

# Description of the Development of Scores for 2011–2012 Program Reviews and an Investigation of their Quantitative Properties

Prepared for: Kentucky Department of Education  
Capital Plaza Tower, 17<sup>th</sup> Floor  
500 Mero Street  
Frankfort, KY 40601

Authors: Andrea L. Sinclair  
Arthur A. Thacker

Date: January 2013



**HumRRO**  
Human Resources Research Organization

Human Resources Research Organization (HumRRO)  
66 Canal Center Plaza, Suite 700, Alexandria, Virginia 22314-1591  
Phone: 703.549.3611 | Fax: 703.549.9661 | [www.humrro.org](http://www.humrro.org)



# Description of the Development of Scores for 2011–2012 Program Reviews and an Investigation of their Quantitative Properties

Prepared for: Kentucky Department of Education  
Capital Plaza Tower, 17<sup>th</sup> Floor  
500 Mero Street  
Frankfort, KY 40601

Authors: Andrea L. Sinclair  
Arthur A. Thacker

Date: January 2013





# Description of the Development of Scores for 2011–2012 Program Reviews and an Investigation of their Quantitative Properties

## Table of Contents

Executive Summary .....	ii
Introduction .....	1
Background.....	2
Section 1: Method for Reporting Program Review Scores and Performance Cuts .....	3
Summary of Method for Reporting Program Review Scores and Performance Cuts .....	4
Section 2: Quantitative Investigation of Program Review Scores.....	6
Investigation of Frequency Distribution of Item Ratings.....	6
Scale Level Analyses.....	10
Follow-up Analyses on Subscales .....	19
Exploratory Principal Components Analysis for PL/CS.....	19
Summary of Quantitative Investigation of the Program Review Scores.....	20
References.....	21
Appendix A: Example 2011–2012 Score Report from Arts/Humanities .....	1
Appendix B: Frequency Counts for Items Rated as “Not Applicable” By a Substantive Number of Schools .....	1

## List of Tables

Table 1. Items for Which the Majority of Schools Selected “Not Applicable” .....	8
Table 2. Items Flagged Due to Little Variance in Ratings.....	10
Table 3. Writing Internal Consistency Reliability Estimates for Scales and Subscales by Grade Level .....	12
Table 4. Arts & Humanities Internal Consistency Reliability Estimates for Scales and Subscales by Grade Level.....	13
Table 5. PL/CS Internal Consistency Reliability Estimates for Scales and Subscales by Grade Level .....	14
Table 6. PL/CS Items Flagged Based on Coefficient Alpha If Item Deleted and Item Total Correlation (Scale Level Analysis) .....	16
Table 7. Item Pairs Flagged Based on Weak Inter-Item Correlations (Scale Level Analysis) .....	18

## Description of the Development of Scores for 2011–2012 Program Reviews and an Investigation of their Quantitative Properties

### Executive Summary

At the conclusion of the 2011–2012 pilot program review, the scores reported to schools were comprised of mean scores on each standard and the overall mean score for the program. The overall mean score for the program was based on the mean of all the characteristics for the program. Classification into Needs Improvement, Proficient, and Distinguished was determined based on mean scores on the four standards for each program. Schools designated as Needs Improvement on two or more of the standards received a Needs Improvement classification for the program area regardless of total mean score. Schools designated as proficient or higher on three or four of the standards were classified as Proficient if the sum of the means on the standards was between 8.0 and less than 10.8, and schools were classified as Distinguished if the sum of the means on the standards was 10.8 or greater.

The quantitative investigation of the program review measures supports the use of the Writing and the Arts & Humanities scores with results reported out at the scale (i.e., “standard”) level by grade. Minor revisions and tweaks of items on these scales by content experts are recommended (see items flagged in the tables and Appendices of this report). The items on the Practical Living and Career Studies (PL/CS) measure, however, would benefit from substantive review and revision by content experts. In its current form, the scales on the PL/CS measure have reasonably high coefficient alphas; however, those values are largely an artifact of the length of the scales and not due to the homogeneity of the items on the scales.

For the 2012–2013 operational program reviews (2013 reporting), HumRRO recommends that the ASSIST web-based tool be “hard coded” so that characteristics that are “not applicable” to certain grade levels are unavailable for entering ratings. Furthermore, HumRRO recommends that an additional quality assurance step be conducted to verify that schools are only able to submit one final set of program review ratings per program area for their school. HumRRO also recommends that the scoring decisions implemented in the pilot year be re-visited in light of the 2012–2013 operational data, which presumably should not contain the limitations of the pilot year data. Finally, HumRRO also recommends that the quantitative investigation of the properties of the program review scores (e.g., estimates of scale reliability) be re-investigated using 2012–2013 data due to the recent changes KDE has made to the 2012–2013 program review materials (e.g., changes to the rubric).

## Introduction

In March 2009, Kentucky's General Assembly passed Senate Bill 1. Passage of this bill established the implementation of a program review to be included as part of a new assessment and accountability model. The Kentucky Department of Education (KDE) defines its program reviews as:

*...a systematic method of analyzing components of an instructional program, including instructional practices, aligned and enacted curriculum, student work samples, formative and summative assessments, professional development and support services, and administrative support and monitoring KRS 158.6453(1)(i)*

The program reviews for the 2011–2012 academic year covers three content areas: Writing, Arts and Humanities, and Practical Living Skills & Career Studies. The KDE identifies the purposes of the program reviews as:

- improving the quality of teaching and learning for all students in all programs
- allowing equal access to all students the skills that will assist them in being productive citizens
- allowing student demonstration of understanding beyond a paper-and-pencil test
- ensuring a school-wide natural integration of the program skills across all contents, beyond the program areas

The 2011–2012 academic year was designated as a pilot year for program reviews. As such, all schools were to submit program review data; however, scores were not included in the Unbridled Learning: College/Career Readiness for All accountability system during its first year (i.e., 2012 reporting).

KDE asked HumRRO to provide input on the program review plan and its implementation. HumRRO provided guidance to KDE on how the 2011–2012 scores could be calculated and reported to schools. This report describes that process, and the end result of how determinations of cut scores for Needs Improvement, Proficient, and Distinguished were determined for each of the program review areas.

In the second section of this report, we investigate the soundness of using the standards as scales for publicly reporting school-level findings. This quantitative investigation stems from a recommendation provided by Kentucky's National Technical Advisory Panel on Assessment and Accountability (NTAPAA) during its meeting with KDE on December 5, 2012. The NTAPAA expressed concerns about the subjectivity of the current program review process (stemming largely from the absence of an audit of the process) coupled with complaints from the field of the overly burdensome nature of the program review in its current format. Consequently, the NTAPAA suggested that rather than using program reviews in the accountability model, the results from the reviews could be publicly reported. In the second section of this report, we investigate the internal consistency of the "characteristics" (i.e., items) for each standard to determine the quantitative soundness of treating each standard as a scale for publicly reporting school-level findings.

## Background

During the 2011–2012 academic year, Kentucky schools participated in a Program Review pilot for the areas of Arts and Humanities, Practical Living Skills & Career Studies, and Writing. Each program review is organized into four standards with supporting demonstrators. Demonstrators are further divided into multiple characteristics. The four standards for each program area and their supporting demonstrators are:

1. Curriculum and Instruction:
  - Student Access
  - Aligned and Rigorous Curriculum
  - Instructional Strategies
  - Student Performance
2. Formative and Summative Assessment
  - Variety of Assessment
  - Expectations for Student Learning
  - Response to Assessment
3. Professional Development
  - Planning
  - Participation
  - Teacher Leadership
4. Administrative/Leadership Support and Monitoring (abbreviated: Administrative Support):
  - Shared Vision
  - Time and Resources
  - Policies and Monitoring
  - Principal Leadership

The characteristics (i.e., items) underlying each demonstrator are not presented here, as they are unique (both in number and in content) to each program area. Schools submitted data to KDE at the characteristic level. All information at the standard or demonstrator level is necessarily generated by aggregation of the characteristic ratings.

### *Caveats about the Data File*

The KDE contracted with AdvancEd to collect schools' program review data and to provide a clean, raw data file. Schools were to enter their program review ratings using software within the Adaptive System of School Improvement Support Tools (ASSIST) system, which is the web-based tool developed by AdvancEd and used by schools for submitting program review ratings. The Program Review Scoring Key in the raw data file from AdvancEd identifies a four point scoring key (i.e., rating scale) for which 0 = "N/A (This characteristic refers to a specific grade level(s) that is not a part of our school)", 1 = "Needs Improvement," 2 = "Proficient," and 3 = "Distinguished." When KDE received the final data file from AdvancEd they found that some schools may have used the '0' option incorrectly as evidenced by the finding that some schools of the same grade level had entered '0' for particular characteristics and other schools had entered a 1, 2, or 3 for those same characteristics. Furthermore, KDE indicated that they had received anecdotal evidence that some schools incorrectly interpreted the 'N/A' as an indication of "No Implementation" as opposed to an indication of "Not Applicable." It was not possible to reconcile which schools provides '0s' as an indication of "Not Applicable" and which schools provided '0s' as an indication of "No Implementation."

