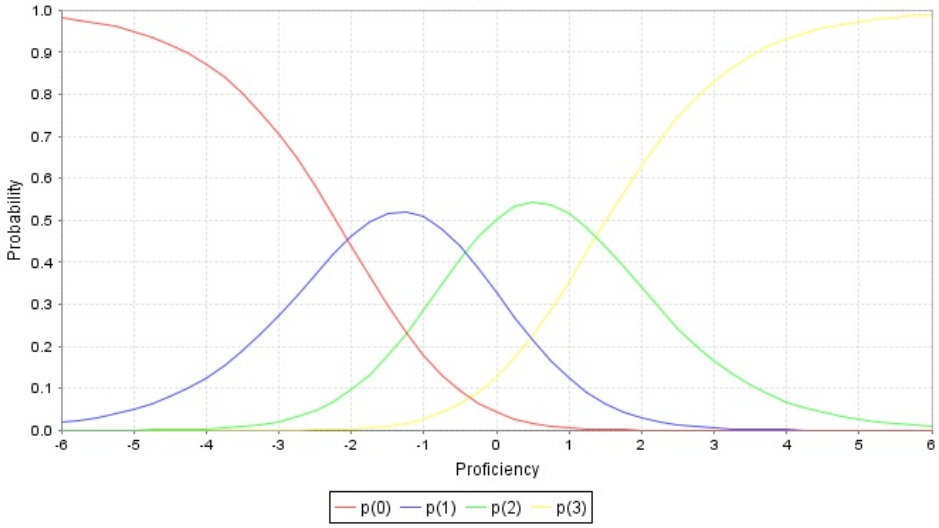
Q27. My teacher respects my cultural background.
Item Set : base
Item : Q27



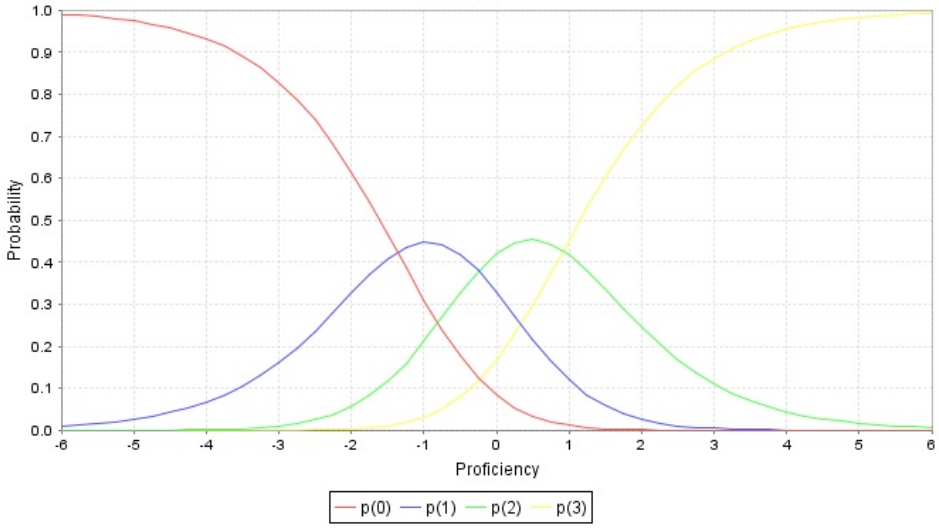
Q28. My teacher respects me as an individual.
Item Set : base
Item : Q28



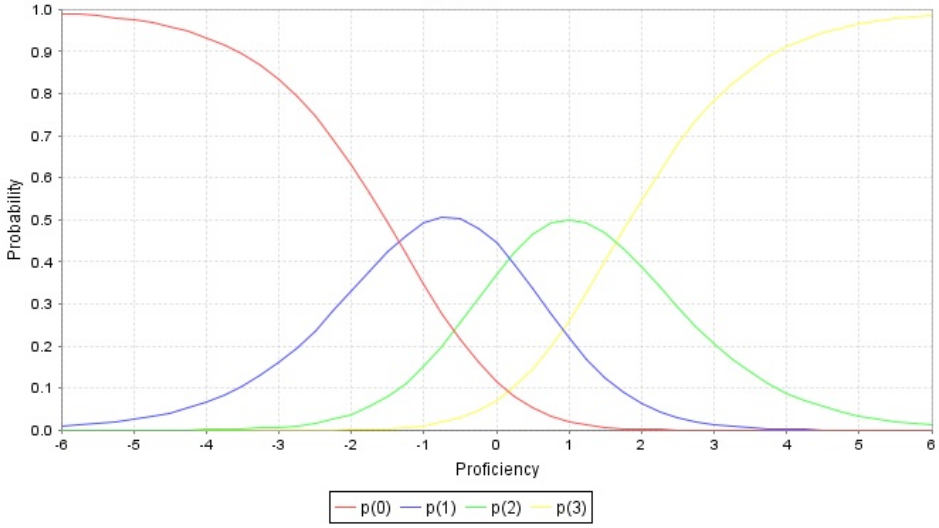
Q29. Students in this class respect each other's differences.
Item Set : base
Item : Q29



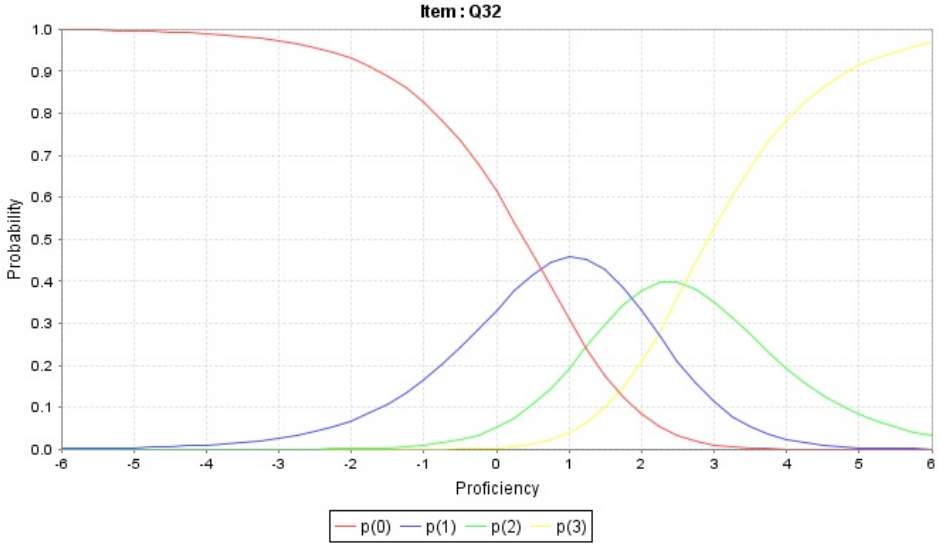
Q30. In this class, I feel like I fit in.
Item Set : base
Item : Q30



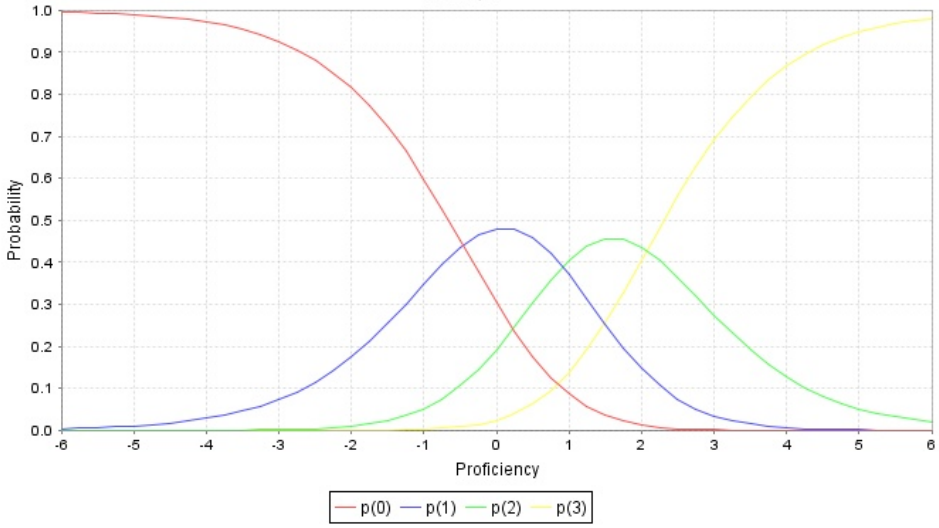
Q31. I feel like an important part of this classroom community.
Item Set : base
Item : Q31



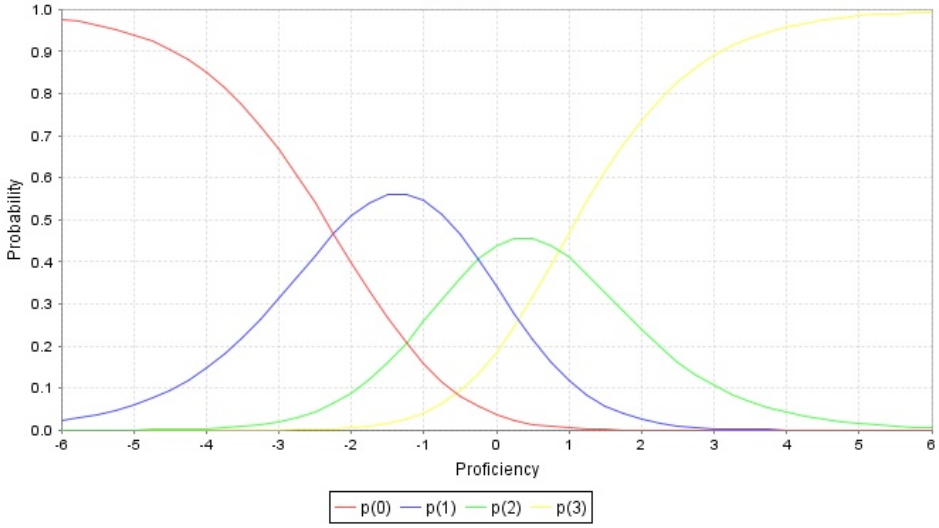
Q32. My teacher knows what my life is like outside of school.
Item Set : base
Item : Q32

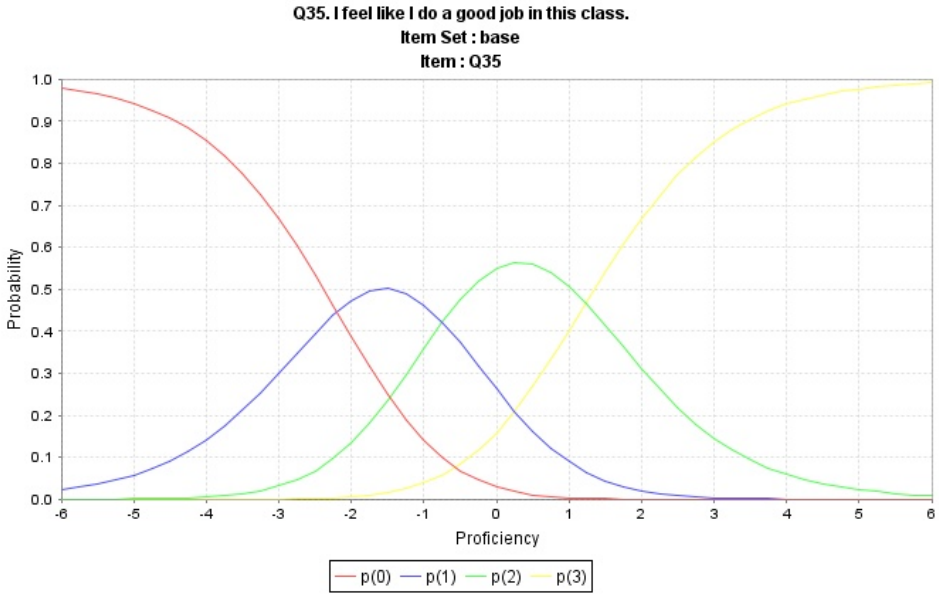


Q33. My teacher knows what is important to me.
Item Set : base
Item : Q33



Q34. I ask for help when I need it.
Item Set : base
Item : Q34




Q35. I feel like I do a good job in this class.
Item Set : base
Item : Q35
Probability
Proficiency
p(0)
p(1)
p(2)
p(3)

Appendix E: *Additional Teacher-Level Reliability Analyses, Grades 3-5*

Scale: STUDENT LEARNING

Table E1.*Teacher-Level Student Learning item statistics, grades 3-5*

	Mean	SD	Item-Total Correlation	α If Item Deleted
1. The schoolwork we do helps me learn.	3.2307	.29493	.792	.943
2. The schoolwork we do is interesting.	2.7880	.33506	.444	.950
3. What I learn in this class is useful to me in my real life.	3.0544	.32647	.736	.944
5. In this class, we learn a lot almost every day.	3.3159	.34498	.827	.942
6. My teacher makes sure that we think hard about things we read and write.	3.4122	.44686	.667	.946
7. When the work is too hard, my teacher helps me keep trying.	3.1994	.38005	.824	.942
8. In this class, it is more important to understand the lesson than to memorize the answers.	3.1590	.29777	.704	.945
9. My teacher uses a lot of different ways to explain things.	3.0869	.33882	.772	.943
10. My teacher knows when we understand the lesson & when we do not.	2.9397	.31226	.793	.943
11. Our classroom materials and supplies have a special place and things are easy to find.	3.2917	.33255	.560	.948
12. In this class, we learn to correct our mistakes.	3.3927	.31141	.815	.942
13. My teacher tells us what we are learning and why.	3.2540	.32451	.736	.944
15. My teacher asks questions to be sure we are following along.	3.1829	.33216	.753	.943
17. My teacher talks to me about my work to help me understand my mistakes.	3.0744	.38730	.848	.941
18. My teacher writes notes on my work that help me do better next time.	2.7195	.50814	.694	.947

Table E2. *Inter-item correlation matrix*

	Q1	Q2	Q3	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q15	Q17	Q18
Q1	1.000	.462	.705	.776	.571	.652	.597	.649	.604	.449	.660	.597	.620	.678	.522
Q2	.462	1.000	.390	.331	.017	.521	.286	.435	.503	.421	.324	.486	.260	.430	.205
Q3	.705	.390	1.000	.685	.522	.593	.527	.606	.583	.450	.641	.587	.553	.641	.504
Q5.	.776	.331	.685	1.000	.714	.631	.647	.685	.609	.459	.732	.620	.686	.679	.602
Q6	.571	.017	.522	.714	1.000	.497	.568	.545	.473	.299	.677	.443	.626	.583	.650
Q7	.652	.521	.593	.631	.497	1.000	.623	.722	.760	.517	.668	.641	.653	.813	.578
Q8.	.597	.286	.527	.647	.568	.623	1.000	.524	.614	.476	.649	.531	.540	.597	.461
Q9	.649	.435	.606	.685	.545	.722	.524	1.000	.674	.363	.614	.630	.685	.666	.532
Q10	.604	.503	.583	.609	.473	.760	.614	.674	1.000	.514	.644	.656	.643	.724	.553
Q11.	.449	.421	.450	.459	.299	.517	.476	.363	.514	1.000	.572	.452	.304	.519	.399
Q12	.660	.324	.641	.732	.677	.668	.649	.614	.644	.572	1.000	.553	.605	.735	.622
Q13	.597	.486	.587	.620	.443	.641	.531	.630	.656	.452	.553	1.000	.572	.638	.557
Q15	.620	.260	.553	.686	.626	.653	.540	.685	.643	.304	.605	.572	1.000	.689	.611
Q17	.678	.430	.641	.679	.583	.813	.597	.666	.724	.519	.735	.638	.689	1.000	.657
Q18	.522	.205	.504	.602	.650	.578	.461	.532	.553	.399	.622	.557	.611	.657	1.000

Scale: STUDENT-CENTERED ENVIRONMENT

Table E3. *Teacher-Level Student-Centered Environment item statistics, grades 3-5*

	Mean	SD	Item-Total Correlation	α If Item Deleted
4. My teacher knows what makes me excited about learning.	2.6588	.38540	.849	.927
14. My teacher wants us to share what we think.	2.9179	.36414	.657	.936
16. Students feel comfortable sharing their ideas in this class.	2.8183	.30792	.719	.934
19. My teacher talks about things we learn in other classes, subjects, and years.	2.5426	.33457	.620	.937
21. If I am sad or angry, my teacher helps me feel better.	2.9414	.49522	.851	.928
22. My teacher would notice if something was bothering me.	2.8137	.38582	.844	.927
28. The people we learn and read about in this class are like me.	1.9588	.27432	.590	.939
29. My teacher teaches us to respect people's differences.	3.2078	.36924	.736	.932
32. My teacher knows what my life is like outside of school.	2.0940	.35959	.780	.930
33. My teacher knows what is important to me.	2.6550	.39907	.878	.925

Table E4. *Inter-item correlation matrix*

	Q4	Q14	Q16	Q19	Q21	Q22	Q28	Q29	Q32	Q33
Q4	1.000	.530	.665	.543	.799	.785	.551	.666	.692	.804
Q14	.530	1.000	.577	.539	.552	.551	.397	.542	.566	.577
Q16	.665	.577	1.000	.434	.633	.628	.458	.526	.638	.654
Q19	.543	.539	.434	1.000	.520	.529	.502	.528	.479	.517
Q21	.799	.552	.633	.520	1.000	.832	.487	.693	.680	.834
Q22	.785	.551	.628	.529	.832	1.000	.482	.681	.674	.806
Q28	.551	.397	.458	.502	.487	.482	1.000	.421	.519	.560
Q29	.666	.542	.526	.528	.693	.681	.421	1.000	.575	.665
Q32	.692	.566	.638	.479	.680	.674	.519	.575	1.000	.794
Q33	.804	.577	.654	.517	.834	.806	.560	.665	.794	1.000

Scale: CLASSROOM COMMUNITY

Table E5. *Teacher-Level Community Classroom item statistics, grades 3-5*

	Mean	SD	Item-Total Correlation	α If Item Deleted
20. My teacher cares about me.	3.4440	.40742	.777	.873
30. In this class, I feel like I fit in.	2.9873	.34672	.822	.856
31. I feel like an important part of my classroom community.	2.8862	.34197	.802	.861
34. I ask for help when I need it.	3.2516	.26136	.628	.898
35. I feel like I do a good job in this class.	3.2375	.26275	.763	.875

Table E6. *Inter-item correlation matrix*

	Q20	Q30	Q31	Q34	Q35
Q20	1.000	.723	.681	.604	.640
Q30	.723	1.000	.797	.522	.714
Q31	.681	.797	1.000	.531	.700
Q34	.604	.522	.531	1.000	.562
Q35	.640	.714	.700	.562	1.000

Scale: CLASSROOM MANAGEMENT

Table E7. *Teacher-Level Classroom Management Item statistics, grades 3-5*

	Mean	SD	Item-Total Correlation	α If Item Deleted
23. Our class stays busy and does not waste time.	2.7380	.32374	.707	.892
24. Students in my class are respectful to our teacher.	2.9029	.36677	.797	.864
25. My classmates behave the way my teacher wants them to.	2.6377	.31104	.845	.842
26. All of the kids in my class know what they are supposed to be doing and learning.	3.0020	.26406	.779	.872