Because the scoring key included in the raw data file indicated that ‘0’ was to indicate not applicable, HumRRO recoded all ‘0s’ as missing data in its calculation of program review mean scores, even though it is likely that for some schools those ‘0s’ were intended to be indications of schools not implementing that characteristic (and not that the characteristic was “not applicable”). This should be considered as a limitation of the 2011–2012 program review pilot data. For future program reviews, KDE should consider hard coding the “not applicable” characteristics so that characteristics that are “not applicable” to a certain grade(s) are not presented to those grades.

Another issue with the final data file was that four schools provided multiple records—three schools had two records and one school had three records in the final data file. In all cases, the multiple records for the same school were *not* identical. AdvancEd was not able to identify which of the records for the school was “most accurate.” Consequently, none of those multiple records were deleted since it was not possible to determine which record was most accurate. This means that three schools were essentially counted twice and one school was counted three times (all based on different sets of ratings) in the overall calculations. This should also be noted as another limitation of the 2011–2012 pilot data. For future program reviews, procedures to ensure that schools are only able to submit one final, accurate record for their school for each program review should be implemented.

Finally, it appears that 12 schools did not submit any program review data for any of the content areas for their school. Consequently, the overall calculations are not based on the full population of schools. This should also be noted as a caveat/limitation of the 2011–2012 program review data.

### **Section 1: Method for Reporting Program Review Scores and Performance Cuts**

Unlike Kentucky’s more traditional test scores, which are reported to schools at the overall score level and sub-score level, the program review ratings were entered by the schools, so all schools have access to their full data at the characteristic level. KDE chose to aggregate the characteristic ratings to provide schools with results at the program-level, and to create “standard-level” scores. The standard- and program-level scores were designed to allow schools to compare their performance to other similar schools and judge their relative standing.

The calculation of standard-level and program-level scores was complicated by the nature of the schools and the data set. Schools in the data file were designated as elementary, middle, high, or combination schools. Combination schools were defined as those that did not exactly fit into the elementary, middle, or high school categories (e.g., K-12 configurations or other less traditional configurations). There was considerable variance in the ratings provided by school type. Consequently, schools were only provided comparison data from their matched type (i.e. elementary, middle, high, or combination).

Aggregate scores can be calculated in essentially two ways. The characteristic ratings can be summed and the total score reported, or the characteristic ratings can be averaged and a mean score reported. Mean scores were selected for the program reviews to limit the impact of the inconsistent use of the 0 scores. The 1, 2, and 3 characteristic ratings were averaged to generate aggregate scores at the standard-level. Scores of 0 were treated as missing data and did not contribute to the mean scores.

Next, a decision was needed for determining how to set the classification cuts for “Needs Improvement,” “Proficient,” and “Distinguished” (NI, P, and D). Initially, we investigated setting

classification cuts based on the percentile ranks on the overall mean ratings (the overall mean rating is the mean based on all characteristics for the program). For example, those schools with an overall mean score in the 90<sup>th</sup> percentile or higher could be identified as Distinguished, those with an overall mean score between the 70<sup>th</sup> percentile and less than the 90<sup>th</sup> percentile could be identified as Proficient, and those with an overall mean score less than the 70<sup>th</sup> percentile could be identified as Needs Improvement. However, KDE decided not to set the classification cuts on the overall mean scores for a couple of reasons. First, when cuts were set on the overall mean scores, this meant that for schools whose overall mean fell very near the 90<sup>th</sup> or 70<sup>th</sup> percentile, there were schools who fell on one side of the cut or the other, but whose overall mean scores differed by a trivial amount (i.e., differences in the 3<sup>rd</sup> and 4<sup>th</sup> decimal places). Another concern with setting cuts on the overall mean scores is that there are more characteristics for some standards than other standards (e.g., for PL/CS there are 52 characteristics that fall under the Curriculum & Instruction standard and there are only 16 characteristics that comprise the Professional Development & Support Services standard). Setting classification cuts on the overall mean ratings would give more weight to those standards that have more characteristics, and KDE did not want to differentially weight the importance of the standards based on the number of characteristics in each standard.

Because of the issues identified above with setting classification cuts on the overall mean scores, another method was developed that guarded against the possibility of schools being placed in different classification categories with only trivial differences in overall mean scores, and that also gave each standard equal weight in terms of its contribution to a school's overall designation as "Needs Improvement," "Proficient," or "Distinguished." Using this method, a school's mean score on *each standard* was used to determine the school's NI, P, D standing on each standard. If the school had a mean rating on a standard of less than 2.0, then the school was designated as being 'NI' on that standard. If the school had an 'NI' on two or more of the four standards, then the school received an overall 'NI' rating for their program review score for that program area. If the school had a mean rating of 2.0 or higher on three or four of the standards, then the school was designated as 'P' in that program area if the sum of the school's mean scores across the four standards was between 8.0 and less than 10.8 (an average of 2 points for each of the four standards would equal 8.). If the sum of their mean scores across the four standards was 10.8 or greater, then the school was designated as Distinguished for that program area (a perfect score for all four standards would equal 12, while an average of 2.7–2 2/3 rounded-- equals 10.8.). An example of a score report that HumRRO produced for Arts & Humanities is provided in Appendix A.

### ***Summary of Method for Reporting Program Review Scores and Performance Cuts***

At the conclusion of the 2011–2012 pilot program review, the scores reported to schools were comprised of mean scores on each standard and the overall mean score for the program. The overall mean score for the program was based on the mean of all the characteristics for the program. Classification into Needs Improvement, Proficient, and Distinguished was determined based on mean scores on the four standards for each program. Schools designated as Needs Improvement on two or more of the standards received a Needs Improvement classification for the program area regardless of total mean score. Schools designated as proficient or higher on three or four of the standards were classified as Proficient if the sum of the means on the standards was between 8.0 and less than 10.8, and schools were classified as Distinguished if the sum of the means on the standards was 10.8 or greater.

For the 2012–2013 operational program reviews (2013 reporting), HumRRO recommends that the ASSIST software system be "hard coded" so that characteristics not applicable to certain

grade levels are unavailable for entering ratings. Furthermore, HumRRO recommends that an additional quality assurance step be conducted to verify that schools are only able to submit one final set of program review ratings per program area for their school. Finally, HumRRO recommends that the scoring decisions implemented in the pilot year be re-visited in light of the operational data, which presumably should not contain the limitations of the pilot year data.

## Section 2: Quantitative Investigation of Program Review Scores

At the recommendation of the NTAPAA during its December 5, 2012 meeting with KDE (which occurred after the program review score reports were publicly released), HumRRO quantitatively investigated the reliability of the program review measures. The purpose of the quantitative analysis was two-fold. The first purpose was to verify that the “scales” on the measure are reliable, and the second purpose was to identify ways to reduce the length of the measure without adversely impacting its reliability. Because some of the subscales are quite long (e.g., the Curriculum and Instruction subscale for PL/CS has 52 items), and because scales with too many items can create problems with respondent fatigue and response biases (e.g., Anastasi, 1976), we explored ways to reduce the number of items on the subscales to improve parsimony. These investigations are described in this section of the report. It should be noted that all analyses in this report were based solely on the 2011–2012 program review data, and its supporting documents for the 2011–2012 academic year (e.g., program review rubrics). HumRRO is aware that KDE has made recent changes to the program review materials—for example, changes have been made to the rubrics for Arts and Humanities, Writing, and PL/CS<sup>1</sup>. In light of the changes that have occurred to the program reviews for the 2012–2013 academic year, we recommend that the same types of quantitative investigations performed on the 2011–2012 data also be conducted on the 2012–2013 data when it becomes available.

Each content area includes four standards (i.e., Curriculum and Instruction, Formative and Summative Assessment, Professional Development, and Administrative Support and Monitoring). These standards are further broken down into demonstrators (listed on page 3 of this report), and each demonstrator has a certain number of characteristics associated with it. In assessment or survey terminology, the standards could be thought of as scales, the demonstrators as subscales, and the characteristics as items. We will use this terminology from this point forward to describe the quantitative properties of the program review scores.

Because the program review scores were reported out by grade level for each content area (i.e., elementary school, middle school, and high school) due to considerable variance in ratings across grade levels, all analyses in this section were also conducted by grade level for each content area<sup>2</sup>. Finally, unless otherwise noted, analyses were conducted on the recoded data (i.e., items for which “not applicable” were recoded as “missing.”).

### ***Investigation of Frequency Distribution of Item Ratings***

First, we investigated the frequency distribution of ratings on the raw, un-recoded data for each item for each content area by grade level. Based on frequency distributions, we flagged items for two reasons. First, we flagged items for which a majority of schools selected the “not applicable” rating. Twenty-two items were flagged as being “not applicable” for the majority of schools in the elementary level; eight items were flagged as “not applicable” for the majority of schools in the middle school level, and two items were flagged as “not applicable” for the high school level. Those items are listed in Table 1, and the actual frequency distributions for those items appear in Appendix B. We recommend that content experts examine Table 1 and Appendix B to verify that the items they believe are “not applicable” to a particular grade level(s) are indeed being rated as “not applicable” by those grades, and also that items they believe are

---

<sup>1</sup> Updated program review rubrics can be found on the KDE website located at: <http://education.ky.gov/curriculum/pgmrev/Pages/default.aspx>

<sup>2</sup> Because this was an exploratory analysis on the pilot data and because the Combination schools made up only 9% of all schools, we did not include Combination schools in this analysis.

“applicable” to a particular grade level(s) are indeed being rated by all or nearly all schools in that grade level. Should content experts find any discrepancies in their expectations, (e.g., an item they believe is *applicable* to a particular grade level, but is being rated as *not applicable* by a substantive number of schools), then this may signal that a revision to the item is needed. One way to address this concern would be to “hard code” the items as they appear in the ASSIST tool so that items that are intended to be “not applicable” to a particular grade level are not presented to schools that include that grade level.

**Table 1. Items for Which the Majority of Schools Selected “Not Applicable”**

Content Area	Scale <sup>a</sup>	# Items	Elementary School	Middle School	High School	
Writing	CI	23	none	none	none	
	FSA	13	none	none	none	
	PD	12	none	none	none	
	AS	13	none	none	none	
AH	CI	22	none	none	none	
	FSA	13	none	none	none	
	PD	7	none	none	none	
	AS	26	AH_3_12_8 AH_3_13_4 <sup>b</sup>	none	none	
PL/CS	CI	52	PL_1_1_2			
			PL_1_1_3			
			PL_1_2_21			
			PL_1_2_23			
			PL_1_2_25		PL_1_1_3	
			PL_1_2_26		PL_1_2_21	
			PL_1_3_13		PL_1_2_23	
			PL_1_4_8		PL_1_2_24	
			PL_1_4_9		PL_1_2_26	PL_1_2_24
			PL_1_4_10		PL_1_4_9	PL_1_4_10
			PL_1_5_3			
			PL_1_6_2			
			FSA	17	PL_1_6_6	none
PD	16	PL_1_9_2				
		PL_1_9_5		none		
		PL_1_12_5		none		
		PL_1_13_1				
AS	30	PL_1_13_5		PL_1_12_5		
		PL_1_13_6		PL_1_13_6		
				none		

Notes. <sup>a</sup>CI = Curriculum and Instruction; FSA = Formative and Summative Assessment; PD = Professional Development; AS = Administrative Leadership and Support.

<sup>b</sup>Codes are the codes provided by AdvancEd in the raw data file.

AH\_3\_12\_8 = To what extent does the school leadership ensure arts courses are not scheduled in conflict with single-section required content area courses?

AH\_3\_13\_4 = To what extent does school leadership ensure the use of various data to inform decisions on arts instructional programs, class offerings and staffing?

PL\_1\_1\_2 = To what extent does the school provide advising programs that include components of the ILP?

PL\_1\_1\_3 = To what extent does the school provide PL/CS interdisciplinary and advanced courses?

PL\_1\_2\_21 = To what extent are career pathways implemented in the state's 14 identified Career Clusters for high school students?

PL\_1\_2\_23 = To what extent are articulation agreements and dual credit agreements a part of career pathways at the high school level?

PL\_1\_2\_24 = To what extent are all students introduced to the 14 career clusters at the elementary level?

PL\_1\_2\_25 = To what extent is a formalized plan for introducing and using the ILP in place for the 6th grade and beyond?

PL\_1\_2\_26 = To what extent do all high school students select and note in their ILP's at least four courses related to their career major and one of the state's 14 Career Clusters?

PL\_1\_3\_13 = To what extent does the school integrate the ILP in some course to use as a research tool for career choices?

PL\_1\_4\_8 = To what extent do teachers & students work together using formative and assessment data to formulate student ILP's?

PL\_1\_4\_9 = To what extent are a variety of work based learning opportunities provided across Career Pathways and majors?

PL\_1\_4\_10 = To what extent is the ILP development for all 6th grade students?

PL\_1\_5\_3 = To what extent are formalized plans in place to monitor the completion of the students' ILP?

PL\_1\_6\_2 = To what extent are middle and high school students required to successfully pass a health education course?

PL\_1\_6\_6 = To what extent are middle and high school students required to pass a physical education course?

PL\_1\_9\_2 = To what extent does the school provide training to staff regarding the use of the ILP to help students select course offerings and develop career goals?

PL\_1\_9\_5 = To what extent does the school provide professional development opportunities to work with postsecondary partners?

PL\_1\_12\_5 = To what extent does the high school utilize the Career & Technical Education and Kentucky Schools Facility Planning Manuals for program planning?

PL\_1\_13\_1 = To what extent does the middle and high school use data from ILP reports to determine PL/CS course offerings?

PL\_1\_13\_5 = To what extent does the school have consistency between career information in the ILP and the Individualized Education Plan (IEP) for students with disabilities?

PL\_1\_13\_6 = To what extent does the high school have measures in place to check for and prevent duplication of coursework between secondary and postsecondary levels?

Next, we investigated the frequency distribution of ratings to identify items with little or no variance in ratings (note that items that were flagged as being “not applicable” by a substantive number of schools were *not* included in this investigation). If there was little variance in ratings—operationalized as 80% or more of schools providing the same rating—then, items were flagged (see Table 2). If an item has little variance (i.e., nearly all schools provide the same rating), then the item may not be very informative. For example, 90% of high schools indicated that their school was “proficient” when asked, “To what extent are students required to successfully pass a health education course?” (PL\_1\_6\_2), and “To what extent are students required to pass a physical education course?” (PL\_1\_6\_6). Because the overwhelming majority of schools provided the same rating, these items contribute very little information with regards to making determinations about the relative level of a school’s performance on the “Formative and Summative Assessment” (FSA) standard of the PL/CS program review. Furthermore, because these items have little variance, they do not co-vary with other items on the FSA standard in a meaningful way. The items flagged in Table 2 should be further evaluated by content experts to determine if the items should be dropped, revised, or retained “as is.” If an item is retained “as is” there should be some indication that more variance is expected in the future or some other compelling reason for retaining the item should be documented.

**Table 2. Items Flagged Due to Little Variance in Ratings**

Content Area	Scale	Items for Which 80% or More of Schools Provided the Same Rating			
		Elementary School	Middle School	High School	
Writing	CI	WR_2_3_6 (80% proficient)	WR_2_3_6 (82% proficient)	none	
		WR_2_4_2 (81% proficient)	WR_2_4_1 (81% proficient)		
	FSA	none	none	none	
	PD	none	none	none	
AH	CI	AH_3_1_2 (81% proficient)	none	none	
		AH_3_6_3 (86% N.I.)	none	none	
	FSA	none	none	none	
	PD	none	none	none	
PL/CS	CI	PL_1_2_15 (82% N.I.)	none	PL_1_5_2 (81% proficient)	
				PL_1_6_2 (90% proficient)	
	FSA	none	PL_1_5_2 (81% proficient)	PL_1_6_7 (81% proficient)	PL_1_6_3 (84% proficient)
					PL_1_6_6 (90% proficient)
	PD	none	PL_1_9_1 (81% N.I.)	PL_1_9_5 (82% N.I.)	PL_1_6_7 (81% proficient)
					PL_1_13_1 (88% N.I.)
	AS	none	PL_1_13_7 (81% N.I.)	none	none
					none

Notes. Codes are the codes provided by AdvancEd in the raw data file.

WR\_2\_3\_6 = To what extent do students' communication strategies demonstrate their understanding of appropriate audience, form and purpose?

WR\_2\_4\_1 = To what extent do students craft an appropriate communication for a specific discipline and purpose?

WR\_2\_4\_2 = To what extent do students build on ideas in their writing and articulate their own ideas when writing?

AH\_3\_1\_2 = To what extent does the school ensure quality instruction for diverse populations?

AH\_3\_6\_3 = To what extent does the school ensure that students develop their own rubrics for creating, performing and responding in the arts?

PL\_1\_2\_15 = To what extent does curriculum include current information on product safety and values?

PL\_1\_2\_21 = To what extent are the career pathways implemented in the state's 14 identified Career Clusters for high school students? (HIGH SCHOOL)

PL\_1\_5\_2 = To what extent do PL/CS assessment measures respond to a variety of learning styles and abilities?

PL\_1\_6\_2 = To what extent are middle and high school students required to successfully pass a health education course?

PL\_1\_6\_3 = To what extent can students demonstrate the knowledge and skills addressed in local, state, and national PL/CS standards?

PL\_1\_6\_6 = To what extent are middle and high school students required to pass a physical education course?

PL\_1\_6\_7 = To what extent do PL/CS teachers develop scoring guides and/or rubrics?

PL\_1\_9\_1 = To what extent do PL/CS teachers collaborate with community partners to participate in various Advisory Committees?

PL\_1\_9\_5 = To what extent does the school provide professional development opportunities to work with postsecondary partners?

PL\_1\_13\_1 = To what extent does the middle and high school use data from ILP reports to determine PL/CS course offerings?