Table E8. *Inter-item correlation matrix*

	Q23	Q24	Q25	Q26
Q23	1.000	.629	.678	.634
Q24	.629	1.000	.794	.708
Q25	.678	.794	1.000	.751
Q26	.634	.708	.751	1.000

Appendix F: *Additional Teacher-Level Reliability Analyses, Grades 6-12*

Scale: STUDENT LEARNING

Table F1. *Teacher-Level Student Learning Item statistics, grades 6-12*

	Mean	SD	Item-Total Correlation	α if Item Deleted
1. My teacher makes learning enjoyable.	2.8806	.49786	.890	.965
2. What I learn in this class is useful to me in my real life.	2.9125	.39405	.767	.967
3. My teacher teaches things that are important to me.	2.8468	.39393	.861	.966
4. My teacher knows the things that make me excited about learning.	2.5351	.47822	.903	.965
5. In this class, we learn a lot every day.	3.0229	.38827	.704	.969
6. In this class, it is more important to understand the lesson than to memorize the answers.	3.1031	.33773	.682	.969
7. When the work is too hard, my teacher helps me keep trying.	3.1379	.42439	.901	.965
8. My teacher accepts nothing less than my best effort.	3.3069	.30653	.752	.968
9. My teacher knows when we understand the lesson and when we do not.	2.9759	.39945	.909	.965
10. If I don't understand something, my teacher explains it a different way.	2.9833	.44343	.916	.965
11. My teacher explains difficult things clearly.	2.9509	.40615	.918	.965
15. In this class, we have a say in what we learn and do.	2.4531	.42804	.756	.968
16. My teacher talks to me about my work to help me understand my mistakes.	2.9203	.41196	.866	.966
17. My teacher writes notes on my work that help me improve.	2.6877	.49889	.587	.972
18. When we study a topic, my teacher makes connections to other subjects or classes.	2.6465	.39794	.797	.967

Table F2. *Inter-item correlation matrix*

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q15	Q16	Q17	Q18
Q1	1.000	.678	.800	.931	.607	.589	.823	.698	.844	.836	.876	.778	.757	.459	.716
Q2	.678	1.000	.835	.680	.590	.573	.664	.617	.691	.692	.694	.604	.653	.463	.654
Q3	.800	.835	1.000	.813	.640	.585	.763	.690	.785	.777	.794	.722	.746	.458	.697
Q4	.931	.680	.813	1.000	.605	.594	.824	.683	.844	.844	.865	.800	.785	.499	.743
Q5	.607	.590	.640	.605	1.000	.617	.658	.609	.661	.664	.671	.357	.608	.519	.585
Q6	.589	.573	.585	.594	.617	1.000	.660	.550	.643	.643	.639	.464	.614	.442	.517
Q7	.823	.664	.763	.824	.658	.660	1.000	.700	.863	.885	.851	.703	.865	.565	.697
Q8	.698	.617	.690	.683	.609	.550	.700	1.000	.745	.671	.708	.526	.698	.441	.573
Q9	.844	.691	.785	.844	.661	.643	.863	.745	1.000	.891	.889	.711	.814	.510	.723
Q10	.836	.692	.777	.844	.664	.643	.885	.671	.891	1.000	.901	.733	.813	.548	.771
Q11	.876	.694	.794	.865	.671	.639	.851	.708	.889	.901	1.000	.720	.788	.548	.746
Q15	.778	.604	.722	.800	.357	.464	.703	.526	.711	.733	.720	1.000	.691	.398	.663
Q16	.757	.653	.746	.785	.608	.614	.865	.698	.814	.813	.788	.691	1.000	.620	.675
Q17	.459	.463	.458	.499	.519	.442	.565	.441	.510	.548	.548	.398	.620	1.000	.586
Q18	.716	.654	.697	.743	.585	.517	.697	.573	.723	.771	.746	.663	.675	.586	1.000

Scale: STUDENT-CENTERED ENVIRONMENT

Table F3. *Teacher-Level Student Centered Environment item statistics, grades 6-12*

	Mean	SD	Item-Total Correlation	α if Item Deleted
12. My classroom is organized and I know where to find what I need.	3.3677	.31031	.659	.965
13. Students feel comfortable sharing their ideas in this class.	3.1486	.37130	.865	.950
14. My teacher respects my opinions and suggestions.	3.2486	.39637	.923	.945
19. My teacher cares about me.	3.1843	.42961	.903	.947
20. My teacher pays attention to what all students are thinking and feeling.	3.0028	.43368	.907	.947
27. My teacher respects my cultural background.	3.4567	.31335	.822	.954
28. My teacher respects me as an individual.	3.4529	.34429	.922	.946

Table F4. *Inter-item correlation matrix*

	Q12	Q13	Q14	Q19	Q20	Q27	Q28
Q12	1.000	.613	.627	.596	.608	.577	.646
Q13	.613	1.000	.844	.799	.822	.746	.815
Q14	.627	.844	1.000	.860	.876	.807	.891
Q19	.596	.799	.860	1.000	.922	.749	.865
Q20	.608	.822	.876	.922	1.000	.730	.850
Q27	.577	.746	.807	.749	.730	1.000	.844
Q28	.646	.815	.891	.865	.850	.844	1.000

Scale: CLASSROOM COMMUNITY

Table F5. *Teacher- Level Classroom Community item statistics, grades 6-12*

	Mean	SD	Item-Total Correlation	α If Item Deleted
21. My teacher would notice if something was bothering me.	2.6824	.46059	.852	.922
26. Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural background.	2.4208	.36522	.624	.937
30. In this class, I feel like I fit in.	3.1156	.31222	.793	.926
31. I feel like an important part of this classroom community.	2.8812	.33605	.885	.919
32. My teacher knows what my life is like outside of school.	2.0690	.40874	.820	.923
33. My teacher knows what is important to me.	2.5558	.42698	.900	.917
34. I ask for help when I need it.	3.1826	.27220	.694	.933
35. I feel like I do a good job in this class.	3.1788	.29013	.692	.933

Table F6. *Inter-item correlation matrix*

	Q21	Q26	Q30	Q31	Q32	Q33	Q34	Q35
Q21	1.000	.542	.663	.779	.802	.887	.604	.615
Q26	.542	1.000	.539	.600	.597	.600	.443	.400
Q30	.663	.539	1.000	.849	.645	.697	.656	.652
Q31	.779	.600	.849	1.000	.733	.811	.658	.709
Q32	.802	.597	.645	.733	1.000	.860	.534	.521
Q33	.887	.600	.697	.811	.860	1.000	.633	.631
Q34	.604	.443	.656	.658	.534	.633	1.000	.647
Q35	.615	.400	.652	.709	.521	.631	.647	1.000

Scale: CLASSROOM MANAGEMENT

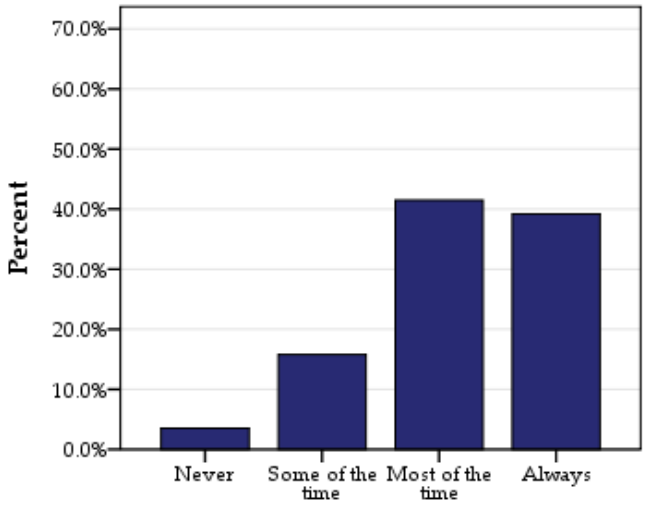
Table F7. *Teacher-Level Classroom Management item statistics, grades 6-12*

	Mean	Std. Deviation	Item-Total Correlation	α If Item Deleted
22. Our class stays busy and does not waste time.	2.8866	.34479	.764	.903
23. Students in this class treat the teacher with respect.	3.0682	.39067	.897	.856
24. The students behave the way my teacher wants them to.	2.8157	.35781	.907	.852
29. Students in this class respect each other's differences.	3.0832	.29075	.677	.931

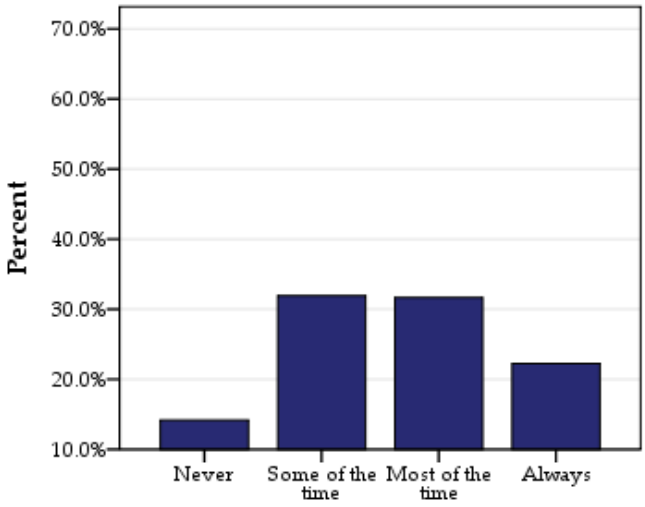
Table F8. *Inter-item correlation matrix*

	Q22	Q23	Q24	Q29
Q22	1.000	.761	.782	.531
Q23	.761	1.000	.911	.687
Q24	.782	.911	1.000	.679
Q29	.531	.687	.679	1.000

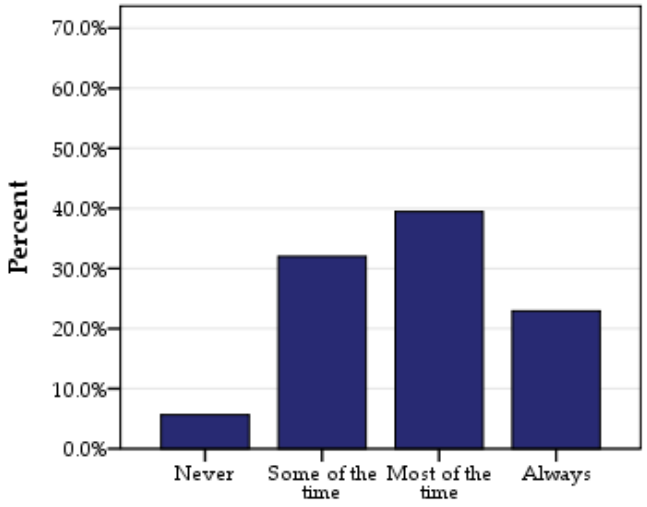
Appendix G. *Item Distributions – Grades 3-5*



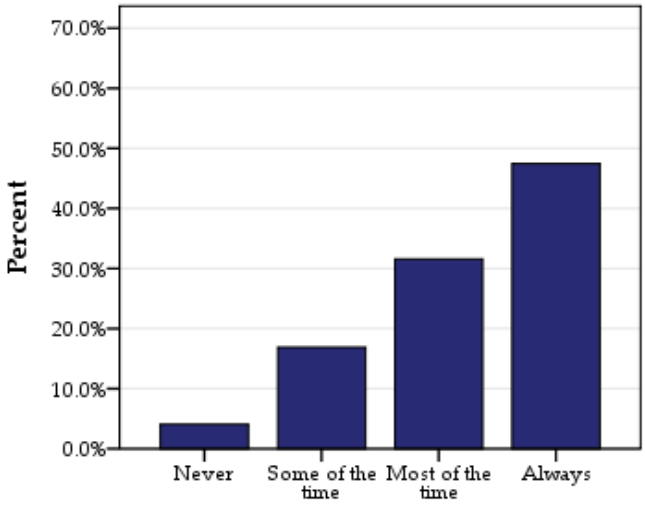
1. The schoolwork we do helps me learn.



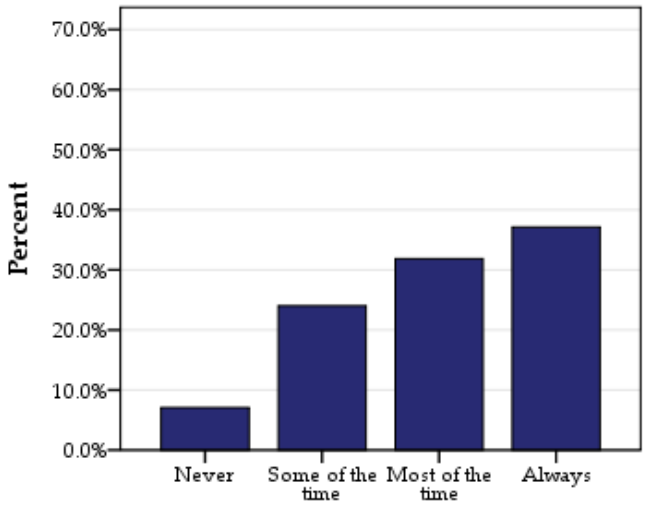
4. My teacher knows what makes me excited about learning.



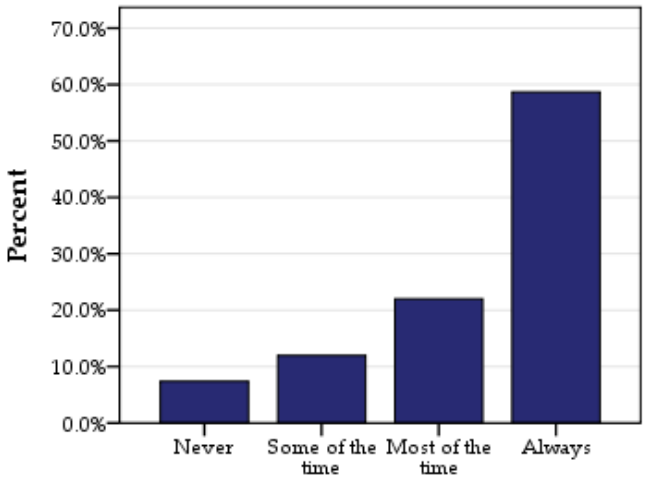
2. The schoolwork we do is interesting.



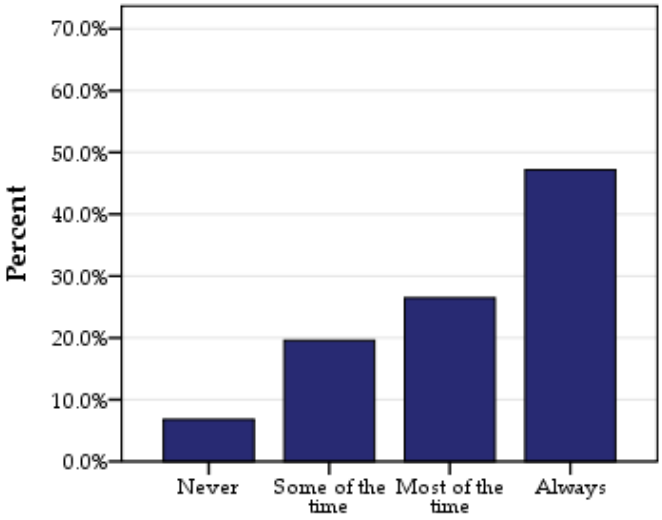
5. In this class, we learn a lot almost every day.



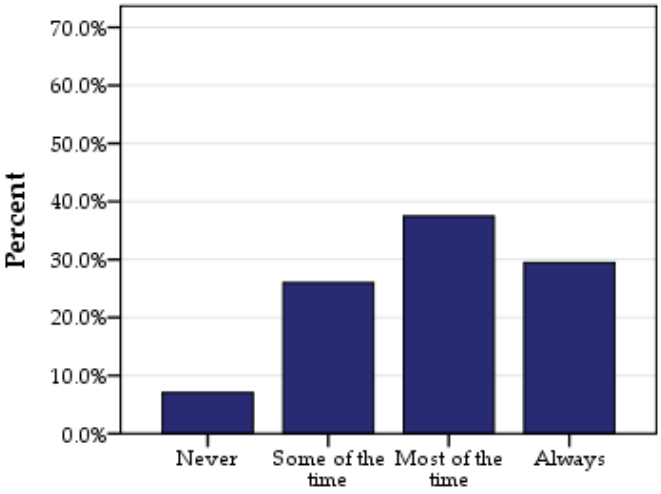
3. What I learn in this class is useful to me in my real life.



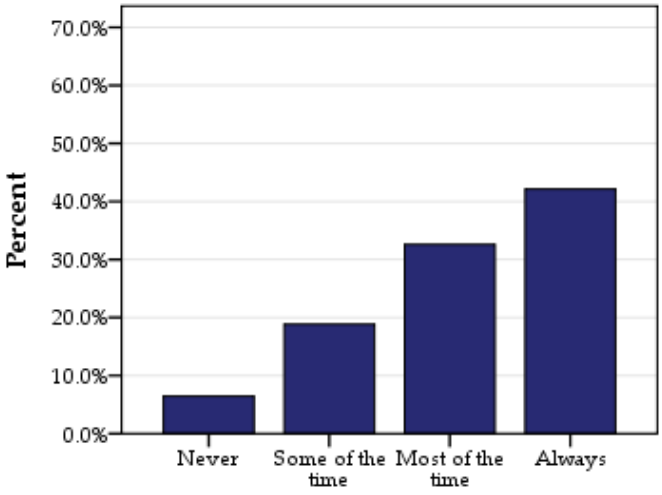
6. My teacher makes sure that we think hard about things we read and write.



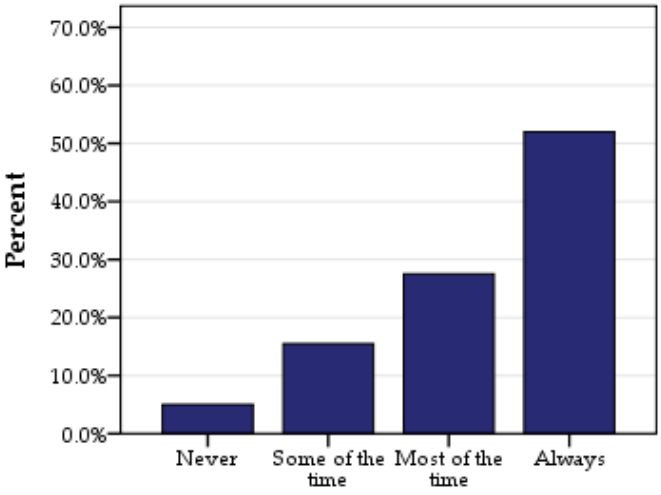
7. When the work is too hard, my teacher helps me keep trying.



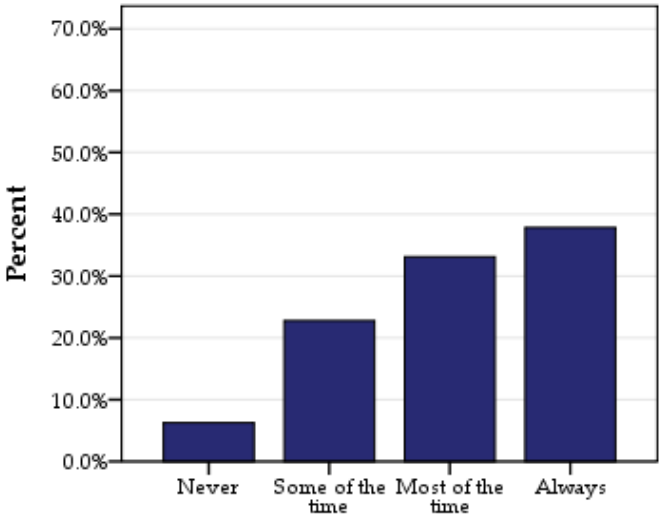
10. My teacher knows when we understand the lesson and when we do not.