PL\_1\_13\_7 = To what extent is the school's wellness policy developed and reviewed annually, and included in the CSIP?

### Scale Level Analyses

Next, we calculated Cronbach's coefficient alpha as an estimate of internal consistency reliability. When individual items all tap the same concept (i.e., typically used as an indication of scale unidimensionality), then the scale will be more reliable. Cronbach's alphas of .80 or more are generally considered to indicate strong internal consistency reliability. The "not applicable" ratings were recoded as missing in this analysis. Because the calculation of Cronbach's coefficient alpha requires complete data, we dropped items for which a majority of schools had missing data (i.e., those items listed in Table 1).

After identifying which items to drop due to missing data, we identified cases (i.e., schools) to drop due to incomplete data. This process of first dropping *items* for which a majority of schools had missing data followed by dropping *schools* with missing data gave us the most complete

dataset with the greatest number of data points on which to calculate internal consistency reliability estimates.

Once we had effectively dealt with the missing data, we were able to calculate coefficient alpha. Because several of the subscales (i.e., demonstrators) have three or fewer items—with two subscales actually only having one item—we feel that the appropriate level for reporting results and for calculating the internal consistency reliability estimates is at the level of the scales (i.e., standards-level). This is also akin to the level at which scores are reported on the K-PREP. However, we also calculated coefficient alpha at the level of the subscales (i.e., demonstrators) so that KDE could get a sense of the internal consistency (or lack thereof) of the items comprising the subscales<sup>3</sup>. We will first discuss findings at the scale- or standards-level, and then follow-up with some additional investigation at the level of the subscales.

The internal consistency reliability estimates for Writing, Arts & Humanities, and PL/CS are presented in Tables 3, 4, and 5, respectively. In all cases, the reliability estimates at the scale level (i.e., those values in bold text in Tables 3, 4, and 5) were above the .80 level, indicating acceptable levels of internal consistency reliability on all the scales for all grade levels and content areas.

---

<sup>3</sup> No reliability estimates for the following subscales were possible due to these subscales only having a single item: Policies and Monitoring (for Writing) and Teacher Leadership (for Arts & Humanities).

**Table 3. Writing Internal Consistency Reliability Estimates for Scales and Subscales by Grade Level**

Grade Level												
Scale <sup>a</sup>	# Scale items	Sub-scale <sup>b</sup>	# Sub-scale items	Elementary School (n = 671)			Middle School (n = 217)			High School (n = 199)		
				# items dropped <sup>c</sup>	# Schools w/complete data <sup>d</sup>	$\alpha$	# items dropped	# Schools w/complete data	$\alpha$	# items dropped	# Schools w/complete data	$\alpha$
<b>CI</b>	<b>23</b>			<b>0</b>	<b>646</b>	<b>.92</b>	<b>0</b>	<b>213</b>	<b>.93</b>	<b>0</b>	<b>196</b>	<b>.94</b>
		SA	2	0	670	.69	0	215	.70	0	198	.73
		ARC	6	0	663	.79	0	214	.80	0	196	.81
		IS	11	0	657	.86	0	214	.87	0	198	.90
		SP	4	0	663	.74	0	215	.79	0	198	.81
<b>FSA</b>	<b>13</b>			<b>0</b>	<b>661</b>	<b>.90</b>	<b>0</b>	<b>208</b>	<b>.90</b>	<b>0</b>	<b>197</b>	<b>.91</b>
		VA	5	0	667	.79	0	214	.80	0	197	.81
		ESL	4	0	667	.74	0	210	.70	0	197	.73
		RA	4	0	665	.77	0	213	.75	0	197	.81
<b>PD</b>	<b>12</b>			<b>0</b>	<b>663</b>	<b>.92</b>	<b>0</b>	<b>215</b>	<b>.93</b>	<b>0</b>	<b>197</b>	<b>.92</b>
		Plan	3	0	668	.86	0	215	.87	0	197	.83
		Part.	4	0	668	.81	0	215	.85	0	198	.80
		TL	5	0	667	.85	0	215	.87	0	198	.88
<b>AS</b>	<b>13</b>			<b>0</b>	<b>654</b>	<b>.92</b>	<b>0</b>	<b>210</b>	<b>.92</b>	<b>0</b>	<b>191</b>	<b>.93</b>
		SV	5	0	666	.82	0	215	.83	0	197	.85
		TR	4	0	662	.88	0	212	.85	0	194	.88
		PM <sup>e</sup>	1	--	--	--	--	--	--	--	--	--
		PL	3	0	670	.78	0	215	.78	0	197	.82

Notes. <sup>a</sup>CI = Curriculum and Instruction; FSA = Formative and Summative Assessment; PD = Professional Development; AS = Administrative Support

<sup>b</sup>SA = Student Access; ARC = Aligned and Rigorous Curriculum; IS = Instructional Strategies; SP = Student Performance; VA = Variety of Assessment; ESL = Expectations for Student Learning; RA = Response to Assessment; Plan = Planning; Part. = Participation; TL = Teacher Leadership; SV = Shared Vision; TR = Time and Resources; PM = Policies and Monitoring; PL = Principal Leadership.

<sup>c</sup>Items for which the majority of schools had missing data were dropped from the reliability analysis.

<sup>d</sup>Only schools with complete data are used in the calculation of the reliability estimate.

<sup>e</sup>There is only one item on the Policies and Monitoring subscale; consequently, it is not possible to compute coefficient alpha.

**Table 4. Arts & Humanities Internal Consistency Reliability Estimates for Scales and Subscales by Grade Level**

Grade Level												
Scale <sup>a</sup>	# Scale items	Sub-scale <sup>b</sup>	# Sub-scale items	Elementary School (n = 671)			Middle School (n = 217)			High School (n = 199)		
				# items dropped <sup>c</sup>	# Schools w/complete data <sup>d</sup>	α	# items dropped	# Schools w/complete data	α	# items dropped	# Schools w/complete data	α
<b>CI</b>	<b>22</b>			<b>0</b>	<b>605</b>	<b>.93</b>	<b>0</b>	<b>205</b>	<b>.94</b>	<b>0</b>	<b>193</b>	<b>.93</b>
		SA	4	0	667	.75	0	214	.81	0	198	.72
		ARC	6	0	640	.82	0	205	.85	0	194	.78
		IS	6	0	661	.79	0	214	.84	0	198	.85
		SP	6	0	636	.83	0	214	.87	0	199	.86
<b>FSA</b>	<b>13</b>			<b>0</b>	<b>597</b>	<b>.89</b>	<b>0</b>	<b>205</b>	<b>.90</b>	<b>0</b>	<b>188</b>	<b>.90</b>
		VA	3	0	659	.78	0	214	.78	0	197	.78
		ESL	3	0	639	.57	0	211	.64	0	197	.56
		RA	7	0	607	.81	0	207	.84	0	190	.84
<b>PD</b>	<b>7</b>			<b>0</b>	<b>648</b>	<b>.86</b>	<b>0</b>	<b>209</b>	<b>.83</b>	<b>0</b>	<b>197</b>	<b>.85</b>
		Plan	4	0	656	.76	0	210	.73	0	197	.71
		Part.	2	0	661	.74	0	212	.66	0	199	.77
		TL <sup>e</sup>	1	--	--	--	--	--	--	--	--	--
<b>AS</b>	<b>26</b>			<b>2</b>	<b>505</b>	<b>.93</b>	<b>0</b>	<b>190</b>	<b>.93</b>	<b>0</b>	<b>183</b>	<b>.93</b>
		SV	4	0	660	.84	0	212	.86	0	196	.85
		TR	10	1	629	.86	0	204	.85	0	194	.86
		PM	4	1	530	.72	0	197	.73	0	189	.75
		PL	6	0	646	.81	0	211	.84	0	194	.82

Notes. <sup>a</sup>CI = Curriculum and Instruction; FSA = Formative and Summative Assessment; PD = Professional Development; AS = Administrative Support

<sup>b</sup>SA = Student Access; ARC = Aligned and Rigorous Curriculum; IS = Instructional Strategies; SP = Student Performance; VA = Variety of Assessment; ESL = Expectations for Student Learning; RA = Response to Assessment; Plan = Planning; Part. = Participation; TL = Teacher Leadership; SV = Shared Vision; TR = Time and Resources; PM = Policies and Monitoring; PL = Principal Leadership.

<sup>c</sup>Items for which the majority of schools had missing data were dropped from the reliability analysis.

<sup>d</sup>Only schools with complete data are used in the calculation of the reliability estimate.

<sup>e</sup>There is only one item on the Teacher Leadership subscale; consequently, it is not possible to compute coefficient alpha.

**Table 5. PL/CS Internal Consistency Reliability Estimates for Scales and Subscales by Grade Level**

Scale <sup>a</sup>	# Scale items	Sub-scale <sup>b</sup>	# Sub-scale items	Grade Level								
				Elementary School (n = 671)			Middle School (n = 217)			High School (n = 199)		
				# items dropped <sup>c</sup>	# Schools w/complete data <sup>d</sup>	$\alpha$	# items dropped	# Schools w/complete data	$\alpha$	# items dropped	# Schools w/complete data	$\alpha$
<b>CI</b>	<b>52</b>			<b>10</b>	<b>343</b>	<b>.94</b>	<b>6</b>	<b>151</b>	<b>.95</b>	<b>2</b>	<b>199</b>	<b>.95</b>
		SA	3	2	-- <sup>e</sup>	-- <sup>e</sup>	1	213	.27	0	198	.59
		ARC	26	4	449	.89	4	189	.92	1	117	.92
		IS	13	1	523	.84	0	197	.86	0	192	.89
		SP	10	3	521	.73	1	177	.77	1	192	.75
<b>FSA</b>	<b>17</b>			<b>3</b>	<b>604</b>	<b>.88</b>	<b>0</b>	<b>210</b>	<b>.87</b>	<b>0</b>	<b>198</b>	<b>.88</b>
		VA	3	1	662	.61	0	214	.56	0	199	.60
		ESL	9	2	642	.76	0	201	.75	0	194	.75
		RA	5	0	632	.78	0	215	.79	0	199	.82
<b>PD</b>	<b>16</b>			<b>2</b>	<b>550</b>	<b>.88</b>	<b>0</b>	<b>201</b>	<b>.83</b>	<b>0</b>	<b>194</b>	<b>.89</b>
		Plan	8	0	646	.81	0	215	.81	0	199	.84
		Part.	6	2	569	.72	0	134	.70	0	188	.76
		TL <sup>e</sup>	2	0	649	.68	0	213	.56	0	198	.69
<b>AS</b>	<b>30</b>			<b>4</b>	<b>432</b>	<b>.91</b>	<b>2</b>	<b>215</b>	<b>.90</b>	<b>0</b>	<b>199</b>	<b>.92</b>
		SV	5	0	664	.86	0	209	.86	0	195	.87
		TR	12	1	618	.82	1	207	.81	0	195	.81
		PM	8	3	465	.68	1	184	.66	0	178	.64
		PL	5	0	610	.82	0	202	.78	0	197	.83

Notes. <sup>a</sup>CI = Curriculum and Instruction; FSA = Formative and Summative Assessment; PD = Professional Development; AS = Administrative Support

<sup>b</sup>SA = Student Access; ARC = Aligned and Rigorous Curriculum; IS = Instructional Strategies; SP = Student Performance; VA = Variety of Assessment; ESL = Expectations for Student Learning; RA = Response to Assessment; Plan = Planning; Part. = Participation; TL = Teacher Leadership; SV = Shared Vision; TR = Time and Resources; PM = Policies and Monitoring; PL = Principal Leadership.

<sup>c</sup>Items for which the majority of schools had missing data were dropped from the reliability analysis.

<sup>d</sup>Only schools with complete data are used in the calculation of the reliability estimate.

<sup>e</sup>There are only 3 items on the CI\_SA scale, and two of those items were dropped from the elementary school level due to a majority of schools having missing data on those two items. Consequently, coefficient alpha could not be computed on CI\_SA at the elementary level.

Having established acceptable levels of internal consistency reliability on the scales, the next step was to conduct quantitative analyses related to data reduction. Because some of the scales are quite long (e.g., the Curriculum and Instruction scale for PL/CS has 52 items), and because scales with too many items can create problems with respondent fatigue and response biases (e.g., Anastasi, 1976), we explored ways to reduce the number of items on the scales to improve parsimony.

First, we investigated the “alpha if item deleted” for each scale. If dropping an item from a scale improved the coefficient alpha by .01 or more, then those items were flagged as items for further review by content experts. Next, we investigated each item’s correlation with the corrected item-total correlation for that scale. If items correlated at .35 or less with the corrected item-total correlation, then those items were also flagged for further review. Both of these criteria serve as indications that items may not be tapping into the same concept as other items on the scale and/or that respondents may be misinterpreting the items. For example, item PL\_1\_6\_6 (*To what extent are middle and high schools required to pass a physical education course?*) appears to be an item that would elicit a “yes” or “no” response as opposed to an item that should be rated on a Likert-type rating scale. Content experts should investigate the flagged items for issues such as this and make a final determination as to whether to drop, revise, or retain the flagged items. Note that none of the items on the scales for Writing nor any of the items on the scales for Arts and Humanities were flagged using these criteria. However, several items on the scales for PL/CS were flagged based on these criteria. The flagged items for PL/CS are presented in Table 6.

**Table 6. PL/CS Items Flagged Based on Coefficient Alpha If Item Deleted and Item Total Correlation (Scale Level Analysis)**

Content Area	Scale <sup>a</sup>	Elementary School		Middle School		High School	
		Items that when dropped improve the internal consistency estimate (by .01 or more)	Items for which the corrected item-total correlation is $r = .35$ or less	Items that when dropped improve the internal consistency estimate (by .01 or more):	Items for which the corrected item-total correlation is $r = .35$ or less	Items that when dropped improve the internal consistency estimate (by .01 or more):	Items for which the corrected item-total correlation is $r = .35$ or less
PL/CS	CI	none	PL_1_2_11	none	PL_1_2_12	none	PL_1_2_12
			PL_1_2_24				PL_1_2_25
	FSA	none	PL_1_4_5	none	PL_1_3_4	none	PL_1_4_5
			PL_1_4_7		PL_1_6_4		PL_1_6_2
	PD	none	PL_1_6_4	none	PL_1_6_6	none	PL_1_6_6
AS	none	none	PL_1_9_5	none	PL_1_9_5	none	none
			PL_1_13_1		PL_1_13_2		
			PL_1_13_7		PL_1_13_8		none

Notes. <sup>a</sup>CI = Curriculum and Instruction; FSA = Formative and Summative Assessment; PD = Professional Development; AS = Administrative Leadership and Support.

<sup>b</sup>Codes are the codes provided by AdvancEd in the raw data file.

PL\_1\_2\_11 = To what extent do the students implement Frequency, Intensity, Type, Time (FITT) Plans?

PL\_1\_2\_12 = To what extent do students participate in a comprehensive school-based physical activity program?

PL\_1\_2\_24 = To what extent are all students introduced to the 14 career clusters at the elementary level?

PL\_1\_2\_25 = To what extent is a formalized plan for introducing and using the ILP in place for the 6<sup>th</sup> grade and beyond?

PL\_1\_3\_4 = To what extent does the school utilize Advisory Committees?

PL\_1\_4\_5 = To what extent are students provided opportunities to earn health related certificates?

PL\_1\_4\_7 = To what extent do students demonstrate the knowledge and skills associated the National Standards for Physical Education?

PL\_1\_6\_2 = To what extent are middle and high school students required to successfully pass a health education course?

PL\_1\_6\_4 = To what extent are students active during physical education lesson time?

PL\_1\_6\_5 = To what extent do students participate in physical education and/or physical activity?

PL\_1\_6\_6 = To what extent are middle and high schools required to pass a physical education course?

PL\_1\_9\_5 = To what extent does the school provide professional development opportunities to work with postsecondary partners?

PL\_1\_12\_6 = To what extent do the school's PL/CS facilities meet industry and classroom standards for size, safety, and accessibility?

PL\_1\_13\_1 = To what extent does the middle and high school use data from ISP reports to determine PL/CS course offerings?

PL\_1\_13\_2 = To what extent does the school use Advisory Committees to ensure quality PL/CS programming and policies?

PL\_1\_13\_7 = To what extent is the schools wellness policy developed and reviewed annually, and included in the CSIP?

PL\_1\_13\_8 = To what extent is the school implementing the district wellness policy via a local wellness policy?

Next, we investigated the correlation matrix for each scale for each content area by grade level. First, we used the correlation matrix for each scale to identify pairs of items that correlated at  $r = .80$  or more. Such a strong correlation would indicate that the items are largely redundant with one another, and that one of the items could be dropped without losing meaningful information. Items for which a majority of schools had missing data (i.e., those listed in Table 1) were not included in this analysis. For Writing, there was only one pair of items that was flagged due to a high correlation ( $r = .81$ ) between WR\_2\_8\_1 (To what extent does the professional development plan support the school's literacy vision?) and WR\_2\_8\_2 (To what extent does the professional development plan meet the needs of the school and the long-term communications program vision?) for the Professional Development (PD) scale at the elementary grade level. Typically, when a pair of items is very highly correlated and we have concerns about the lengthiness of the scale, then we would recommend dropping one of the items in order to reduce the length of the scale. However, because the PD scale contains the fewest items of all the scales, data reduction is not a major concern with the PD scale. No other item pairs for any of the scales for any of the content areas were flagged due to correlations of  $r = .80$  or more.