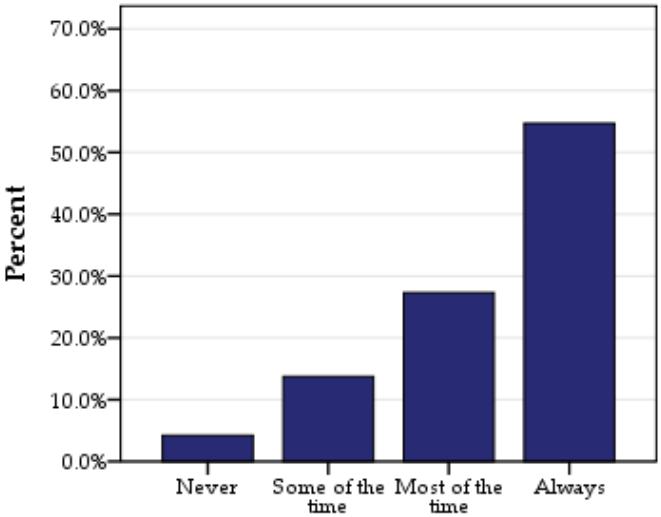
8. In this class, it is more important to understand the lesson than to memorize the answers.



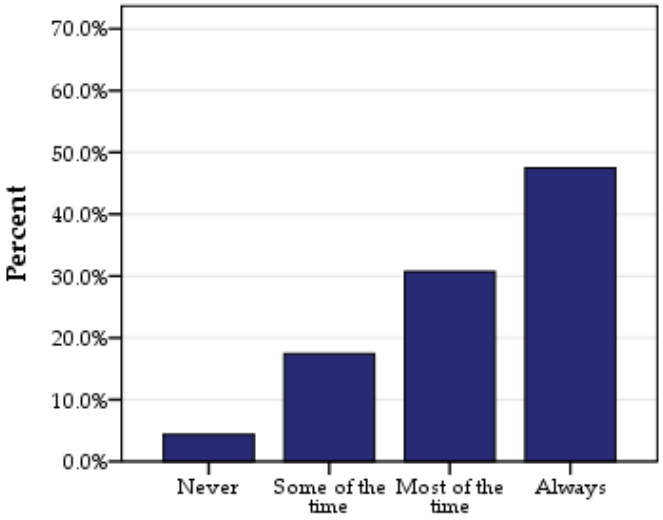
11. Our classroom materials and supplies have a special place and things are easy to find.



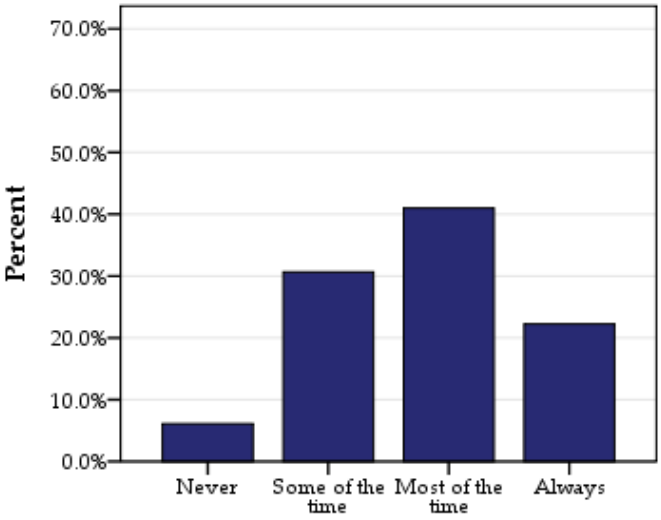
9. My teacher uses a lot of different ways to explain things.



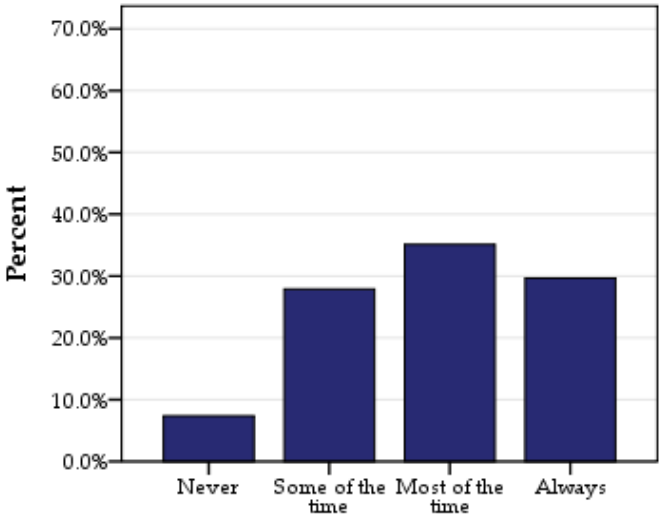
12. In this class, we learn to correct our mistakes.



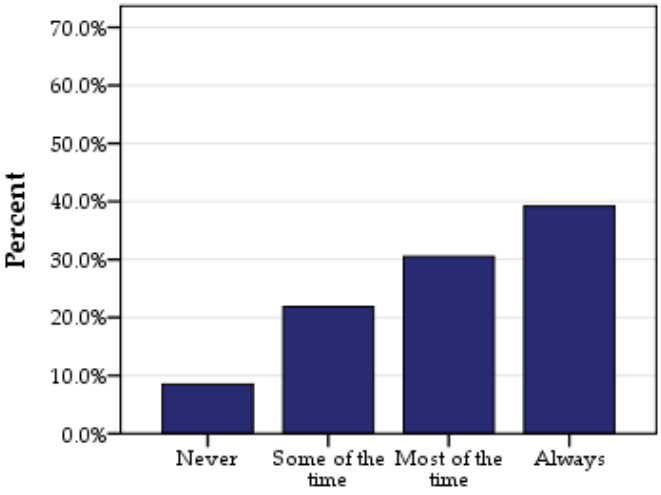
13. My teacher tells us what we are learning and why.



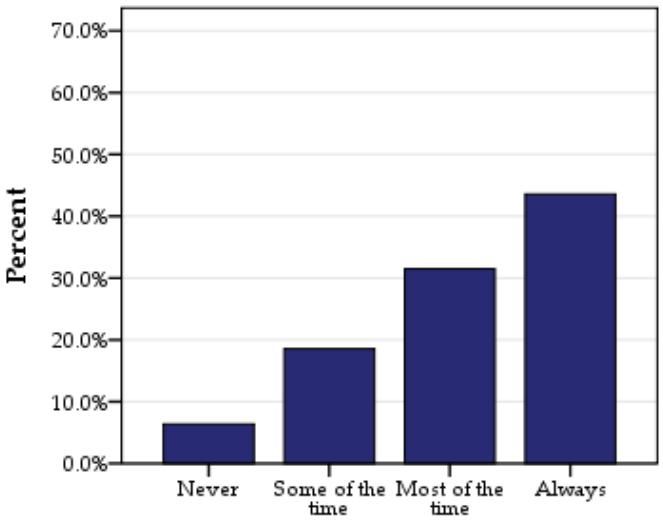
16. Students feel comfortable sharing their ideas in this class.



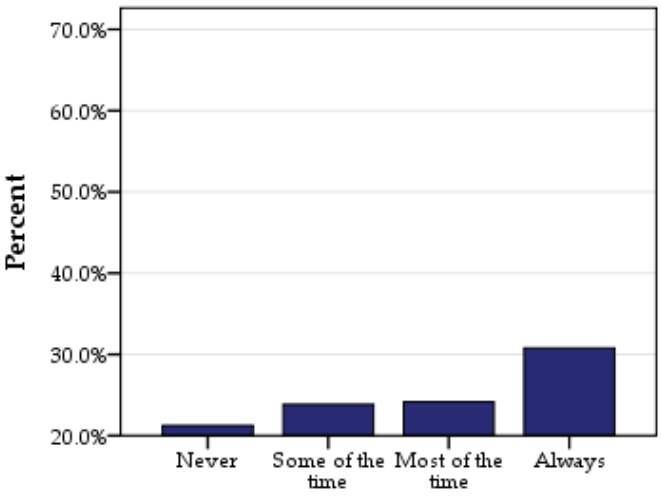
14. My teacher wants us to share what we think.



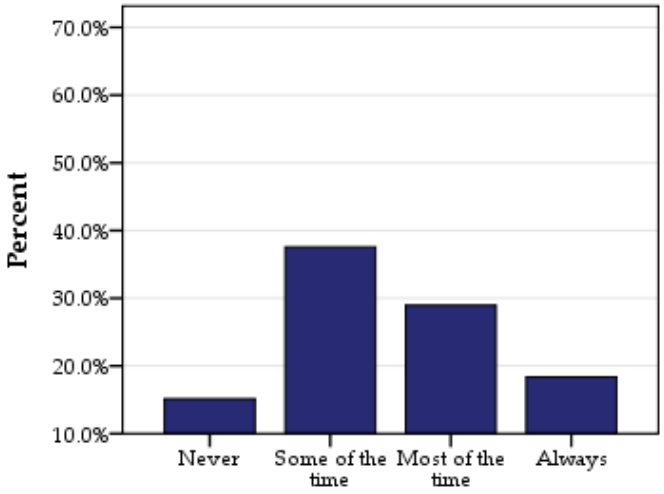
17. My teacher talks to me about my work to help me understand my mistakes.



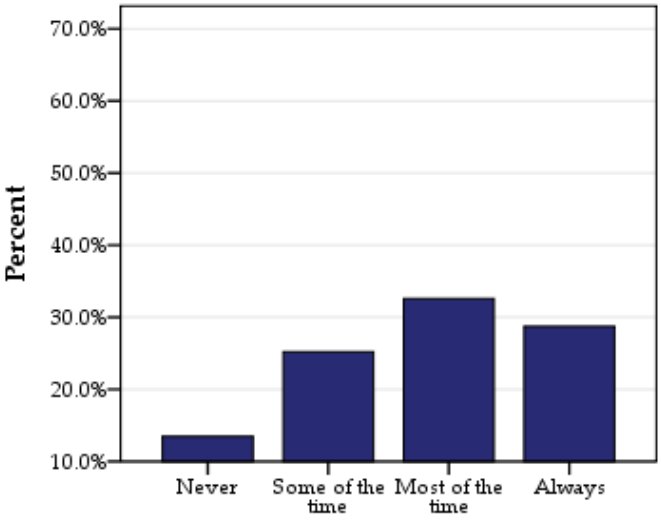
15. My teacher asks questions to be sure we are following along.



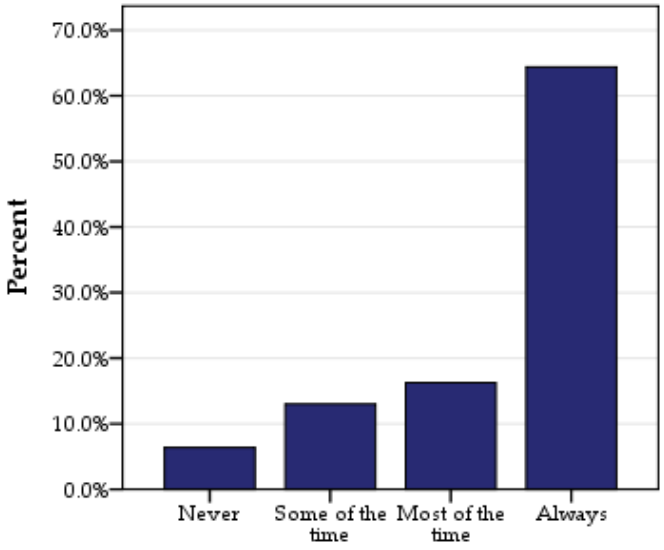
18. My teacher writes notes on my work that help me do better next time.



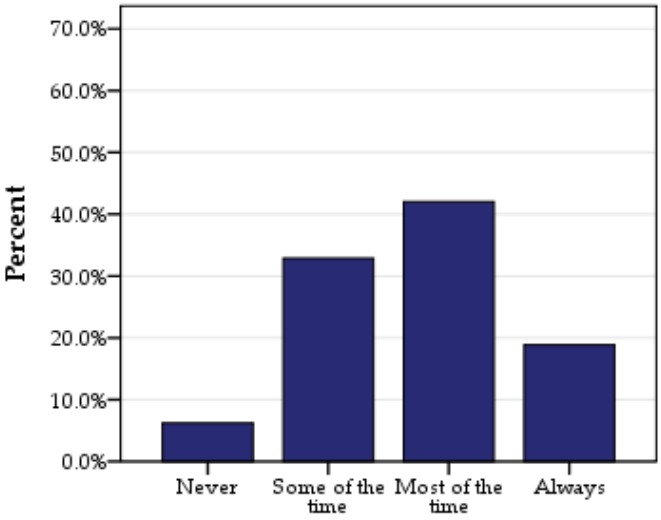
19. My teacher talks about things we learn in other classes, subjects, and years.



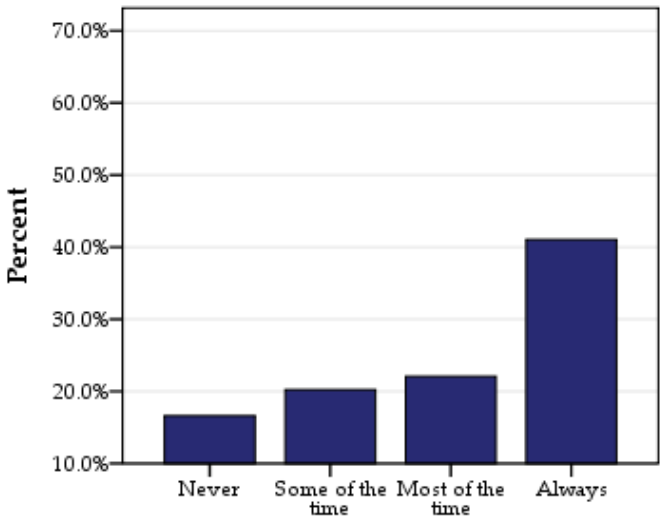
22. My teacher would notice if something was bothering me.



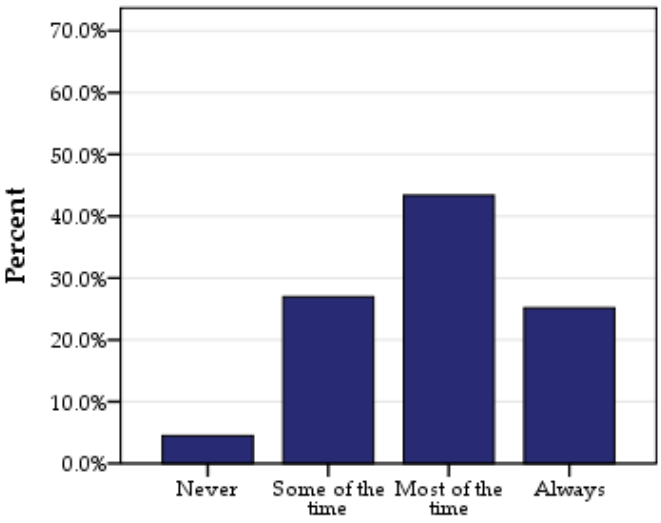
20. My teacher cares about me.



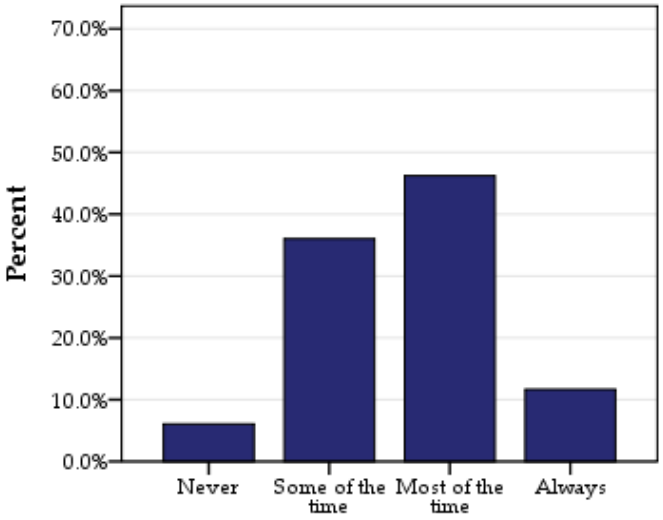
23. Our class stays busy and does not waste time.



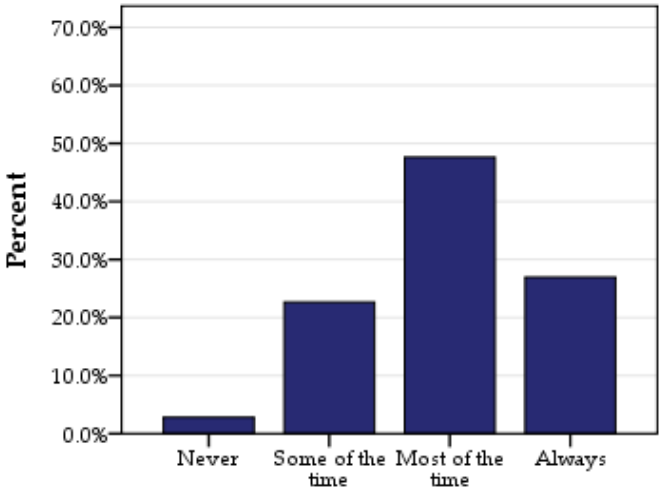
21. If I am sad or angry, my teacher helps me feel better.



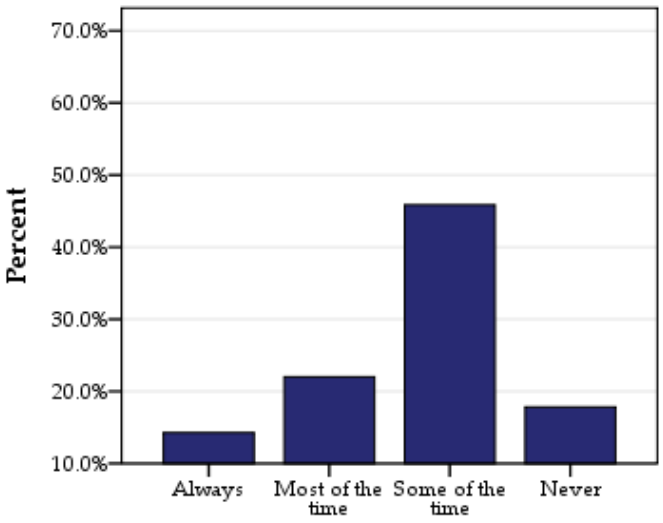
24. Students in my class are respectful to our teacher.



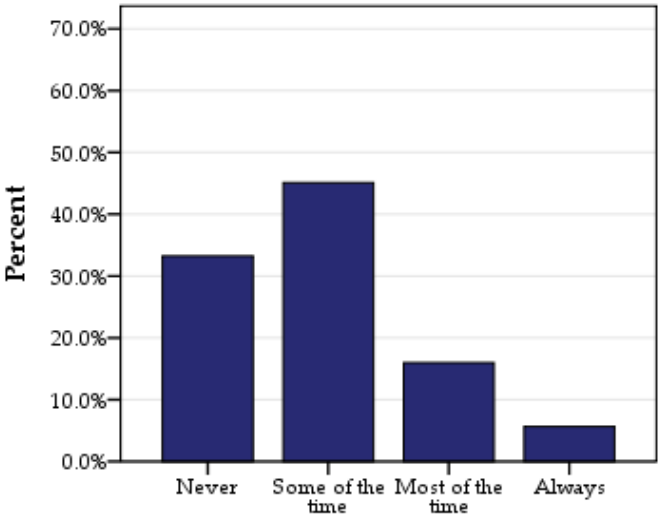
25. My classmates behave the way my teacher wants them to.



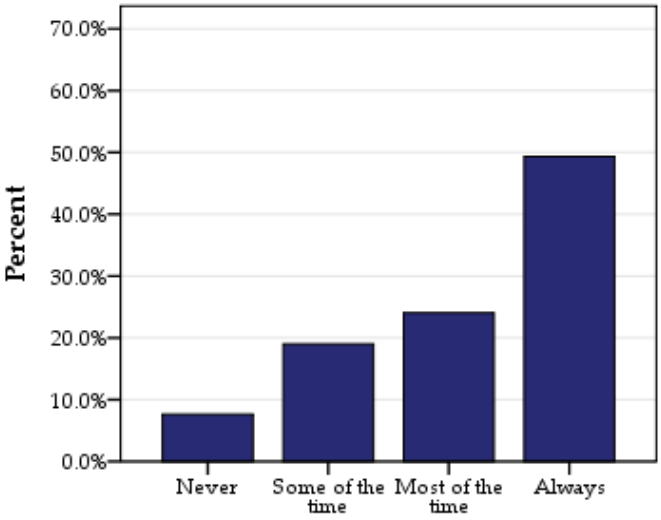
26. All of the kids in my class know what they are supposed to be doing and learning.



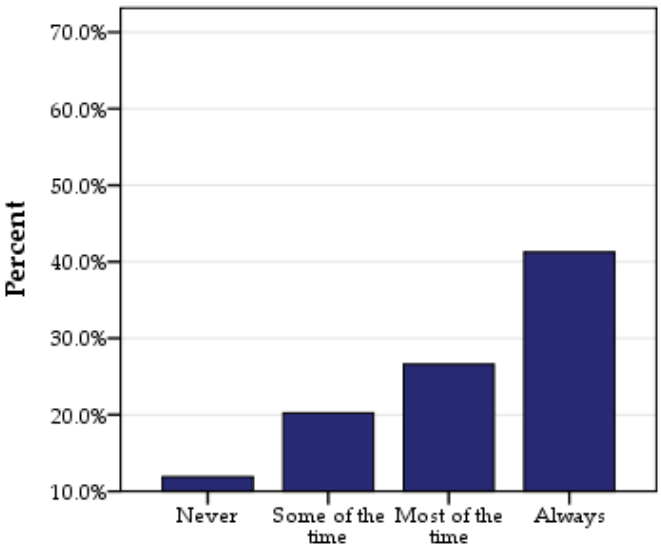
27. The way students behave in this class makes it hard to learn.



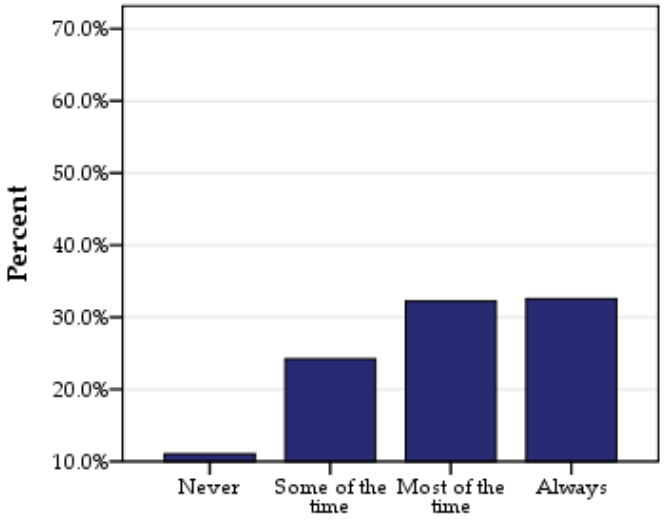
28. The people we learn and read about in this class are like me.



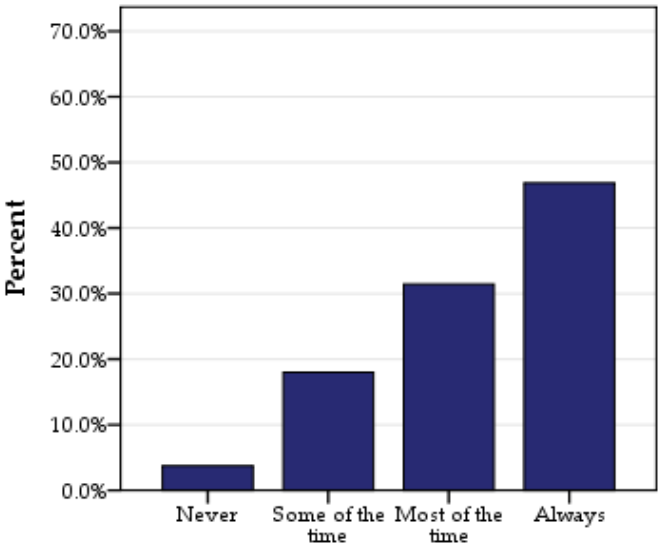
29. My teacher teaches us to respect people's differences.



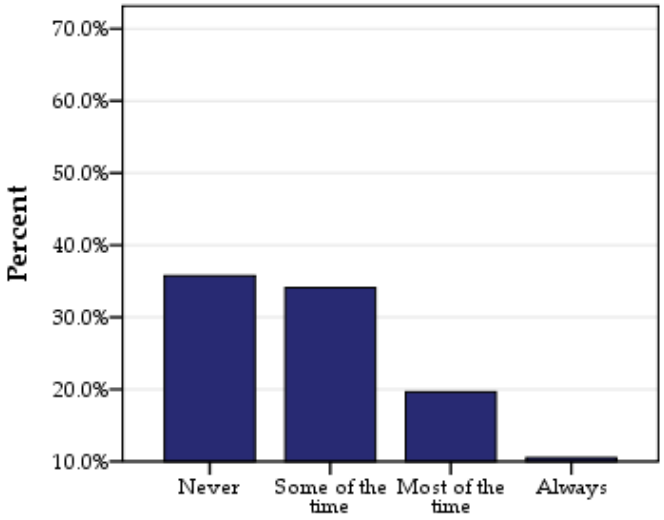
30. In this class, I feel like I fit in.



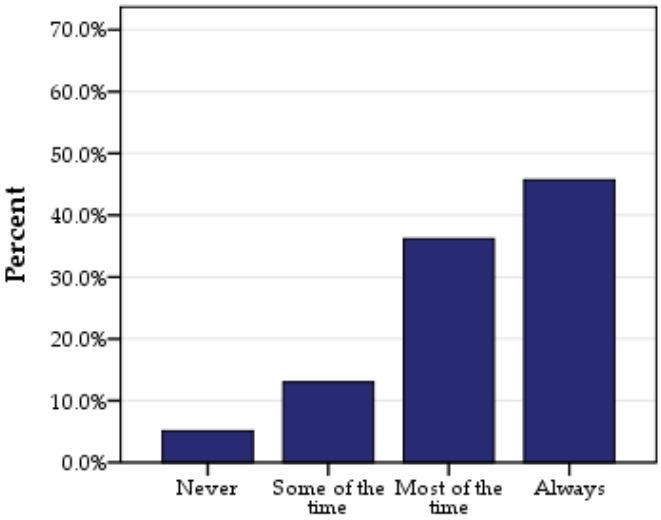
31. I feel like an important part of my classroom community.



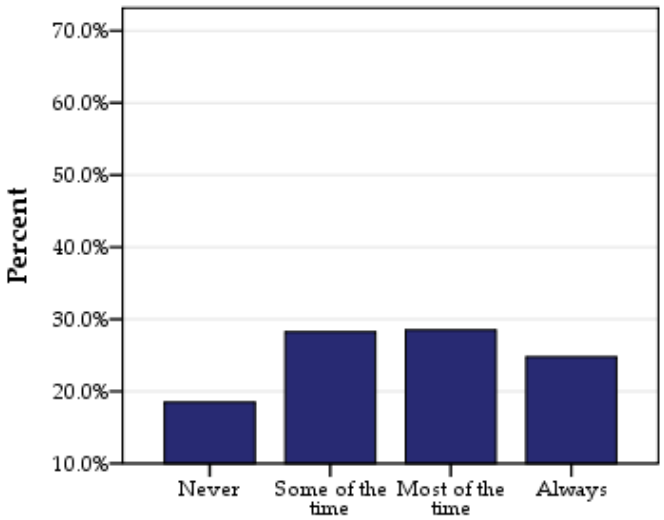
34. I ask for help when I need it.



32. My teacher knows what my life is like outside of school.

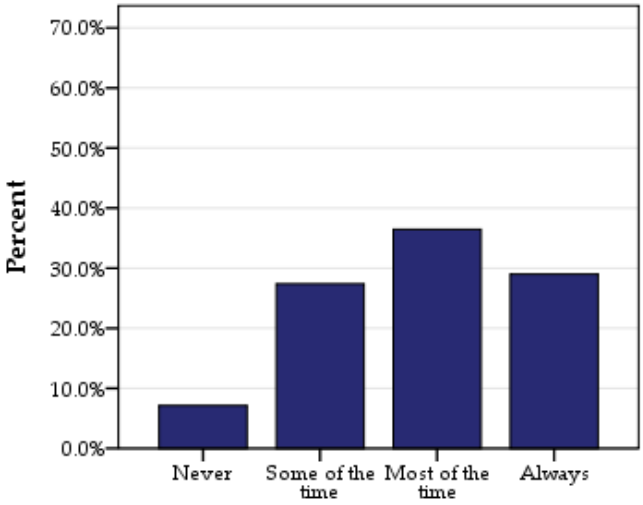


35. I feel like I do a good job in this class.

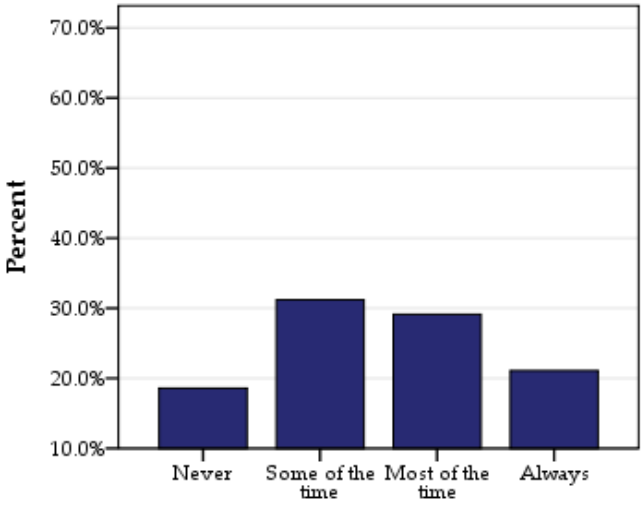


33. My teacher knows what is important to me.

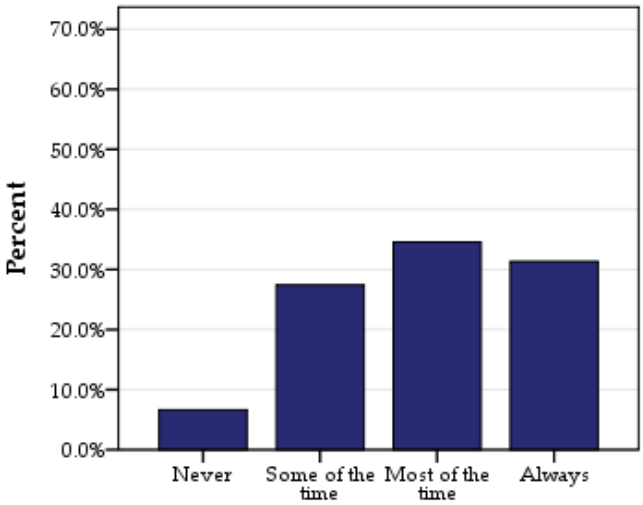
Appendix H. *Item Distributions – Grades 6-12*



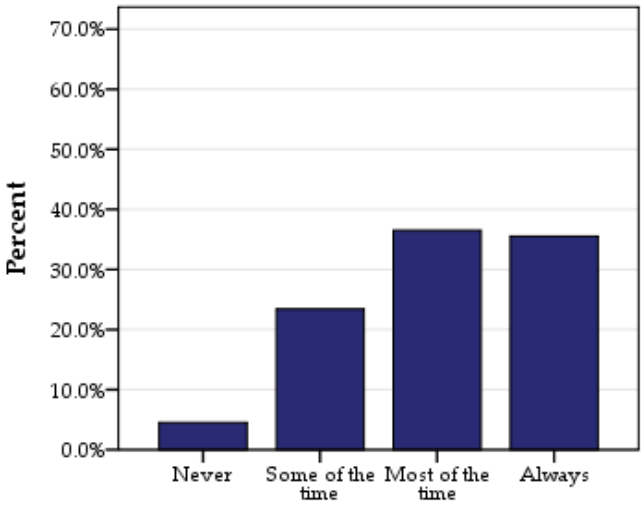
1. My teacher makes learning enjoyable.



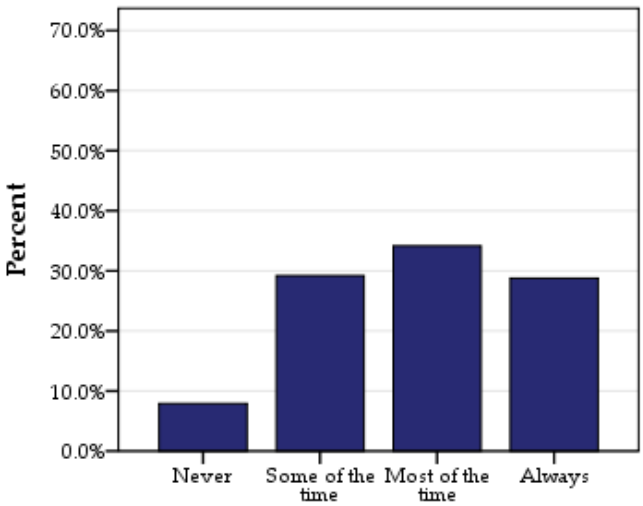
4. My teacher knows the things that make me excited about learning.



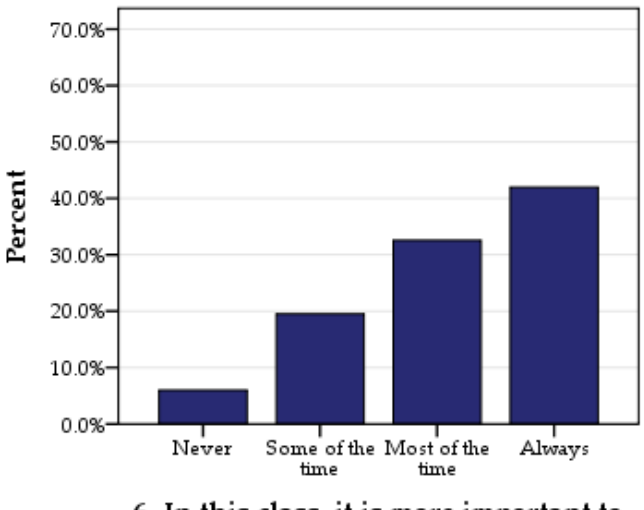
2. What I learn in this class is useful to me in my real life.



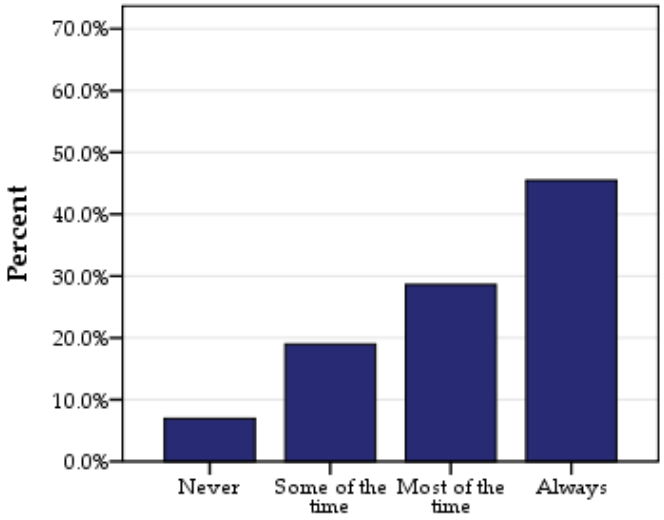
5. In this class, we learn a lot every day.



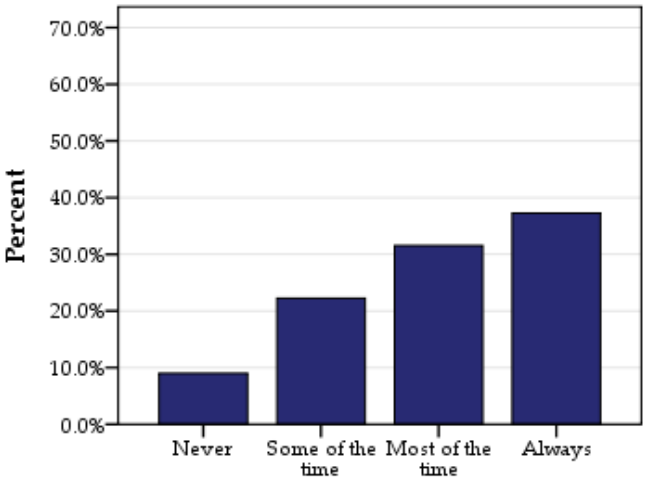
3. My teacher teaches things that are important to me.



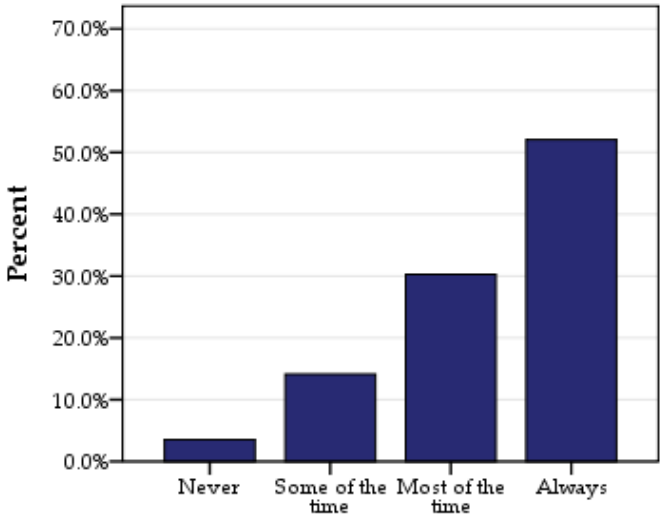
6. In this class, it is more important to understand the lesson than to memorize the answers.