Next, we used the correlation matrices for each scale to identify pairs of items that correlated at  $r = .15$  or less. As mentioned above, all of the scales have coefficient alphas greater than  $.80$ . However, it is possible to get a reliable scale using items with reasonably poor internal consistency if the scale contains enough items. For example, 10 items that have an average inter-item correlation of only  $.2$  will still produce a scale with a reliability of approximately  $.71$ . Similarly, if the average correlation among five items is  $.5$ , the alpha coefficient will be approximately  $0.83$ , but if the number of items is 10—with the same average correlation—the alpha coefficient will be approximately  $0.91$ . Because each of the scales for each content area is fairly lengthy (with the exception of the Professional Development scale for Arts and Humanities, which only has 7 items), we investigated the correlation matrices for each scale to investigate whether the high coefficient alphas could have resulted from having many items on the scale with relatively weak inter-correlations. For Writing, none of the item pairs on any of the scales were flagged for having inter-correlations of  $r = .15$  or less; this provides support that the coefficient alphas for the scales for Writing are not simply a function of the lengthiness of the scales.

For Arts and Humanities, for the elementary school level, the Administrative Support and Leadership (AS) scale had a pair of items with a very low inter-correlation. Item AH\_3\_12\_5 (*To what extent does the school leadership ensure that arts classrooms are adequate for the specific arts disciplines?*) and item AH\_3\_13\_5 (*To what extent does the school leadership ensure differentiated learning opportunities for students identified as gifted and talented in the arts?*) are correlated at  $r = .14$ . This is an indication that these items may not be tapping into the same concept. In looking at the inter-correlations of each of these items with the other items on the AS scale, we see that AH\_3\_13\_5 tends to have weaker correlations than AH\_3\_12\_5 with the other items on the scale. Consequently, based on this quantitative evidence, AH\_3\_13\_5 appears to be the better candidate to drop from the scale. However, the final determination should be made based on content experts' reviews of the item. None of the items were flagged in Arts and Humanities at the middle school level due to having inter-correlations of  $r = .15$  or less. At the high school level, three pairs of items were flagged for low inter-correlations on the Curriculum and Instruction (CI) scale (see Table 7). Items AH\_3\_1\_1 (*To what extent does the school leadership ensure that the arts are included in the school vision?*) and AH\_3\_2\_3 (*To what extent does the school ensure instruction addressing the interrelationships of the arts?*) are weakly correlated at  $r = .11$ . Item AH\_3\_1\_1 is also weakly correlated with AH\_3\_3\_3 (*To what extent does the school ensure teachers incorporate student/peer reviews of students'*

products and performances?) at  $r = .14$ . Item AH\_3\_1\_1 tends to be fairly weakly correlated with other items on the CI scale as well; therefore, we recommend that content experts further review AH\_3\_1\_1 to determine if it should be dropped. Item AH\_3\_2\_1 (*To what extent does the school ensure that an aligned arts curriculum is articulated and meets accepted instructional standards in the arts?*) and AH\_3\_3\_6 (*To what extent does the school ensure the utilization of outside resources for instructional purposes in the arts?*) are also weakly correlated at  $r = .14$ . Neither of these items stands out as being more poorly correlated with other items on the CI scale than the other. Consequently, there is no strong quantitative evidence to support dropping one item over the other. For the Administrative Leadership and Support Scale (AS), two pairs of items are weakly correlated. The first pair is AH\_3\_11\_1 (*To what extent does the school leadership ensure that the arts are included in the school vision?*) and AH\_3\_13\_3 (*To what extent does school leadership ensure adequate instructional assistance to arts teachers?*) ( $r = .12$ ), and the second pair is AH\_3\_13\_3 and AH\_3\_11\_2 (*To what extent does the school leadership ensure that the school vision supports learning to accepted arts instructional standards by all students?*) ( $r = .09$ ). Because item AH\_3\_13\_3 is flagged in both item pairs, this item may be the better candidate for dropping from the scale. However, content experts should make the final determination.

Finally, for PL/CS there were too many item pairs that were flagged due to weak inter-item correlations on each scale for each grade level to report in a table. As such, there appears to be accumulating quantitative evidence that the PL/CS measure could benefit from substantive revision.

**Table 7. Item Pairs Flagged Based on Weak Inter-Item Correlations (Scale Level Analysis)**

Content Area	Standard <sup>a</sup>	Elementary School	Middle School	High School	
<b>Writing</b>					
<b>No Items Flagged Based on this Criterion For Writing</b>					
A&H	CI	none	none	AH_3_1_1 & 3_2_3 ( $r = .11$ ) AH_3_1_1 & 3_3_3 ( $r = .14$ ) AH_3_2_1 & 3_3_6 ( $r = .14$ )	
	FSA	none	none	none	
	PD	none	none	none	
	AS	AH_3_12_5 & 3_13_5 ( $r = .14$ )	none	none	AH_3_11_1 & 3_13_3 ( $r = .12$ ) AH_3_11_2 & 3_13_3 ( $r = .09$ )
<b>PL/CS</b>		Too many items to list			

Notes. <sup>a</sup>CI = Curriculum and Instruction; FSA = Formative and Summative Assessment; PD = Professional Development; AS = Administrative Leadership and Support.

AH\_3\_1\_1 = To what extent does the school ensure that all students have equitable access for instruction in all four arts disciplines?

AH\_3\_2\_1 = To what extent does the school ensure that an aligned arts curriculum is articulated and meets accepted instructional standards in the arts?

AH\_3\_2\_3 = To what extent does the school ensure instruction addressing the interrelationships of the arts?

AH\_3\_3\_3 = To what extent does the school ensure teachers incorporate student/peer reviews of students' products and performances?

AH\_3\_3\_6 = To what extent does the school ensure the utilization of outside resources for instructional purposes in the arts?

AH\_3\_11\_1 = To what extent does the school leadership ensure that the arts are included in the school vision?

AH\_3\_11\_2 = To what extent does the school leadership ensure that the school vision supports learning to accepted arts instructional standards by all students?

AH\_3\_12\_5 = To what extent does the school leadership ensure that arts classrooms are adequate for the specific arts disciplines?

AH\_3\_13\_3 = To what extent does school leadership ensure adequate instructional assistance to arts teachers?

AH\_3\_13\_5 = To what extent does the school leadership ensure differentiated learning opportunities for students identified as gifted and talented in the arts?

### ***Follow-up Analyses on Subscales***

To reiterate, we believe that the scale level is the appropriate level at which results should be reported (the 2011–2012 score reports were indeed reported out at the standard (scale) level). However, to provide some additional insight on the functioning of the measure, to provide guidance on which scales could be shortened, and to provide further guidance on which areas of the measure need the most substantive review and revision, we further investigated the properties of the subscales (see Tables 4, 5, and 6).

First, from a data reduction standpoint, many of the subscales have more than 10 items and have coefficient alphas of .80 or more. Consequently, these would be scales to target for data reduction. The subscales that meet this criterion are: Instructional Strategies (for Writing), Time and Resources (for Arts & Humanities), Aligned and Rigorous Curriculum (for PL/CS), Instructional Strategies (for PL/CS), and Time and Resources (for PL/CS). While the investigation of the scale level correlation matrices revealed very little instances of redundant items (i.e., inter-time correlations of  $r = .80$  or more), these are the best candidates for subscales where an item or two could be dropped without having a notable negative impact on reliability.

Likely of greater concern are those subscales that have approximately *six or more items*, and that still have alphas less than .80. This is an indication that these items may not be tapping into the same construct. None of the subscales for Writing nor for Arts & Humanities are subject to this concern. There are four subscales for PL/CS that are subject to this concern. They are: Student Performance, Expectations for Student Learning, Participation, and Policies and Monitoring. Again, there is accumulating evidence that the PL/CS measure could benefit from substantive revision, particularly for the four subscales listed here.

### ***Exploratory Principal Components Analysis for PL/CS***

The quantitative evidence for Writing and for Arts and Humanities, based on the 2011–2012 data, supports the use of the items with results reported out at the scale level. However, the quantitative evidence for the PL/CS measure indicates that the reasonably high alpha coefficients on the PL/CS scales are largely due to the length of the scales and not the homogeneity of the items on the scale. Consequently, the PL/CS measure could benefit from additional review and revision by content experts.

As an exploratory analysis, we conducted a principal components analysis for the items on the PL/CS measure in order to see if a reasonable component solution would emerge<sup>4</sup>. We used the varimax (orthogonal) rotation method to simplify the interpretation of the extracted factors. Based on eigenvalues greater than one, 22 factors were extracted. The rotated component matrix did not demonstrate a simple structure. That is, the PL/CS items had relatively high factor loadings on multiple components (as opposed to just one), and some components had only one or two items loading on them. Consequently, the exploratory principal components analysis did not provide clarity on an alternative component structure for PL/CS, and further re-iterated the need for a review and revision of the PL/CS measure by content experts.

---

<sup>4</sup> Missing values were imputed with the mean.

***Summary of Quantitative Investigation of the Program Review Scores***

The quantitative investigation of the program review measures supports the use of the Writing and the Arts & Humanities scores with results reported out at the scale level by grade. Minor revisions and tweaks of items on these scales by content experts are recommended (see items flagged in the tables and Appendices of this report). The items on the PL/CS measure, however, would benefit from substantive review and revision by content experts. In its current form, the scales on the PL/CS measure have reasonably high coefficient alphas; however, those values are largely an artifact of the length of the scales and not due to the homogeneity of the items on the scales.