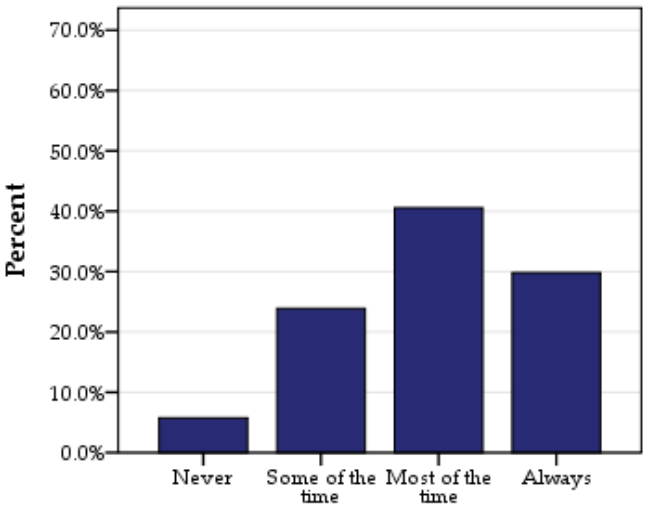
7. When the work is too hard, my teacher helps me keep trying.



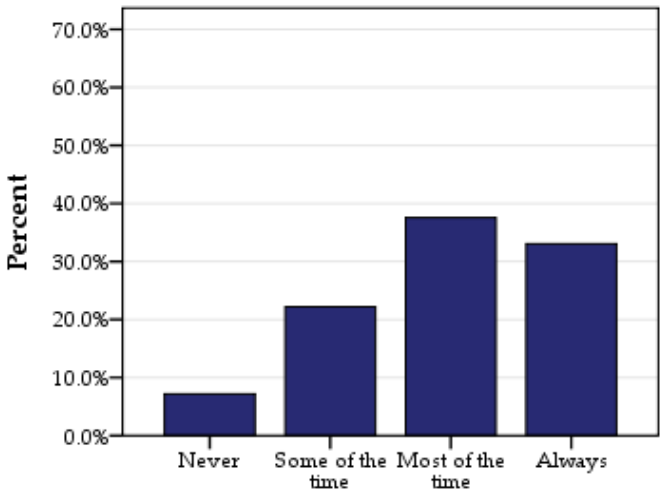
10. If I don't understand something, my teacher explains it a different way.



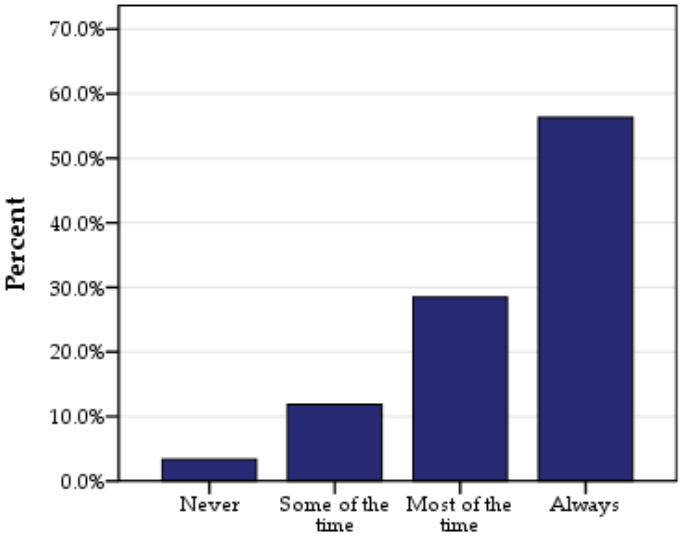
8. My teacher accepts nothing less than my best effort.



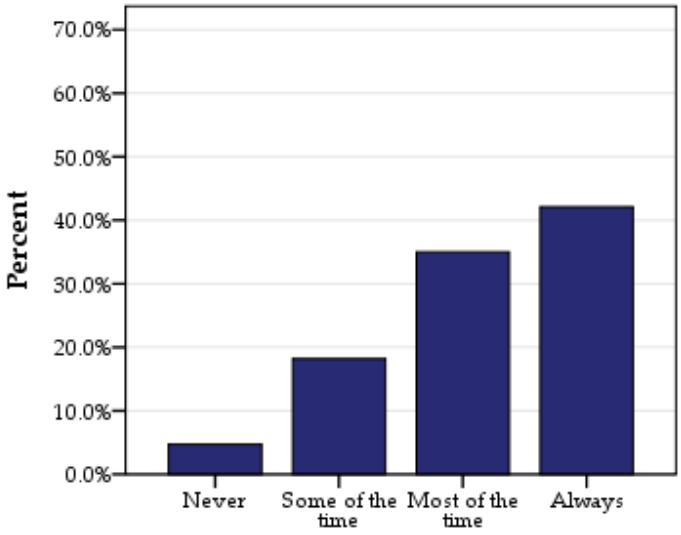
11. My teacher explains difficult things clearly.



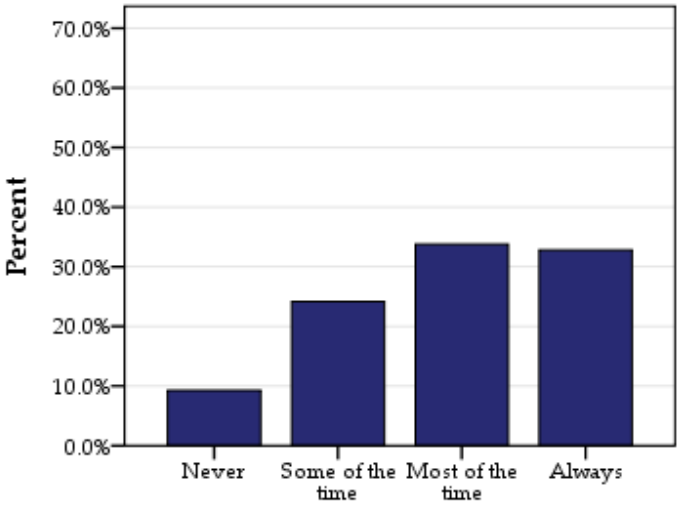
9. My teacher knows when we understand the lesson and when we do not.



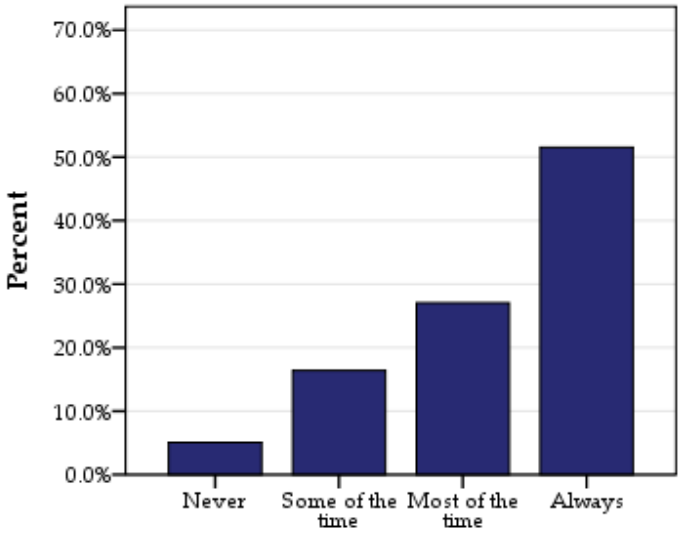
12. My classroom is organized and I know where to find what I need.



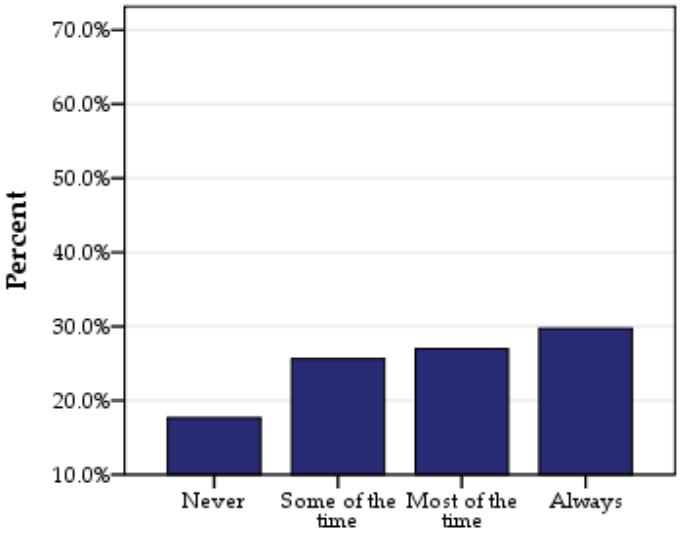
13. Students feel comfortable sharing their ideas in this class.



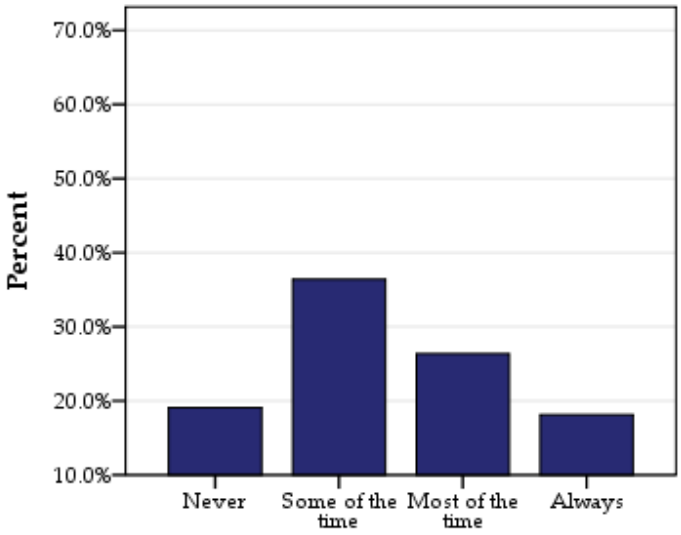
16. My teacher talks to me about my work to help me understand my mistakes.



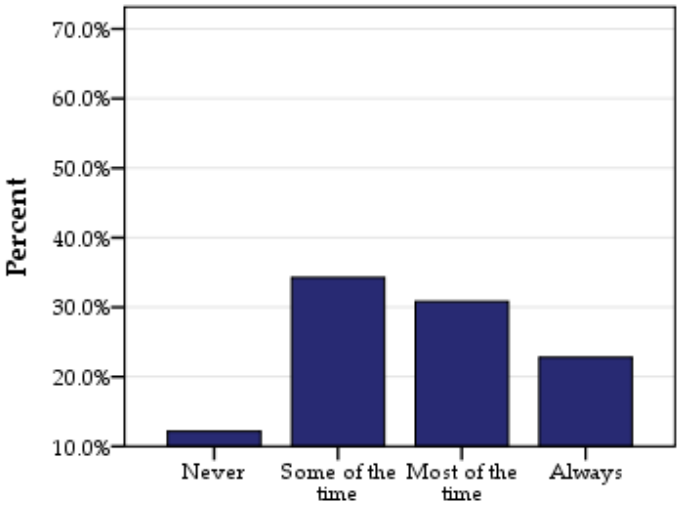
14. My teacher respects my opinions and suggestions.



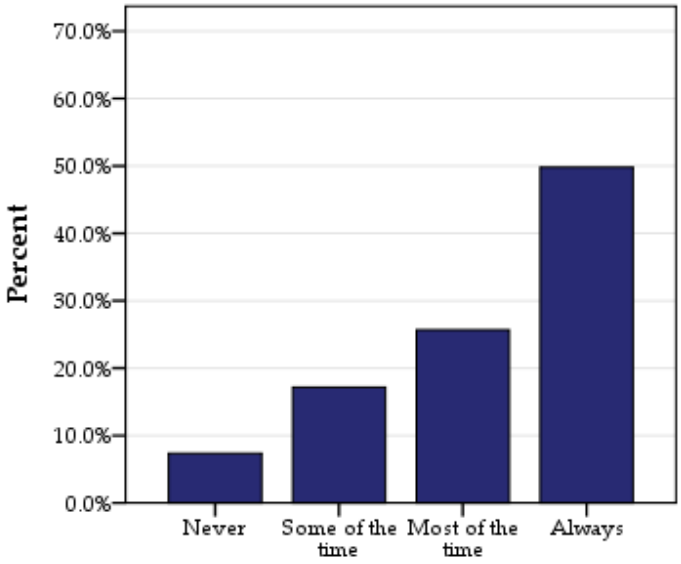
17. My teacher writes notes on my work that help me improve.



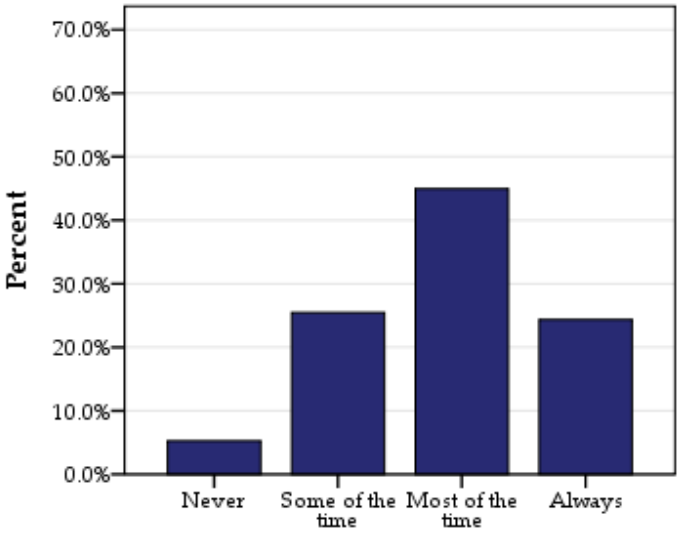
15. In this class, we have a say in what we learn and do.



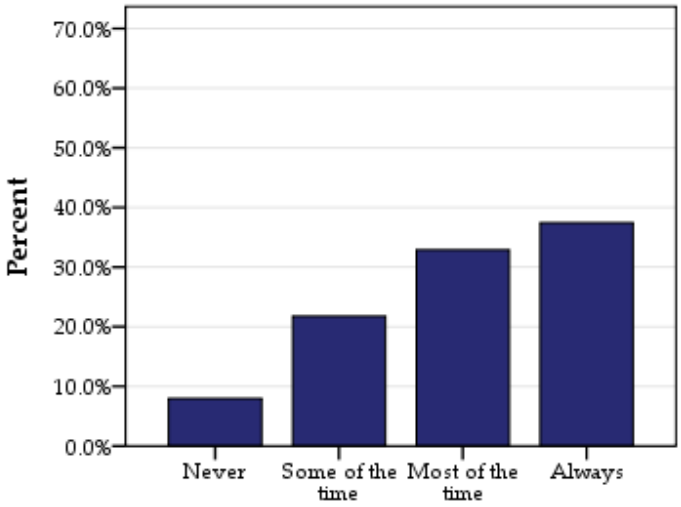
18. When we study a topic, my teacher makes connections to other subjects or classes.



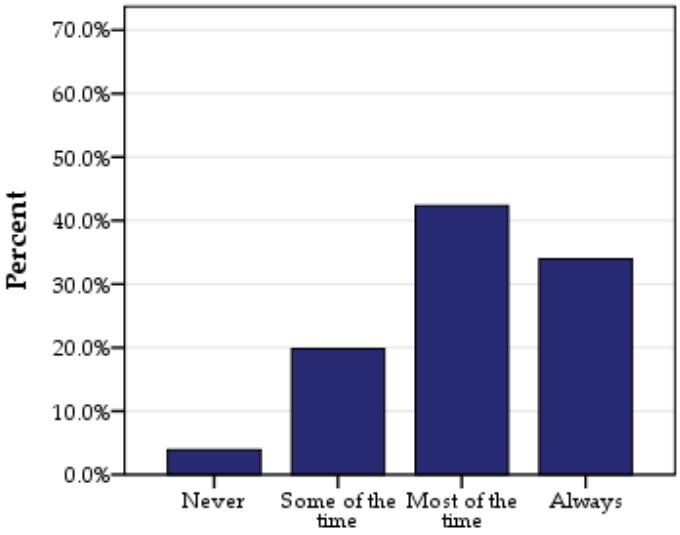
19. My teacher cares about me.



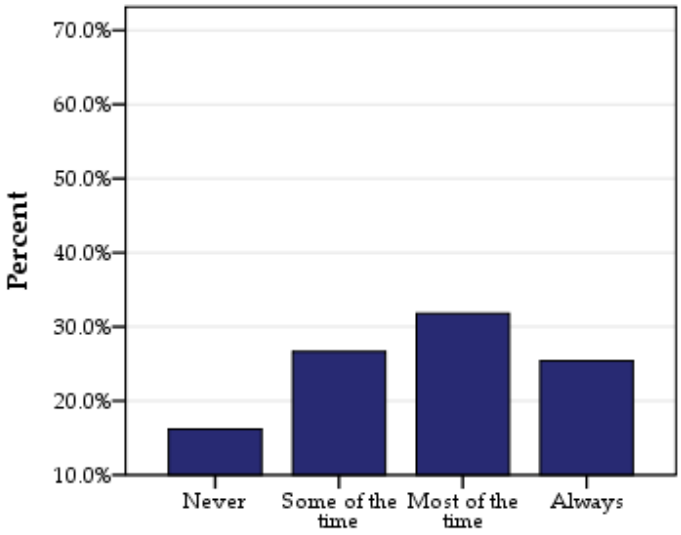
22. Our class stays busy and does not waste time.



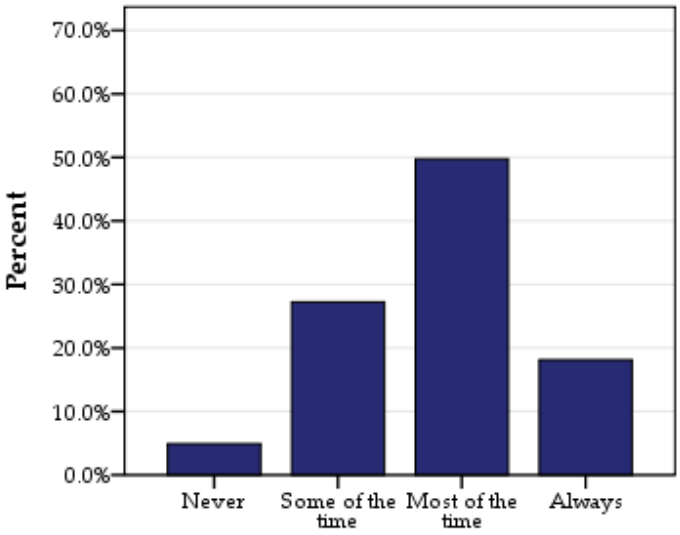
20. My teacher pays attention to what all students are thinking and feeling.



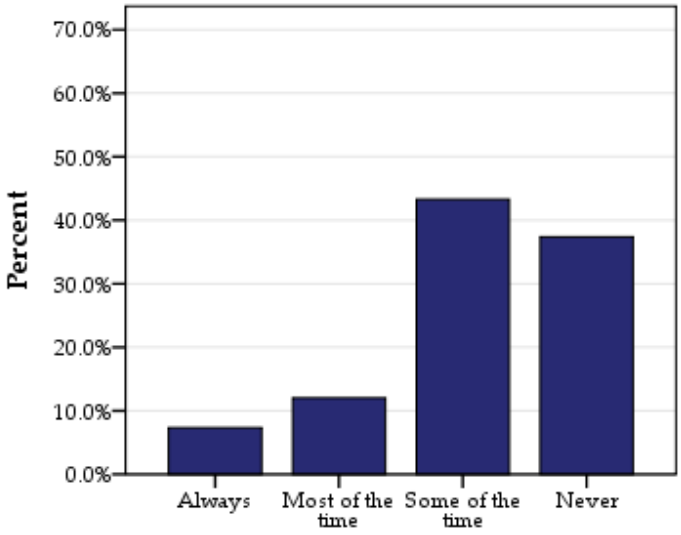
23. Students in this class treat the teacher with respect.



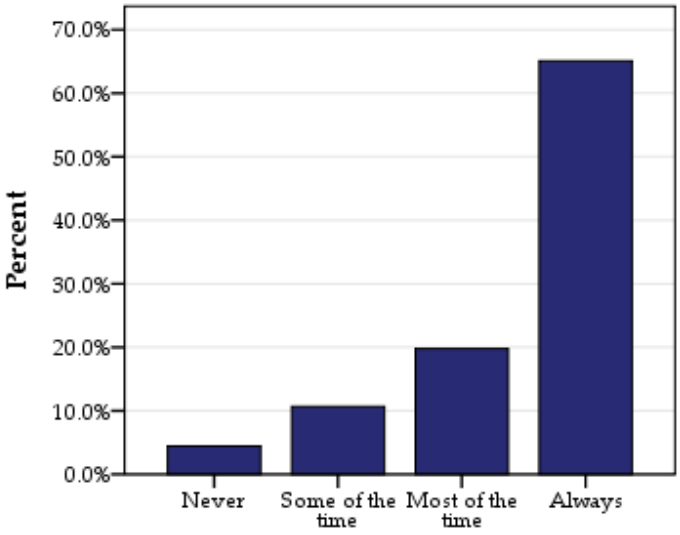
21. My teacher would notice if something was bothering me.



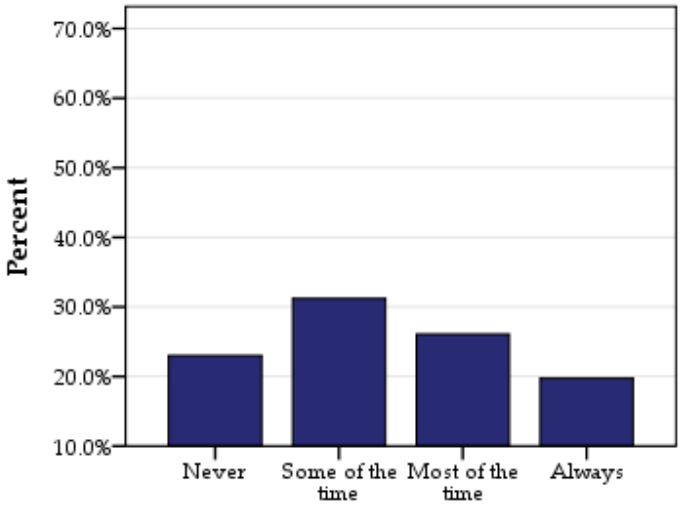
24. The students behave the way my teacher wants them to.



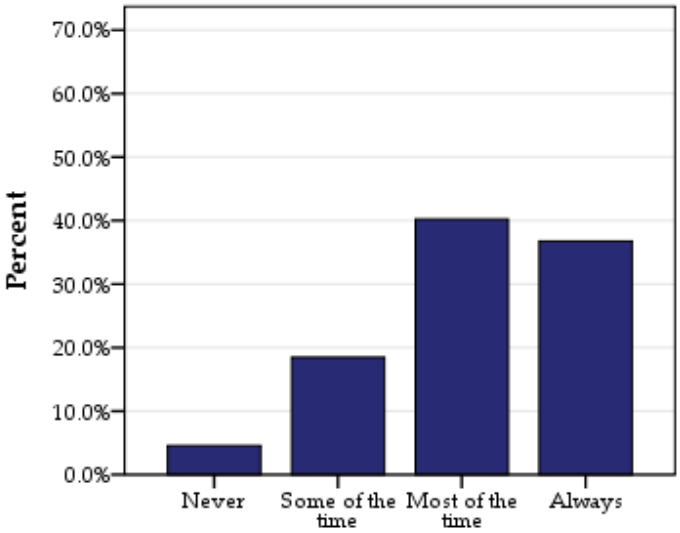
25. The way students behave in this class makes it hard to learn.*



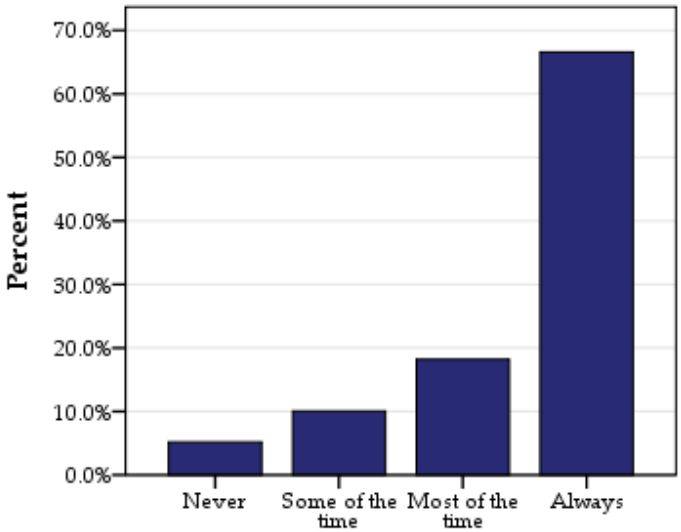
28. My teacher respects me as an individual.



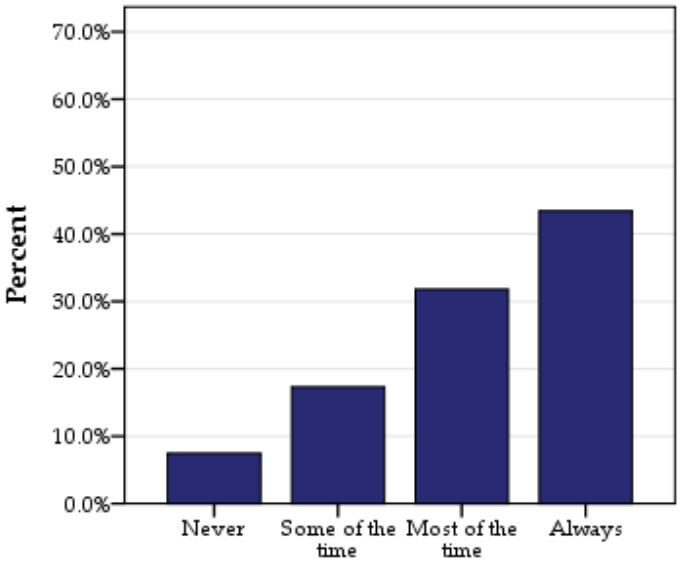
26. Our classroom materials (books, articles, videos, art, music, posters, etc.) reflect my cultural...



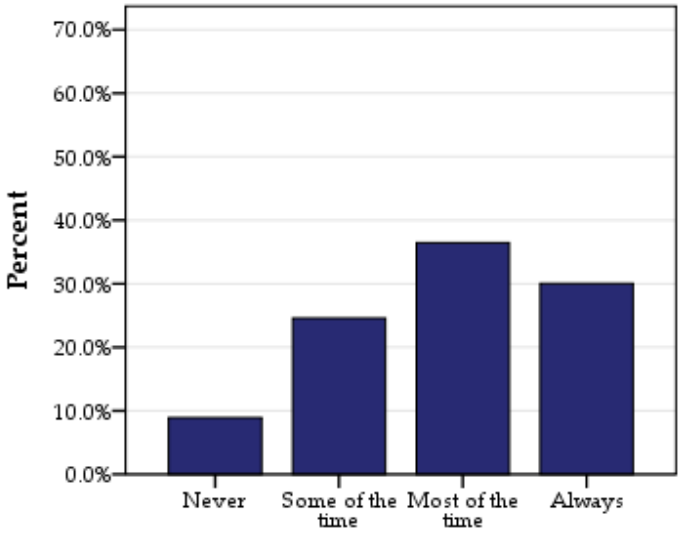
29. Students in this class respect each other's differences.



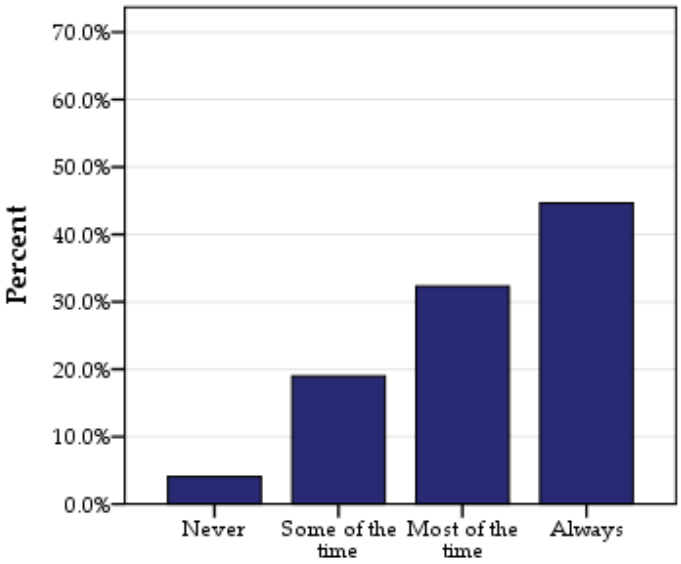
27. My teacher respects my cultural background.



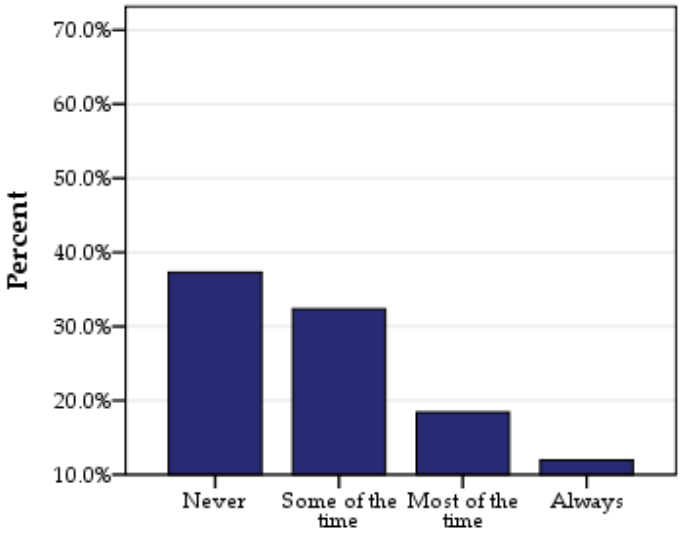
30. In this class, I feel like I fit in.



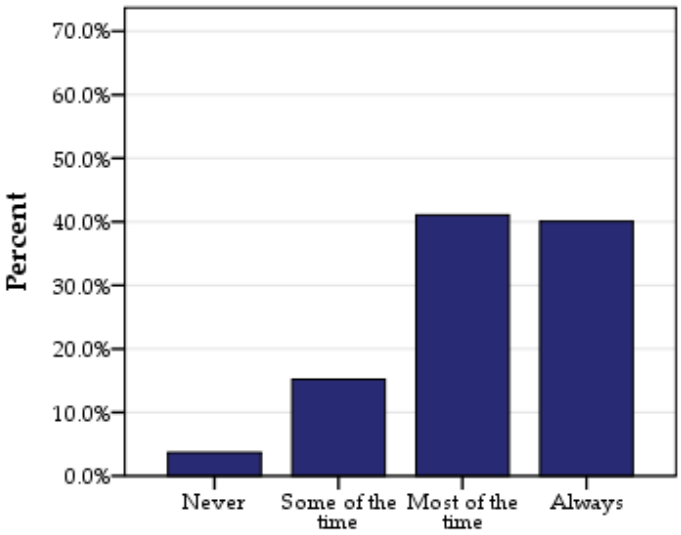
31. I feel like an important part of this classroom community.



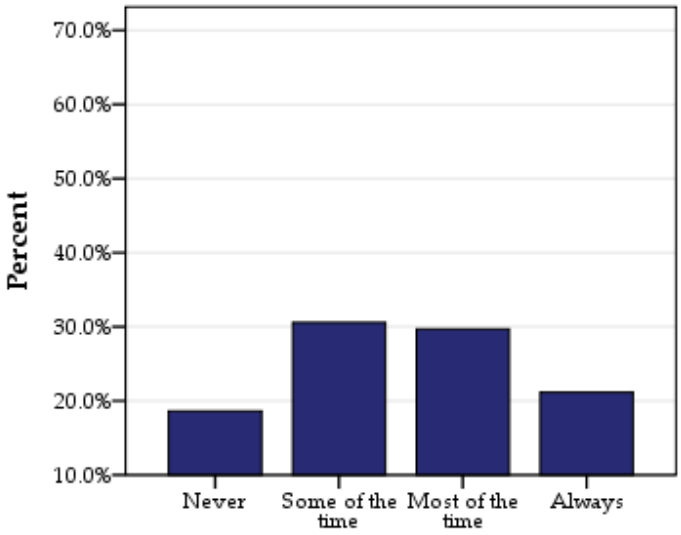
34. I ask for help when I need it.



32. My teacher knows what my life is like outside of school.



35. I feel like I do a good job in this class.



33. My teacher knows what is important to me.

```
------------------------------------------------------------------------------
                            Item Statistics (EAP)
Number of Active Items = 34
Students = 6986
------------------------------------------------------------------------------
..............................................................................

Item: Q01      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 0.97 t = -1.71 Outfit MNSQ = 0.93 t = -4.03

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count                216  1,060  2,868  2,773      69
Percent (%)         3.12  15.32  41.46  40.09
Pt-Biserial        -0.36  -0.39  -0.06   0.48
Mean Ability       -1.14  -0.06   0.65   1.31      NA
SD Abilities        0.24   0.22   0.22   0.25      NA
Step Difficulties          -2.01  -0.66   0.99
Thresholds            NA  -2.21  -0.62   1.15
Error                 NA   0.02   0.04   0.05

==============================================================================

Item: Q02      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 0.98 t = -1.48 Outfit MNSQ = 0.97 t = -1.89

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count                375  2,204  2,756  1,570      81
Percent (%)         5.43  31.92  39.91  22.74
Pt-Biserial        -0.39  -0.34   0.21   0.34
Mean Ability       -0.73   0.30   0.95   1.39      NA
SD Abilities        0.23   0.22   0.23   0.26      NA
Step Difficulties          -1.88   0.37   1.75
Thresholds            NA  -1.97   0.27   1.94
Error                 NA   0.02   0.05   0.05

==============================================================================

Item: Q03      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 1.10 t = 5.61 Outfit MNSQ = 1.12 t = 6.53

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count                462  1,612  2,210  2,607      95
Percent (%)         6.70  23.39  32.07  37.83
Pt-Biserial        -0.32  -0.33   0.01   0.44
Mean Ability       -0.34   0.22   0.73   1.29      NA
SD Abilities        0.23   0.22   0.22   0.25      NA
Step Difficulties          -1.42   0.17   0.87
Thresholds            NA  -1.59   0.02   1.20
Error                 NA   0.02   0.04   0.05

==============================================================================
```

Item: Q04 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.98 t = -0.99 Outfit MNSQ = 0.97 t = -1.67

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	954	2,185	2,202	1,577	68
Percent (%)	13.79	31.58	31.83	22.80	
Pt-Biserial	-0.54	-0.24	0.24	0.44	
Mean Ability	-0.44	0.42	1.03	1.55	NA
SD Abilities	0.22	0.21	0.23	0.27	NA
Step Difficulties		-0.73	0.69	1.57	
Thresholds	NA	-0.92	0.60	1.86	
Error	NA	0.02	0.04	0.05	

==========

Item: Q05 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.08 t = 4.34 Outfit MNSQ = 1.03 t = 1.68

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	257	1,113	2,190	3,354	72
Percent (%)	3.72	16.10	31.67	48.51	
Pt-Biserial	-0.38	-0.39	-0.04	0.47	
Mean Ability	-1.05	-0.03	0.67	1.21	NA
SD Abilities	0.24	0.22	0.22	0.25	NA
Step Difficulties		-1.87	-0.36	0.49	
Thresholds	NA	-2.05	-0.46	0.78	
Error	NA	0.02	0.04	0.05	

==========

Item: Q06 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.21 t = 10.47 Outfit MNSQ = 1.20 t = 10.62

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	455	778	1,486	3,985	282
Percent (%)	6.79	11.61	22.17	59.44	
Pt-Biserial	-0.37	-0.32	-0.09	0.47	
Mean Ability	-0.50	-0.01	0.58	1.11	NA
SD Abilities	0.23	0.22	0.22	0.24	NA
Step Difficulties		-0.89	-0.36	-0.14	
Thresholds	NA	-1.30	-0.43	0.35	
Error	NA	0.02	0.04	0.05	

==========

Item: Q07 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.98 t = -1.33 Outfit MNSQ = 0.94 t = -3.79

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	448	1,316	1,840	3,300	82
Percent (%)	6.49	19.06	26.65	47.80	
Pt-Biserial	-0.47	-0.41	-0.01	0.56	
Mean Ability	-0.85	0.02	0.69	1.30	NA
SD Abilities	0.23	0.21	0.22	0.25	NA
Step Difficulties		-1.32	0.07	0.37	
Thresholds	NA	-1.53	-0.14	0.80	
Error	NA	0.02	0.04	0.05	

==========

Item: Q08 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.18 t = 9.56 Outfit MNSQ = 1.16 t = 8.67

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	422	1,252	2,227	2,917	168
Percent (%)	6.19	18.36	32.66	42.78	
Pt-Biserial	-0.37	-0.27	0.03	0.36	
Mean Ability	-0.55	0.24	0.76	1.15	NA
SD Abilities	0.23	0.22	0.22	0.25	NA
Step Difficulties		-1.34	-0.16	0.71	
Thresholds	NA	-1.57	-0.21	1.00	
Error	NA	0.02	0.04	0.05	

==========

Item: Q09 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.10 t = 5.71 Outfit MNSQ = 1.04 t = 2.56

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	404	1,492	2,315	2,684	91
Percent (%)	5.86	21.64	33.58	38.93	
Pt-Biserial	-0.45	-0.35	0.09	0.43	
Mean Ability	-0.89	0.16	0.82	1.27	NA
SD Abilities	0.23	0.21	0.22	0.25	NA
Step Difficulties		-1.52	0.01	0.87	
Thresholds	NA	-1.70	-0.09	1.16	
Error	NA	0.02	0.04	0.05	

==========

Item: Q10 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.02 t = 1.42 Outfit MNSQ = 0.99 t = -0.44

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	458	1,766	2,595	2,083	84
Percent (%)	6.64	25.59	37.60	30.18	
Pt-Biserial	-0.45	-0.38	0.17	0.43	
Mean Ability	-0.78	0.16	0.90	1.40	NA
SD Abilities	0.23	0.21	0.23	0.26	NA
Step Difficulties		-1.48	0.15	1.32	
Thresholds	NA	-1.64	0.09	1.55	
Error	NA	0.02	0.05	0.05	

==========

Item: Q11 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.12 t = 6.61 Outfit MNSQ = 1.12 t = 6.56