## References

- Anastasi, A. (1976). Psychological testing, 4th ed. New York: Macmillan.
- Kentucky General Assembly. (2009). Senate Bill 1, KRS 158.6453(1)(i).



### Appendix A: Example 2011–2012 Score Report from Arts/Humanities

CODE	District	School	Type*	AH Curriculum & Instruction Mean	AH Formative & Summative Mean	AH Professional Development Mean	AH Administrative Support Mean	Overall AH Sum	Overall AH Mean	AHCI_ind	AHFS_ind	AHPD_ind	AHAS_ind	AH Ind_Total	AH Indicator 2	AH Total	AH Total * Ind_2	AH Performance Category	AH Missing Data Flag
006010	Anchorage	Anchorage Independent Pu	C	1.590909	1.615385	2.142857	1.880000	118	1.761194	0	0	1	0	1	0	7.229151	0	NI	
013011	Augusta In	Augusta Independent Schoo	C	1.681818	1.538462	1.714286	1.461538	107	1.573529	0	0	0	0	0	0	6.396104	0	NI	
016011	Barbourvil	Barbourville City School	C	0.000000	0.000000	0.000000	0.000000	0	0.000000	0	0	0	0	0	0	0	0	NI	Missing Data
162020	Fairview Ir	Fairview High School	C	1.772727	1.615385	2.000000	1.653846	117	1.720588	0	0	1	0	1	0	7.041958	0	NI	
165018	Fayette Cc	Scapa At Bluegrass	C	2.636364	2.153846	1.428571	2.038462	149	2.191176	1	1	0	1	3	1	8.257243	8.257243	P	
175020	Floyd Cour	Allen Elementary School	C	1.727273	1.461538	1.714286	1.846154	117	1.720588	0	0	0	0	0	0	6.749251	0	NI	
236030	Harlan Ind	Harlan High School	C	1.818182	1.615385	2.142857	2.192308	133	1.955882	0	0	1	1	2	0	7.768732	0	NI	
245010	Hart Coun	Bonnieville Elementary Sch	C	2.000000	1.846154	2.000000	2.269231	141	2.073529	1	0	1	1	3	1	8.115385	8.115385	P	
245030	Hart Coun	Cub Run Elementary Schoo	C	2.090909	1.538462	1.142857	2.400000	134	2.000000	1	0	0	1	2	0	7.172228	0	NI	
245050	Hart Coun	Legrande Elementary Schoo	C	1.772727	1.846154	2.142857	1.730769	123	1.808824	0	0	1	0	1	0	7.492507	0	NI	
245070	Hart Coun	Memorial Elementary Scho	C	1.818182	1.923077	1.571429	2.153846	132	1.941176	0	0	0	1	1	0	7.466534	0	NI	
245090	Hart Coun	Munfordville Elementary Sc	C	2.136364	2.000000	2.857143	2.080000	145	2.164179	1	1	1	1	4	1	9.073507	9.073507	P	
261050	Hickman C	Hickman County High Schoo	C	1.545455	1.461538	1.428571	1.192308	94	1.382353	0	0	0	0	0	0	5.627872	0	NI	
265006	Hopkins Cc	West Hopkins School	C	1.136364	1.384615	1.571429	1.076923	82	1.205882	0	0	0	0	0	0	5.169331	0	NI	
165007	Fayette Cc	Ashland Elementary School	E	2.500000	2.153846	1.428571	2.500000	153	2.318182	1	1	0	1	3	1	8.582417	8.582417	P	
165012	Fayette Cc	Veterans Park Elementary S	E	1.909091	1.769231	1.857143	2.208333	131	1.984848	0	0	0	1	1	0	7.743798	0	NI	
165032	Fayette Cc	Cassidy Elementary School	E	3.000000	3.000000	2.857143	3.000000	203	2.985294	1	1	1	1	4	1	11.85714	11.85714	D	
165035	Fayette Cc	Deep Springs Elementary Sc	E	1.727273	1.538462	2.000000	1.708333	113	1.712121	0	0	1	0	1	0	6.974068	0	NI	

Notes. \*Type: C = Combination school, E = Elementary school, M = Middle school, H = High school.

AH Curriculum & Instruction Mean = mean of all the ratings for the 22 characteristics comprising the Curriculum & Instruction standard (with NA ratings coded as missing).

AH Formative & Summative Mean = mean of all the ratings for the 13 characteristics comprising the Formative & Summative Assessment standard (with NA ratings coded as missing).

AH Professional Development Mean = mean of all the ratings for the 7 characteristics comprising the Professional Development standard (with NA ratings coded as missing).

AH Administrative Support Mean = mean of all the ratings for the 26 characteristics comprising the Administrative Support standard (with NA ratings coded as missing).

Overall AH Sum = Sum of all the ratings across all characteristics for Arts & Humanities.

Overall AH Mean = Mean of all the ratings across all characteristics for Arts & Humanities (with NA ratings coded as missing).

AHCI\_ind, AHFS\_ind, AHPD\_ind, AHAS\_ind = 1 if standard mean ≥ 2.0, or = 0 if standard mean < 2.0.

AH Total = sum of AHCI + AHFS + AHPD + AHAS + AHAS

AH Indicator 2 = 1 if AH Total ≥ 3, or = 0 if AH Total < 3

AH Total = AH Curriculum & Instruction Mean + AH Formative & Summative Assessment Mean + AH Professional Development Mean + AH Administrative Support Mean

AH Total \* Ind\_2 = AH Indicator 2 X AH Total

The Performance Categories were determined as such: AH Total < 8.0 = Needs Improvement (NI); AH Total ≥ 8.0 = Proficient (P); AH Total ≥ 10.8 = Distinguished (D).



**Appendix B: Frequency Counts for Items Rated as “Not Applicable”<sup>5</sup> By a Substantive Number of Schools**

**Elementary School Level:**

AH\_3\_12\_8 Administrative/Leadership Support and Monitoring: Time and Resources\_To what extent does the school leadership ensure arts courses are not scheduled in conflict with single-section required content area courses?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	601	89.6	89.6	89.6
1 Needs Improvement	19	2.8	2.8	92.4
2 Proficient	34	5.1	5.1	97.5
3 Distinguished	17	2.5	2.5	100.0
Total	671	100.0	100.0	

AH\_3\_13\_4 Administrative/Leadership Support and Monitoring: Policies and Monitoring\_To what extent does school leadership ensure the use of various data to inform decisions on arts instructional programs, class offerings and staffing?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	610	90.9	90.9	90.9
1 Needs Improvement	43	6.4	6.4	97.3
2 Proficient	12	1.8	1.8	99.1
3 Distinguished	6	.9	.9	100.0
Total	671	100.0	100.0	

PL\_1\_1\_2 Curriculum and Instruction: Student Access\_To what extent does the school provide advising programs that include components of the ILP?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	485	72.3	72.3	72.3
1 Needs Improvement	74	11.0	11.0	83.3
2 Proficient	84	12.5	12.5	95.8
3 Distinguished	28	4.2	4.2	100.0
Total	671	100.0	100.0	

<sup>5</sup> There is evidence that some schools used the “0” rating to indicate “Needs Improvement” as opposed to “Not Applicable.” Consequently, this is a limitation of the 2011-2012 pilot data.

PL\_1\_1\_3 Curriculum and Instruction: Student Access\_To what extent does the school provide PL/CS interdisciplinary and advanced courses?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	648	96.6	96.6	96.6
1 Needs Improvement	14	2.1	2.1	98.7
2 Proficient	7	1.0	1.0	99.7
3 Distinguished	2	.3	.3	100.0
Total	671	100.0	100.0	

PL\_1\_2\_21 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent are the career pathways implemented in the states 14 identified Career Clusters for high school students? (HIGH SCHOOL)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	667	99.4	99.4	99.4
1 Needs Improvement	1	.1	.1	99.6
2 Proficient	3	.4	.4	100.0
Total	671	100.0	100.0	

PL\_1\_2\_23 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent are articulation agreements and dual credit agreements a part of career pathways at the high school level? (HIGH SCHOOL)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	664	99.0	99.0	99.0
1 Needs Improvement	5	.7	.7	99.7
3 Distinguished	2	.3	.3	100.0
Total	671	100.0	100.0	

PL\_1\_2\_25 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent is a formalized plan for introducing and using the ILP in place for the 6th grade and beyond? (MIDDLE AND HIGH SCHOOL)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	539	80.3	80.3	80.3
1 Needs Improvement	15	2.2	2.2	82.6
2 Proficient	84	12.5	12.5	95.1
3 Distinguished	33	4.9	4.9	100.0
Total	671	100.0	100.0	

PL\_1\_2\_26 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent to all high school students select and note in their ILPs at least four course related to their career major and one of the states 14 Career Clusters?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	662	98.7	98.7	98.7
1 Needs Improvement	5	.7	.7	99.4
2 Proficient	3	.4	.4	99.9
3 Distinguished	1	.1	.1	100.0
Total	671	100.0	100.0	