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	334	1,045	1,892	3,626	89
Percent (%)	4.84	15.15	27.43	52.57	
Pt-Biserial	-0.36	-0.30	-0.00	0.37	
Mean Ability	-0.70	0.12	0.72	1.09	NA
SD Abilities	0.23	0.22	0.22	0.24	NA
Step Difficulties		-1.51	-0.28	0.24	
Thresholds	NA	-1.74	-0.41	0.62	
Error	NA	0.02	0.05	0.05	

==========

Item: Q12 Item Set: base Variable: Construct 1

(by parameter) Infit MNSQ = 1.06 t = 3.34 Outfit MNSQ = 1.01 t = 0.50

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	259	902	1,887	3,833	105
Percent (%)	3.76	13.11	27.42	55.70	
Pt-Biserial	-0.37	-0.39	-0.10	0.50	
Mean Ability	-1.00	-0.11	0.58	1.16	NA
SD Abilities	0.23	0.22	0.22	0.24	NA
Step Difficulties		-1.71	-0.48	0.14	
Thresholds	NA	-1.95	-0.59	0.49	
Error	NA	0.02	0.04	0.05	

==========

Item: Q13 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.08 t = 4.23 Outfit MNSQ = 1.03 t = 1.49

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	293	1,173	2,149	3,306	65
Percent (%)	4.23	16.95	31.05	47.77	
Pt-Biserial	-0.38	-0.38	-0.01	0.45	
Mean Ability	-0.89	0.01	0.70	1.19	NA
SD Abilities	0.23	0.22	0.22	0.25	NA
Step Difficulties		-1.75	-0.26	0.49	
Thresholds	NA	-1.93	-0.38	0.80	
Error	NA	0.02	0.04	0.05	

==========

Item: Q14 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.13 t = 7.68 Outfit MNSQ = 1.11 t = 6.18

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	480	1,882	2,446	2,091	87
Percent (%)	6.96	27.28	35.45	30.31	
Pt-Biserial	-0.41	-0.28	0.13	0.37	
Mean Ability	-0.63	0.33	0.88	1.30	NA
SD Abilities	0.23	0.22	0.23	0.25	NA
Step Difficulties		-1.48	0.29	1.27	
Thresholds	NA	-1.62	0.17	1.53	
Error	NA	0.02	0.04	0.05	

==========

Item: Q15 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.15 t = 8.02 Outfit MNSQ = 1.09 t = 5.32

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	397	1,234	2,163	3,087	105
Percent (%)	5.77	17.93	31.43	44.86	
Pt-Biserial	-0.42	-0.30	0.00	0.43	
Mean Ability	-0.79	0.18	0.72	1.20	NA
SD Abilities	0.23	0.22	0.22	0.25	NA
Step Difficulties		-1.41	-0.17	0.60	
Thresholds	NA	-1.64	-0.25	0.92	
Error	NA	0.02	0.04	0.05	

==========

Item: Q16 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.04 t = 2.51 Outfit MNSQ = 1.03 t = 1.52

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	410	2,086	2,840	1,552	98
Percent (%)	5.95	30.28	41.23	22.53	
Pt-Biserial	-0.36	-0.34	0.17	0.37	
Mean Ability	-0.56	0.29	0.90	1.46	NA
SD Abilities	0.23	0.22	0.23	0.26	NA
Step Difficulties		-1.73	0.28	1.79	
Thresholds	NA	-1.84	0.23	1.96	
Error	NA	0.02	0.05	0.05	

==========

Item: Q17 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.01 t = 0.54 Outfit MNSQ = 0.95 t = -2.67

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	546	1,465	2,121	2,738	116
Percent (%)	7.95	21.32	30.87	39.85	
Pt-Biserial	-0.51	-0.35	0.03	0.54	
Mean Ability	-0.79	0.15	0.75	1.39	NA
SD Abilities	0.23	0.21	0.22	0.25	NA
Step Difficulties		-1.14	0.11	0.77	
Thresholds	NA	-1.37	0.00	1.11	
Error	NA	0.02	0.04	0.05	

==========

Item: Q18 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.20 t = 11.45 Outfit MNSQ = 1.20 t = 10.62

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,356	1,626	1,658	2,112	234
Percent (%)	20.08	24.08	24.56	31.28	
Pt-Biserial	-0.48	-0.14	0.13	0.42	
Mean Ability	-0.10	0.51	0.92	1.36	NA
SD Abilities	0.22	0.22	0.23	0.25	NA
Step Difficulties		-0.03	0.67	0.94	
Thresholds	NA	-0.39	0.58	1.41	
Error	NA	0.02	0.04	0.05	

==========

Item: Q19 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.18 t = 10.13 Outfit MNSQ = 1.16 t = 8.93

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	970	2,535	1,996	1,289	196
Percent (%)	14.29	37.33	29.40	18.98	
Pt-Biserial	-0.41	-0.19	0.19	0.37	
Mean Ability	-0.16	0.51	0.99	1.55	NA
SD Abilities	0.22	0.22	0.23	0.27	NA
Step Difficulties		-0.80	1.00	1.74	
Thresholds	NA	-0.95	0.84	2.05	
Error	NA	0.02	0.05	0.05	

==========

Item: Q20 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.87 t = -6.93 Outfit MNSQ = 0.83 t = -10.30

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	407	861	1,090	4,479	149
Percent (%)	5.95	12.59	15.94	65.51	
Pt-Biserial	-0.48	-0.39	-0.15	0.62	
Mean Ability	-0.94	-0.13	0.43	1.16	NA
SD Abilities	0.23	0.21	0.21	0.24	NA
Step Difficulties		-1.14	0.01	-0.60	
Thresholds	NA	-1.43	-0.40	0.15	
Error	NA	0.02	0.05	0.05	

==========

Item: Q21 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.92 t = -5.08 Outfit MNSQ = 0.88 t = -7.21

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,100	1,352	1,504	2,828	202
Percent (%)	16.21	19.93	22.17	41.69	
Pt-Biserial	-0.58	-0.26	0.08	0.58	
Mean Ability	-0.42	0.29	0.84	1.40	NA
SD Abilities	0.22	0.21	0.22	0.25	NA
Step Difficulties		-0.17	0.47	0.44	
Thresholds	NA	-0.56	0.32	1.01	
Error	NA	0.02	0.04	0.05	

==========

Item: Q22 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.99 t = -0.58 Outfit MNSQ = 0.96 t = -2.59

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	905	1,705	2,247	2,020	109
Percent (%)	13.16	24.79	32.67	29.37	
Pt-Biserial	-0.53	-0.27	0.18	0.47	
Mean Ability	-0.46	0.31	0.93	1.47	NA
SD Abilities	0.22	0.21	0.23	0.26	NA
Step Difficulties		-0.60	0.34	1.26	
Thresholds	NA	-0.89	0.34	1.55	
Error	NA	0.02	0.04	0.05	

==========

Item: Q23 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.07 t = 4.18 Outfit MNSQ = 1.06 t = 3.67

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	425	2,256	2,912	1,328	65
Percent (%)	6.14	32.60	42.07	19.19	
Pt-Biserial	-0.30	-0.27	0.19	0.26	
Mean Ability	-0.34	0.40	0.93	1.32	NA
SD Abilities	0.22	0.22	0.23	0.26	NA
Step Difficulties		-1.75	0.37	2.01	
Thresholds	NA	-1.85	0.32	2.16	
Error	NA	0.02	0.05	0.05	

==========

Item: Q24 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.06 t = 3.24 Outfit MNSQ = 1.05 t = 2.74

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	307	1,831	2,993	1,777	78
Percent (%)	4.44	26.51	43.33	25.72	
Pt-Biserial	-0.29	-0.29	0.10	0.31	
Mean Ability	-0.49	0.31	0.83	1.30	NA
SD Abilities	0.23	0.22	0.23	0.26	NA
Step Difficulties		-1.99	0.04	1.65	
Thresholds	NA	-2.10	-0.01	1.80	
Error	NA	0.02	0.05	0.05	

==========

Item: Q25 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.05 t = 3.11 Outfit MNSQ = 1.05 t = 2.76

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	421	2,456	3,207	832	70
Percent (%)	6.09	35.51	46.37	12.03	
Pt-Biserial	-0.30	-0.29	0.25	0.27	
Mean Ability	-0.35	0.39	0.97	1.55	NA
SD Abilities	0.23	0.22	0.23	0.27	NA
Step Difficulties		-1.80	0.41	2.66	
Thresholds	NA	-1.90	0.41	2.75	
Error	NA	0.02	0.06	0.05	

==========

Item: Q26 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.02 t = 1.45 Outfit MNSQ = 1.00 t = 0.11

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	179	1,531	3,317	1,886	73
Percent (%)	2.59	22.15	47.98	27.28	
Pt-Biserial	-0.32	-0.34	0.08	0.35	
Mean Ability	-1.13	0.17	0.79	1.33	NA
SD Abilities	0.24	0.22	0.23	0.26	NA
Step Difficulties		-2.49	-0.32	1.64	
Thresholds	NA	-2.59	-0.34	1.76	
Error	NA	0.02	0.05	0.05	

==========

Item: Q28 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.10 t = 5.96 Outfit MNSQ = 1.11 t = 6.07

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	2,174	3,069	1,099	394	250
Percent (%)	32.27	45.56	16.32	5.85	
Pt-Biserial	-0.44	0.13	0.22	0.26	
Mean Ability	0.17	0.85	1.20	1.94	NA
SD Abilities	0.22	0.23	0.24	0.30	NA
Step Difficulties		0.15	2.11	2.72	
Thresholds	NA	0.02	1.91	3.06	
Error	NA	0.02	0.05	0.06	

==========

Item: Q29 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.10 t = 5.71 Outfit MNSQ = 1.05 t = 2.64

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	477	1,248	1,648	3,462	151
Percent (%)	6.98	18.26	24.11	50.65	
Pt-Biserial	-0.42	-0.30	-0.06	0.49	
Mean Ability	-0.64	0.18	0.63	1.21	NA
SD Abilities	0.23	0.21	0.22	0.25	NA
Step Difficulties		-1.20	0.11	0.19	
Thresholds	NA	-1.43	-0.13	0.69	
Error	NA	0.02	0.04	0.05	

==========

Item: Q30 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.80 t = -12.83 Outfit MNSQ = 0.83 t = -10.40

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	810	1,380	1,836	2,856	104
Percent (%)	11.77	20.05	26.68	41.50	
Pt-Biserial	-0.44	-0.27	0.02	0.49	
Mean Ability	-0.33	0.26	0.74	1.31	NA
SD Abilities	0.22	0.21	0.22	0.25	NA
Step Difficulties		-0.60	0.23	0.59	
Thresholds	NA	-0.92	0.13	1.03	
Error	NA	0.02	0.04	0.05	

==========

Item: Q31 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.73 t = -17.63 Outfit MNSQ = 0.73 t = -17.29

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	744	1,629	2,210	2,286	117
Percent (%)	10.83	23.72	32.17	33.28	
Pt-Biserial	-0.45	-0.31	0.09	0.49	
Mean Ability	-0.40	0.24	0.83	1.42	NA
SD Abilities	0.22	0.21	0.22	0.26	NA
Step Difficulties		-0.82	0.26	1.07	
Thresholds	NA	-1.08	0.21	1.38	
Error	NA	0.02	0.04	0.05	

==========

Item: Q32 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.10 t = 6.19 Outfit MNSQ = 1.11 t = 6.27

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	2,394	2,348	1,366	739	139
Percent (%)	34.96	34.29	19.95	10.79	
Pt-Biserial	-0.50	0.06	0.28	0.31	
Mean Ability	0.14	0.79	1.23	1.73	NA
SD Abilities	0.22	0.22	0.24	0.28	NA
Step Difficulties		0.48	1.55	2.16	
Thresholds	NA	0.21	1.46	2.52	
Error	NA	0.02	0.04	0.06	

==========

Item: Q33 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.91 t = -5.34 Outfit MNSQ = 0.90 t = -5.80

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,207	1,904	1,976	1,736	163
Percent (%)	17.69	27.91	28.96	25.44	
Pt-Biserial	-0.56	-0.19	0.20	0.48	
Mean Ability	-0.32	0.45	0.99	1.57	NA
SD Abilities	0.22	0.21	0.23	0.27	NA
Step Difficulties		-0.31	0.67	1.36	
Thresholds	NA	-0.59	0.62	1.70	
Error	NA	0.02	0.04	0.05	

==========

Item: Q34 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.95 t = -2.76 Outfit MNSQ = 0.93 t = -3.92

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	254	1,235	2,163	3,265	69
Percent (%)	3.67	17.85	31.27	47.20	
Pt-Biserial	-0.33	-0.33	-0.02	0.40	
Mean Ability	-0.82	0.12	0.69	1.16	NA
SD Abilities	0.23	0.22	0.22	0.25	NA
Step Difficulties		-1.96	-0.22	0.52	
Thresholds	NA	-2.11	-0.38	0.83	
Error	NA	0.02	0.05	0.05	

==========

Item: Q35 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.96 t = 0.51 Outfit MNSQ = 0.98 t = 1.41

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	348	887	2,474	3,179	98
Percent (%)	5.05	12.88	35.92	46.15	
Pt-Biserial	-0.36	-0.28	-0.10	0.44	
Mean Ability	-0.72	0.10	0.60	1.21	NA
SD Abilities	0.23	0.22	0.22	0.25	NA
Step Difficulties		-1.30	-0.70	0.67	
Thresholds	NA	-1.66	-0.56	0.88	
Error	NA	0.01	0.03	0.04	

==========

The following statistics include complete cases only.

Cronbach's Alpha	0.94
Student Count	5599

==========

Item: Q01 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.93 t = -7.64 Outfit MNSQ = 0.92 t = -8.50

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,557	5,974	7,943	6,326	27
Percent (%)	7.14	27.40	36.44	29.02	
Pt-Biserial	-0.49	-0.45	0.13	0.58	
Mean Ability	-1.26	0.02	1.02	2.16	NA
SD Abilities	0.26	0.23	0.24	0.32	NA
Step Difficulties		-1.73	0.28	1.70	
Thresholds	NA	-1.84	0.21	1.89	
Error	NA	0.01	0.03	0.03	

==========

Item: Q02 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.08 t = 8.53 Outfit MNSQ = 1.07 t = 7.36

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,441	5,944	7,496	6,799	147
Percent (%)	6.65	27.42	34.58	31.36	
Pt-Biserial	-0.40	-0.37	0.08	0.49	
Mean Ability	-0.96	0.16	0.97	1.91	NA
SD Abilities	0.26	0.23	0.25	0.31	NA
Step Difficulties		-1.83	0.31	1.54	
Thresholds	NA	-1.93	0.20	1.76	
Error	NA	0.01	0.03	0.03	

==========

Item: Q03 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.03 t = 2.71 Outfit MNSQ = 1.01 t = 1.30

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,719	6,336	7,410	6,245	117
Percent (%)	7.92	29.18	34.13	28.77	
Pt-Biserial	-0.45	-0.37	0.12	0.52	
Mean Ability	-1.01	0.18	1.04	2.05	NA
SD Abilities	0.25	0.23	0.25	0.32	NA
Step Difficulties		-1.63	0.45	1.67	
Thresholds	NA	-1.74	0.34	1.89	
Error	NA	0.01	0.03	0.03	

==========

Item: Q04 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.95 t = -4.93 Outfit MNSQ = 0.97 t = -2.96

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	4,041	6,776	6,329	4,582	99
Percent (%)	18.60	31.19	29.13	21.09	
Pt-Biserial	-0.61	-0.21	0.26	0.53	
Mean Ability	-0.64	0.47	1.32	2.39	NA
SD Abilities	0.24	0.23	0.25	0.34	NA
Step Difficulties		-0.48	0.96	2.06	
Thresholds	NA	-0.66	0.91	2.30	
Error	NA	0.01	0.03	0.03	

```
================================================================================

Item: Q05      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 1.15 t = 15.22 Outfit MNSQ = 1.12 t = 12.24

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count                982  5,099  7,945  7,731      70
Percent (%)         4.51  23.44  36.52  35.53
Pt-Biserial        -0.39  -0.43   0.05   0.51
Mean Ability       -1.41  -0.03   0.91   1.84      NA
SD Abilities        0.27   0.23   0.24   0.30      NA
Step Difficulties          -2.24  -0.03   1.36
Thresholds            NA  -2.34  -0.12   1.55
Error                 NA   0.01   0.03   0.03

================================================================================

Item: Q06      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 1.25 t = 23.75 Outfit MNSQ = 1.29 t = 27.52

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count              1,300  4,225  7,064  9,097     141
Percent (%)         5.99  19.48  32.57  41.95
Pt-Biserial        -0.36  -0.32  -0.01   0.43
Mean Ability       -0.91   0.12   0.85   1.60      NA
SD Abilities        0.26   0.24   0.25   0.29      NA
Step Difficulties          -1.75  -0.16   0.99
Thresholds            NA  -1.91  -0.22   1.23
Error                 NA   0.01   0.03   0.03

================================================================================

Item: Q07      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 0.97 t = -2.85 Outfit MNSQ = 0.94 t = -6.62

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count              1,506  4,111  6,208  9,867     135
Percent (%)         6.94  18.95  28.62  45.49
Pt-Biserial        -0.49  -0.44  -0.03   0.63
Mean Ability       -1.33  -0.19   0.77   1.81      NA
SD Abilities        0.26   0.23   0.24   0.29      NA
Step Difficulties          -1.54  -0.06   0.75
Thresholds            NA  -1.73  -0.17   1.05
Error                 NA   0.01   0.03   0.03

================================================================================

Item: Q08      Item Set: base      Variable: Construct 1
 (by parameter) Infit MNSQ = 1.11 t = 10.31 Outfit MNSQ = 1.08 t = 8.57

Categories             0      1      2      3 missing
Responses              1      2      3      4
Count                767  3,069  6,566 11,300     125
Percent (%)         3.53  14.14  30.26  52.07
Pt-Biserial        -0.35  -0.40  -0.08   0.48
Mean Ability       -1.48  -0.28   0.71   1.53      NA
SD Abilities        0.27   0.23   0.24   0.28      NA
Step Difficulties          -2.24  -0.62   0.52
Thresholds            NA  -2.41  -0.69   0.76
Error                 NA   0.01   0.03   0.03
```

Item: Q09 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.01 t = 1.06 Outfit MNSQ = 0.97 t = -2.82

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,560	4,814	8,143	7,168	142
Percent (%)	7.19	22.20	37.55	33.06	
Pt-Biserial	-0.49	-0.42	0.10	0.54	
Mean Ability	-1.27	-0.05	0.98	1.97	NA
SD Abilities	0.26	0.23	0.24	0.31	NA
Step Difficulties		-1.58	-0.04	1.51	
Thresholds	NA	-1.75	-0.04	1.68	
Error	NA	0.01	0.03	0.03	

==========

Item: Q10 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.03 t = 2.84 Outfit MNSQ = 1.00 t = -0.19

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,943	4,828	6,842	8,078	136
Percent (%)	8.96	22.26	31.54	37.24	
Pt-Biserial	-0.52	-0.39	0.05	0.59	
Mean Ability	-1.11	0.01	0.92	1.94	NA
SD Abilities	0.25	0.23	0.24	0.31	NA
Step Difficulties		-1.30	0.15	1.19	
Thresholds	NA	-1.48	0.09	1.44	
Error	NA	0.01	0.03	0.03	

==========

Item: Q11 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.95 t = -5.55 Outfit MNSQ = 0.92 t = -8.92