PL\_1\_3\_13 Curriculum and Instruction: Instructional Strategies\_To what extent does the school integrate the ILP in some course to use as a research tool for career choices?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	510	76.0	76.0	76.0
1 Needs Improvement	82	12.2	12.2	88.2
2 Proficient	66	9.8	9.8	98.1
3 Distinguished	13	1.9	1.9	100.0
Total	671	100.0	100.0	

PL\_1\_4\_8 Curriculum and Instruction: Student Performance\_To what extent do teachers and students work together using formative and assessment data to formulate student ILPs?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	492	73.3	73.3	73.3
1 Needs Improvement	96	14.3	14.3	87.6
2 Proficient	74	11.0	11.0	98.7
3 Distinguished	9	1.3	1.3	100.0
Total	671	100.0	100.0	

PL\_1\_4\_9 Curriculum and Instruction: Student Performance\_To what extent are a variety of work based learning opportunities provided across Career Pathways and majors?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	663	98.8	98.8	98.8
1 Needs Improvement	5	.7	.7	99.6
2 Proficient	3	.4	.4	100.0
Total	671	100.0	100.0	

PL\_1\_4\_10 Curriculum and Instruction: Student Performance\_To what extent is the ILP development for all 6th grade students?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	533	79.4	79.4	79.4
1 Needs Improvement	10	1.5	1.5	80.9
2 Proficient	86	12.8	12.8	93.7
3 Distinguished	42	6.3	6.3	100.0
Total	671	100.0	100.0	

PL\_1\_5\_3 Formative and Summative Assessment: Variety of Assessment\_To what extent are formalized plans in place to monitor the completion of the students ILP?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	501	74.7	74.7	74.7
1 Needs Improvement	52	7.7	7.7	82.4
2 Proficient	92	13.7	13.7	96.1
3 Distinguished	26	3.9	3.9	100.0
Total	671	100.0	100.0	

PL\_1\_6\_2 Formative and Summative Assessment: Expectations for Student Learning\_To what extent are middle and high school students required to successfully pass a health education course?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	620	92.4	92.4	92.4
1 Needs Improvement	37	5.5	5.5	97.9
2 Proficient	14	2.1	2.1	100.0
Total	671	100.0	100.0	

PL\_1\_6\_3 Formative and Summative Assessment: Expectations for Student Learning\_To what extent can students demonstrate the knowledge and skills addressed in local, state, and national PL/CS standards?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	6	.9	.9	.9
1 Needs Improvement	204	30.4	30.4	31.3
2 Proficient	424	63.2	63.2	94.5
3 Distinguished	37	5.5	5.5	100.0
Total	671	100.0	100.0	

PL\_1\_6\_6 Formative and Summative Assessment: Expectations for Student Learning\_  
To what extent are middle and high school students required to pass a physical education course?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	616	91.8	91.8	91.8
1 Needs Improvement	37	5.5	5.5	97.3
2 Proficient	17	2.5	2.5	99.9
3 Distinguished	1	.1	.1	100.0
Total	671	100.0	100.0	

PL\_1\_9\_2 Professional Development and Support Services: Participation\_  
To what extent does the school provide training to staff regarding the use of the ILP to help students select course offerings and develop career goals?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	535	79.7	79.7	79.7
1 Needs Improvement	92	13.7	13.7	93.4
2 Proficient	37	5.5	5.5	99.0
3 Distinguished	7	1.0	1.0	100.0
Total	671	100.0	100.0	

PL\_1\_9\_5 Professional Development and Support Services: Participation\_  
To what extent does the school provide professional development opportunities to work with postsecondary partners?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	402	59.9	59.9	59.9
1 Needs Improvement	234	34.9	34.9	94.8
2 Proficient	34	5.1	5.1	99.9
3 Distinguished	1	.1	.1	100.0
Total	671	100.0	100.0	

PL\_1\_12\_5 Administrative/Leadership Support and Monitoring: Time and Resources\_  
To what extent does the high school utilize the Career & Technical Education and Kentucky Schools Facility Planning Manuals for program planning?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	662	98.7	98.7	98.7
1 Needs Improvement	3	.4	.4	99.1
2 Proficient	5	.7	.7	99.9
3 Distinguished	1	.1	.1	100.0
Total	671	100.0	100.0	

PL\_1\_13\_1 Administrative/Leadership Support and Monitoring: Policies and Monitoring\_To what extent does the middle and high school use data from ILP reports to determine PL/CS course offerings?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	620	92.4	92.4	92.4
1 Needs Improvement	41	6.1	6.1	98.5
2 Proficient	8	1.2	1.2	99.7
3 Distinguished	2	.3	.3	100.0
Total	671	100.0	100.0	

PL\_1\_13\_5 Administrative/Leadership Support and Monitoring: Policies and Monitoring\_To what extent does the school have consistency between career information in the ILP and the Individualized Education Plan (IEP) for students with disabilities?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	506	75.4	75.4	75.4
1 Needs Improvement	68	10.1	10.1	85.5
2 Proficient	73	10.9	10.9	96.4
3 Distinguished	24	3.6	3.6	100.0
Total	671	100.0	100.0	

PL\_1\_13\_6 Administrative/Leadership Support and Monitoring: Policies and Monitoring\_To what extent does the high school have measures in place to check for and prevent duplication of coursework between secondary and postsecondary levels?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	666	99.3	99.3	99.3
1 Needs Improvement	2	.3	.3	99.6
2 Proficient	3	.4	.4	100.0
Total	671	100.0	100.0	

**Middle School Level:**

PL\_1\_1\_3 Curriculum and Instruction: Student Access\_To what extent does the school provide PL/CS interdisciplinary and advanced courses?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	104	47.9	48.4	48.4
	1 Needs Improvement	50	23.0	23.3	71.6
	2 Proficient	39	18.0	18.1	89.8
	3 Distinguished	22	10.1	10.2	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_2\_21 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent are the career pathways implemented in the states 14 identified Career Clusters for high school students? (HIGH SCHOOL)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	213	98.2	99.1	99.1
	2 Proficient	2	.9	.9	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_2\_23 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent are articulation agreements and dual credit agreements a part of career pathways at the high school level? (HIGH SCHOOL)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	210	96.8	97.7	97.7
	1 Needs Improvement	2	.9	.9	98.6
	2 Proficient	1	.5	.5	99.1
	3 Distinguished	2	.9	.9	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_2\_24 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent are all students introduced to the 14 career clusters at the elementary level? (ELEMENTARY)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	194	89.4	90.2	90.2
	1 Needs Improvement	5	2.3	2.3	92.6
	2 Proficient	6	2.8	2.8	95.3
	3 Distinguished	10	4.6	4.7	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_2\_26 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent to all high school students select and note in their ILPs at least four course related to their career major and one of the states 14 Career Clusters?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	207	95.4	96.3	96.3
	1 Needs Improvement	1	.5	.5	96.7
	2 Proficient	6	2.8	2.8	99.5
	3 Distinguished	1	.5	.5	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_4\_9 Curriculum and Instruction: Student Performance\_To what extent are a variety of work based learning opportunities provided across Career Pathways and majors?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	210	96.8	97.7	97.7
	1 Needs Improvement	3	1.4	1.4	99.1
	3 Distinguished	2	.9	.9	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_12\_5 Administrative/Leadership Support and Monitoring: Time and Resources\_  
To what extent does the high school utilize the Career & Technical Education and Kentucky Schools Facility Planning Manuals for program planning?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	207	95.4	96.3	96.3
	1 Needs Improvement	3	1.4	1.4	97.7
	2 Proficient	2	.9	.9	98.6
	3 Distinguished	3	1.4	1.4	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

PL\_1\_13\_6 Administrative/Leadership Support and Monitoring: Policies and Monitoring\_  
To what extent does the high school have measures in place to check for and prevent duplication of coursework between secondary and postsecondary levels?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 N/A	211	97.2	98.1	98.1
	1 Needs Improvement	1	.5	.5	98.6
	2 Proficient	2	.9	.9	99.5
	3 Distinguished	1	.5	.5	100.0
	Total	215	99.1	100.0	
Missing	System	2	.9		
Total		217	100.0		

**High School Level:**

PL\_1\_2\_24 Curriculum and Instruction: Aligned and Rigorous Curriculum\_To what extent are all students introduced to the 14 career clusters at the elementary level?  
(ELEMENTARY)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	144	72.4	72.4	72.4
1 Needs Improvement	28	14.1	14.1	86.4
2 Proficient	20	10.1	10.1	96.5
3 Distinguished	7	3.5	3.5	100.0
Total	199	100.0	100.0	

PL\_1\_4\_10 Curriculum and Instruction: Student Performance\_To what extent is the ILP development for all 6th grade students?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 N/A	121	60.8	60.8	60.8
1 Needs Improvement	9	4.5	4.5	65.3
2 Proficient	53	26.6	26.6	92.0
3 Distinguished	16	8.0	8.0	100.0
Total	199	100.0	100.0	