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,251	5,184	8,803	6,466	123
Percent (%)	5.76	23.88	40.56	29.79	
Pt-Biserial	-0.47	-0.49	0.12	0.57	
Mean Ability	-1.45	-0.13	0.99	2.12	NA
SD Abilities	0.27	0.23	0.24	0.32	NA
Step Difficulties		-1.92	-0.04	1.74	
Thresholds	NA	-2.05	-0.05	1.88	
Error	NA	0.01	0.03	0.03	

==========

Item: Q12 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.17 t = 15.61 Outfit MNSQ = 1.13 t = 12.81

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	729	2,572	6,177	12,213	136
Percent (%)	3.36	11.86	28.48	56.30	
Pt-Biserial	-0.36	-0.36	-0.13	0.48	
Mean Ability	-1.60	-0.26	0.61	1.47	NA
SD Abilities	0.27	0.23	0.24	0.28	NA
Step Difficulties		-2.19	-0.82	0.31	
Thresholds	NA	-2.39	-0.86	0.55	
Error	NA	0.01	0.03	0.03	

==========

Item: Q13 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.08 t = 7.80 Outfit MNSQ = 1.05 t = 4.73

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,019	3,943	7,588	9,132	145
Percent (%)	4.70	18.19	35.00	42.12	
Pt-Biserial	-0.40	-0.42	-0.04	0.54	
Mean Ability	-1.40	-0.16	0.77	1.76	NA
SD Abilities	0.27	0.23	0.24	0.30	NA
Step Difficulties		-2.02	-0.35	1.04	
Thresholds	NA	-2.17	-0.39	1.23	
Error	NA	0.01	0.03	0.03	

==========

Item: Q14 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.00 t = -0.17 Outfit MNSQ = 0.95 t = -5.43

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,096	3,549	5,851	11,153	178
Percent (%)	5.06	16.39	27.03	51.52	
Pt-Biserial	-0.45	-0.47	-0.11	0.64	
Mean Ability	-1.53	-0.33	0.62	1.72	NA
SD Abilities	0.27	0.23	0.23	0.29	NA
Step Difficulties		-1.87	-0.28	0.46	
Thresholds	NA	-2.05	-0.41	0.78	
Error	NA	0.01	0.03	0.03	

==========

Item: Q15 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.19 t = 19.10 Outfit MNSQ = 1.20 t = 19.75

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	4,133	7,885	5,707	3,928	174
Percent (%)	19.09	36.42	26.36	18.14	
Pt-Biserial	-0.52	-0.12	0.22	0.43	
Mean Ability	-0.40	0.66	1.29	2.27	NA
SD Abilities	0.24	0.24	0.26	0.34	NA
Step Difficulties		-0.54	1.29	2.21	
Thresholds	NA	-0.68	1.16	2.48	
Error	NA	0.01	0.03	0.03	

==========

Item: Q16 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.05 t = 4.79 Outfit MNSQ = 1.01 t = 1.07

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	2,004	5,237	7,324	7,109	153
Percent (%)	9.25	24.16	33.79	32.80	
Pt-Biserial	-0.49	-0.39	0.08	0.58	
Mean Ability	-1.00	0.06	0.96	2.04	NA
SD Abilities	0.25	0.23	0.24	0.31	NA
Step Difficulties		-1.30	0.22	1.45	
Thresholds	NA	-1.48	0.18	1.67	
Error	NA	0.01	0.03	0.03	

==========

Item: Q17 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.31 t = 29.85 Outfit MNSQ = 1.33 t = 30.62

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	3,813	5,535	5,832	6,423	224
Percent (%)	17.65	25.62	27.00	29.73	
Pt-Biserial	-0.45	-0.22	0.11	0.48	
Mean Ability	-0.31	0.43	1.06	1.94	NA
SD Abilities	0.24	0.24	0.25	0.31	NA
Step Difficulties		-0.43	0.70	1.46	
Thresholds	NA	-0.67	0.63	1.78	
Error	NA	0.01	0.03	0.03	

==========

Item: Q18 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.18 t = 18.08 Outfit MNSQ = 1.17 t = 16.37

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	2,626	7,406	6,667	4,931	197
Percent (%)	12.14	34.24	30.82	22.80	
Pt-Biserial	-0.47	-0.25	0.16	0.47	
Mean Ability	-0.65	0.44	1.13	2.17	NA
SD Abilities	0.25	0.24	0.25	0.33	NA
Step Difficulties		-1.15	0.89	1.97	
Thresholds	NA	-1.27	0.77	2.21	
Error	NA	0.01	0.03	0.03	

==========

Item: Q19 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.94 t = -6.28 Outfit MNSQ = 0.90 t = -10.53

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,583	3,697	5,521	10,716	310
Percent (%)	7.36	17.18	25.66	49.80	
Pt-Biserial	-0.52	-0.43	-0.07	0.65	
Mean Ability	-1.35	-0.21	0.69	1.77	NA
SD Abilities	0.26	0.23	0.23	0.29	NA
Step Difficulties		-1.40	-0.09	0.50	
Thresholds	NA	-1.62	-0.22	0.86	
Error	NA	0.01	0.03	0.03	

==========

Item: Q20 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.98 t = -2.17 Outfit MNSQ = 0.94 t = -6.41

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,716	4,694	7,102	8,077	238
Percent (%)	7.95	21.74	32.90	37.41	
Pt-Biserial	-0.53	-0.44	0.05	0.62	
Mean Ability	-1.28	-0.09	0.90	1.99	NA
SD Abilities	0.26	0.23	0.24	0.31	NA
Step Difficulties		-1.44	0.05	1.21	
Thresholds	NA	-1.62	0.00	1.45	
Error	NA	0.01	0.03	0.03	

==========

Item: Q21 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.05 t = 4.86 Outfit MNSQ = 1.03 t = 2.90

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	3,495	5,753	6,862	5,483	234
Percent (%)	16.19	26.64	31.78	25.39	
Pt-Biserial	-0.57	-0.26	0.21	0.52	
Mean Ability	-0.67	0.35	1.19	2.19	NA
SD Abilities	0.24	0.23	0.25	0.33	NA
Step Difficulties		-0.57	0.59	1.84	
Thresholds	NA	-0.80	0.61	2.06	
Error	NA	0.01	0.03	0.03	

==========

Item: Q22 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.16 t = 15.94 Outfit MNSQ = 1.13 t = 13.05

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,139	5,523	9,745	5,282	138
Percent (%)	5.25	25.46	44.93	24.35	
Pt-Biserial	-0.39	-0.36	0.16	0.38	
Mean Ability	-1.18	0.14	1.07	1.90	NA
SD Abilities	0.26	0.23	0.25	0.32	NA
Step Difficulties		-2.06	-0.03	2.12	
Thresholds	NA	-2.18	-0.02	2.23	
Error	NA	0.01	0.03	0.03	

==========

Item: Q23 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.18 t = 17.26 Outfit MNSQ = 1.15 t = 15.05

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	846	4,297	9,175	7,357	152
Percent (%)	3.90	19.82	42.33	33.94	
Pt-Biserial	-0.33	-0.35	-0.02	0.45	
Mean Ability	-1.25	0.06	0.82	1.78	NA
SD Abilities	0.27	0.23	0.24	0.30	NA
Step Difficulties		-2.29	-0.39	1.55	
Thresholds	NA	-2.42	-0.39	1.67	
Error	NA	0.01	0.03	0.03	

==========

Item: Q24 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.13 t = 13.29 Outfit MNSQ = 1.11 t = 10.66

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,064	5,893	10,773	3,921	176
Percent (%)	4.91	27.22	49.76	18.11	
Pt-Biserial	-0.35	-0.36	0.18	0.38	
Mean Ability	-1.07	0.17	1.08	2.14	NA
SD Abilities	0.26	0.23	0.25	0.34	NA
Step Difficulties		-2.18	-0.00	2.64	
Thresholds	NA	-2.28	0.03	2.70	
Error	NA	0.01	0.03	0.03	

==========

Item: Q26 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.40 t = 37.66 Outfit MNSQ = 1.45 t = 40.08

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	4,817	6,538	5,470	4,135	867
Percent (%)	22.98	31.19	26.10	19.73	
Pt-Biserial	-0.38	-0.13	0.16	0.37	
Mean Ability	0.02	0.63	1.20	2.03	NA
SD Abilities	0.24	0.24	0.26	0.33	NA
Step Difficulties		-0.16	1.15	2.09	
Thresholds	NA	-0.38	1.09	2.37	
Error	NA	0.01	0.03	0.03	

==========

Item: Q27 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.16 t = 13.58 Outfit MNSQ = 1.11 t = 10.51

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,087	2,119	3,827	14,015	779
Percent (%)	5.16	10.07	18.18	66.59	
Pt-Biserial	-0.41	-0.34	-0.16	0.54	
Mean Ability	-1.31	-0.31	0.46	1.41	NA
SD Abilities	0.26	0.23	0.23	0.28	NA
Step Difficulties		-1.52	-0.58	-0.40	
Thresholds	NA	-1.83	-0.74	0.09	
Error	NA	0.01	0.03	0.03	

==========

Item: Q28 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.97 t = -2.52 Outfit MNSQ = 0.92 t = -7.98

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	960	2,299	4,263	14,016	289
Percent (%)	4.46	10.67	19.79	65.08	
Pt-Biserial	-0.44	-0.44	-0.22	0.65	
Mean Ability	-1.64	-0.56	0.33	1.52	NA
SD Abilities	0.27	0.23	0.23	0.28	NA
Step Difficulties		-1.76	-0.61	-0.28	
Thresholds	NA	-2.02	-0.77	0.15	
Error	NA	0.01	0.03	0.03	

==========

Item: Q29 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.17 t = 16.82 Outfit MNSQ = 1.15 t = 14.66

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	981	3,991	8,689	7,937	229
Percent (%)	4.54	18.48	40.23	36.75	
Pt-Biserial	-0.33	-0.31	-0.01	0.40	
Mean Ability	-1.03	0.12	0.83	1.66	NA
SD Abilities	0.26	0.23	0.24	0.30	NA
Step Difficulties		-2.05	-0.43	1.38	
Thresholds	NA	-2.21	-0.41	1.52	
Error	NA	0.01	0.03	0.03	

==========

Item: Q30 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.95 t = -5.44 Outfit MNSQ = 0.94 t = -6.25

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,613	3,744	6,877	9,385	208
Percent (%)	7.46	17.32	31.81	43.41	
Pt-Biserial	-0.41	-0.33	-0.05	0.51	
Mean Ability	-0.92	0.04	0.75	1.70	NA
SD Abilities	0.25	0.23	0.24	0.29	NA
Step Difficulties		-1.36	-0.25	0.92	
Thresholds	NA	-1.61	-0.24	1.16	
Error	NA	0.01	0.03	0.03	

==========

Item: Q31 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.83 t = -19.02 Outfit MNSQ = 0.81 t = -20.91

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	1,923	5,305	7,872	6,483	244
Percent (%)	8.91	24.58	36.47	30.04	
Pt-Biserial	-0.47	-0.37	0.09	0.55	
Mean Ability	-0.96	0.10	0.97	2.09	NA
SD Abilities	0.25	0.23	0.24	0.32	NA
Step Difficulties		-1.35	0.18	1.66	
Thresholds	NA	-1.52	0.17	1.84	
Error	NA	0.01	0.03	0.03	

==========

Item: Q32 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 1.21 t = 20.41 Outfit MNSQ = 1.23 t = 22.34

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	8,025	6,963	3,963	2,574	302
Percent (%)	37.28	32.35	18.41	11.96	
Pt-Biserial	-0.53	0.08	0.24	0.39	
Mean Ability	0.04	0.98	1.51	2.55	NA
SD Abilities	0.24	0.25	0.27	0.37	NA
Step Difficulties		0.61	1.86	2.60	
Thresholds	NA	0.39	1.77	2.92	
Error	NA	0.01	0.03	0.04	

==========

Item: Q33 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.97 t = -3.41 Outfit MNSQ = 0.96 t = -4.21

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	4,004	6,579	6,390	4,550	304
Percent (%)	18.60	30.57	29.69	21.14	
Pt-Biserial	-0.57	-0.21	0.23	0.53	
Mean Ability	-0.55	0.47	1.25	2.39	NA
SD Abilities	0.24	0.23	0.25	0.34	NA
Step Difficulties		-0.46	0.91	2.08	
Thresholds	NA	-0.66	0.88	2.31	
Error	NA	0.01	0.03	0.03	

==========

Item: Q34 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.87 t = -13.28 Outfit MNSQ = 0.86 t = -14.85

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	880	4,104	6,988	9,652	203
Percent (%)	4.07	18.98	32.32	44.64	
Pt-Biserial	-0.32	-0.35	-0.08	0.48	
Mean Ability	-1.09	0.03	0.70	1.63	NA
SD Abilities	0.26	0.23	0.24	0.29	NA
Step Difficulties		-2.24	-0.25	0.87	
Thresholds	NA	-2.36	-0.37	1.10	
Error	NA	0.01	0.03	0.03	

==========

Item: Q35 Item Set: base Variable: Construct 1
(by parameter) Infit MNSQ = 0.93 t = 0.15 Outfit MNSQ = 0.95 t = 0.20

Categories	0	1	2	3	missing
Responses	1	2	3	4	
Count	794	3,288	8,891	8,679	175
Percent (%)	3.67	15.19	41.06	40.08	
Pt-Biserial	-0.34	-0.35	-0.07	0.46	
Mean Ability	-1.37	-0.09	0.74	1.69	NA
SD Abilities	0.26	0.23	0.24	0.30	NA
Step Difficulties		-2.20	-0.74	1.24	
Thresholds	NA	-2.38	-0.68	1.36	
Error	NA	0.00	0.02	0.02	

==========

The following statistics include complete cases only.

Cronbach's Alpha	0.96
Student Count	19027

Appendix K: *Correlational Results from MET Study*

Table 7. Pairwise Correlations with Teacher Value-Added: Math

	VALUE-ADDED ON STATE MATHEMATICS TEST			VALUE ADDED ON BALANCED ASSESSMENT IN MATH		NUMBER OF TEACHERS
	SAME SECTION	DIFFERENT SECTION	PRIOR YEAR	SAME SECTION	DIFFERENT SECTION	
VA on State Math Test	1.000	0.380 ***	0.404 ***	0.377 ***	0.161 ***	1011
Disattenuated					0.542	
TRIPOD:						
Sum of 7 C's	0.212 ***	0.218 ***	0.203 ***	0.107 ***	0.114 ***	952
Disattenuated		0.433	0.346		0.296	
Care	0.158 ***	0.155 ***	0.146 ***	0.073 **	0.096 ***	952
Disattenuated		0.307	0.265		0.246	
Clarify	0.208 ***	0.237 ***	0.189 ***	0.093 ***	0.105 ***	952
Disattenuated		0.487	0.336		0.281	
Control	0.224 ***	0.171 ***	0.180 ***	0.182 ***	0.143 ***	952
Disattenuated		0.384	0.352		0.420	
Challenge	0.219 ***	0.216 ***	0.232 ***	0.080 **	0.115 ***	952
Disattenuated		0.436	0.404		0.301	
Captivate	0.158 ***	0.197 ***	0.152 ***	0.080 **	0.082 **	952
Disattenuated		0.388	0.258		0.210	
Confer	0.135 ***	0.166 ***	0.157 ***	0.049	0.091 ***	952
Disattenuated		0.336	0.275		0.241	
Consolidate	0.142 ***	0.181 ***	0.153 ***	0.052	0.050	952
Disattenuated		0.367	0.268		0.132	
Control+Challenge	0.256 ***	0.219 ***	0.235 ***	0.160 ***	0.149 ***	952
Disattenuated		0.465	0.435		0.414	
Other 5 C's	0.173 ***	0.201 ***	0.173 ***	0.075 **	0.091 ***	952
Disattenuated		0.395	0.298		0.234	

Note: A *,**, or *** indicates a correlation that is significantly different from zero at the .10, .05 and .01 level respectively. The correlation for "different section" was for at most one video observation in another section, so is likely to increase as more videos are scored. Disattenuated correlations under State Mathematics Test for different section and prior year are not reported as they are by definition 1.

Table 8. Pairwise Correlations with Teacher Value-Added: ELA

	VALUE-ADDED ON STATE ENGLISH LANGUAGE ARTS			VALUE ADDED ON STANFORD 9 OE ELA		NUMBER OF TEACHERS
	SAME SECTION	DIFFERENT SECTION	PRIOR YEAR	SAME SECTION	DIFFERENT SECTION	
VA on State ELA Test	1.000	0.179 ***	0.195 ***	0.221 ***	0.093 ***	1096
Disattenuated					0.367	
TRIPOD:						
Sum of 7 C's	0.095 ***	0.070 **	0.099 ***	0.135 ***	0.063 **	1026
Disattenuated		0.195	0.250		0.121	
Care	0.029	0.027	0.039	0.081 **	0.002	1026
Disattenuated		0.075	0.105		0.004	
Clarify	0.087 ***	0.072 **	0.073 **	0.123 ***	0.064 **	1026
Disattenuated		0.198	0.186		0.121	
Control	0.142 ***	0.099 ***	0.084 **	0.158 ***	0.088 ***	1026
Disattenuated		0.294	0.236		0.183	
Challenge	0.128 ***	0.111 ***	0.162 ***	0.147 ***	0.092 ***	1026
Disattenuated		0.316	0.427		0.178	
Captivate	0.050	0.049	0.078 **	0.102 ***	0.041	1026
Disattenuated		0.136	0.196		0.080	
Confer	0.040	0.002	0.070 **	0.078 **	0.017	1026
Disattenuated		0.007	0.184		0.036	
Consolidate	0.080 **	0.067 **	0.092 ***	0.106 ***	0.079 **	1026
Disattenuated		0.189	0.238		0.154	
Control+Challenge	0.158 ***	0.118 ***	0.138 ***	0.180 ***	0.101 ***	1026
Disattenuated		0.342	0.374		0.203	
Other 5 C's	0.060 *	0.045	0.076 **	0.106 ***	0.042	1026
Disattenuated		0.124	0.194		0.080	

Note: A *,**, or *** indicates a correlation that is significantly different from zero at the .10, .05 and .01 level respectively. The correlation for "different section" was for at most one video observation in another section, so is likely to increase as more videos are scored.

Source: Bill & Melinda Gates Foundation. (2010). *Learning about Teaching: Initial Findings from the Measures of Effective Teaching Project.* The full report is available here.

Appendix L: *Quick Links to CEI Student Perception Survey Documents*

Links to Student Perception Survey Resources

- Student Perception Survey homepage
  - Here you will find:
    - The Full technical report
- Toolkit homepage
  - Here you will find:
    - The Planning guide
- Planning homepage
  - Here you will find:
    - Data checklists
    - Communication materials
    - Sample presentations
- Administration homepage
  - Here you will find:
    - Paper/pencil versions of the survey
    - Sample proctor guides
    - Sample building coordinator guides
- Results and Reflection homepage
  - Here you will find:
    - Sample reports
    - Communication materials
    - Guidance for using results
    - Norming Data
- Additional Resources for Teachers
  - Here you will find:
    - The survey instruments
    - Guidance for using results
    - Teacher reflections on the survey
    - Additional research on student surveys

Additional Resources

- The Teacher Perception Survey homepage
- The Colorado Education Initiative's Transforming School Climate Toolkit
- The Colorado Department of Education's State Model Evaluation System