

Wyoming Legislative Service Office Program Evaluation Report



Proficiency Assessments for Wyoming Students (PAWS)

Management Audit Committee
December 15, 2010

As they work to allocate tax dollars effectively and make government more efficient, legislators and administrators need objective information. Program evaluation reports from the Legislative Service Office (LSO) are a source for timely, accurate, and unbiased information on state government performance. These reports assist the Legislature in performing its function of oversight: decisionmakers need to understand the operations of state government in order to make informed decisions on the laws they pass and the financial decisions they make.

Wyoming's legislative evaluation activities began in 1971 with the creation of LSO and establishment of the legislative auditor to examine agencies' accounts and operations. In subsequent years, the Legislature changed the section's orientation from financial to sunset auditing, and then to program evaluation. These reviews compare what a program is accomplishing to what the Legislature intended the program to accomplish. W.S. 28-8-107 through 113 authorizes the program evaluation function.

The Management Audit Committee chooses state government programs for review and approves the final reports for release. An eleven-member bi-partisan committee, it has representation from the Senate and the House of Representatives.

LSO program evaluators research, analyze, and write reports on the assigned topics. The reports assess effectiveness and efficiency, examine whether intended results are being achieved, and include non-binding recommendations for change in administrative policies as well as for statutory changes.

Management Audit Committee

Representative David Miller, Chairman

Senator John Hines, Vice Chairman

Senator Bruce Burns

Senator Floyd Esquibel

Senator Tony Ross

Senator John Schiffer

Senator Kathryn Sessions

Representative James Byrd

Representative Michael Madden

Representative John Patton

Representative Bill Thompson

Program Evaluation Staff

Gerald W. Hoppmann, Program Evaluation Manager

Joy Hill, Program Evaluator

William Freeman, Associate Program Evaluator

Anthony Sara, Technical Assistance and Graphics



Wyoming Legislative Service Office

EXECUTIVE SUMMARY

Proficiency Assessments for Wyoming Students (PAWS)

Program Evaluation Section

December 23, 2010

Purpose

On February 4, 2010, the Management Audit Committee directed staff to conduct an audit of the Proficiency Assessments for Wyoming Students (PAWS). PAWS is used to assess students' proficiency in reading, math, writing, and science (as of Spring 2008). The overall objective of this audit is to examine and comment on the costs associated with PAWS, and to explore other alternatives for the future.

Wyoming has 48 independent school districts, which are required by W.S. 21-3-110 (a) (xxiv) to develop student assessment systems to measure student performance against student content and performance standards, relative to the common core of knowledge and skills prescribed under W.S. 21-9-101 (b). Their systems should be integrated with the statewide assessment system (e.g. PAWS).

The Committee's request was based on concerns that PAWS was too expensive, took too long to administer, and was not administered in a manner that provides timely scores to school districts. There were also concerns about the subjectivity of scoring certain portions of PAWS, such as the writing component. It also questioned whether different assessments already being used in Wyoming's school districts could be used in place of PAWS.

Background

The Elementary and Secondary Education Act or ESEA (Pub. L. 89-10, 79 Stat. 27, 20 U.S.C. Ch. 70) was originally enacted on April 11, 1965. The purpose of the Act is to fund primary and secondary education in the United States. The Act is comprised of seven titles setting forth the major provisions that include the following: professional development; instructional materials; resources to support educational programs; and parental involvement.

The Act was originally authorized from 1965 through 1970 and has been reauthorized every five years thereafter. The current reauthorization is the No Child Left Behind Act of 2001 (NCLB), which requires children to become proficient in reading and math by 2014. NCLB places focus on setting high standards and creating measurable goals that will improve education. It also requires states to create assessments to be administered in certain grades that measure against state content and achievement standards.

In Wyoming, student proficiency is assessed against Wyoming's "*common core of knowledge*" as annotated in W.S. 21-9-101 (b). They focus primarily on reading, writing, and mathematics for grades one (1) through eight (8). It should be noted, that pursuant to W.S. 21-2-304 (a) (iii) "*student content and performance standards*" are also required in order to measure against the

“common core of knowledge” and *“common core of skills”* annotated in W.S. 21-9-101 (b) and WDE Chapter 31 Rules (Graduation Requirements), Section 8.

Wyoming statutes, as discussed above, are generally consistent with the No Child Left Behind (NCLB) Act of 2002. More specifically, 20 USC 6311, Section 1111 (b) (3) requires a single statewide assessment for students in grades 3 through 8 in mathematics, reading, language arts and science (no later than SY 2007-2008). NCLB also requires that assessments measure against challenging state academic content and achievement standards (Section 111 (b) (2)).

The Committee also inquired about the impact of pending reauthorization of ESEA, originally passed in 1965. The most recent reauthorization of the Act is No Child Left Behind (NCLB), passed in 2001. Pending reauthorization will more than likely change federal requirements related to Title I (Improving the Academic Achievement of Disadvantaged Students) and other federal funds.

Results in Brief

The Proficiency Assessment for Wyoming Students (PAWS) is not being fully administered as an integrated statewide assessment. The original intent was to measure twice per year what students know, by using multiple summative statewide assessments that measure proficiency against student content and performance standards. It is also used for federal accountability purposes.

Although envisioned as such, PAWS is not fully utilized as an instructionally supportive assessment, nor does it include a formative assessment component. Wyoming chose to go beyond the requirements of NCLB by

envisioning an integrated statewide assessment.

WDE does not specifically know the extent by which schools and districts use various instructionally supportive reports and suggestions based on PAWS scores, nor are formative assessments used as part of PAWS, to supplement the summative assessment administered each year.

As a result, districts may not be using PAWS to support instructional decisions, which call into question the \$78,561,068 Wyoming will spend on PAWS through BFY 2011 for an integrated statewide assessment. Other similar states we surveyed appear to spend less on their statewide assessments.

For example, North Dakota stated it expends approximately \$3 million annually for administration of its assessment. South Dakota stated it expends approximately \$850,000 per year for development and \$3.8 million annually for administration. Finally, Vermont expends \$2 million per year for development and \$2.25 million annually for administration. It should also be noted that South Dakota also uses NCS Pearson, Inc. (the company that administers PAWS in Wyoming).

General consensus from the districts surveyed express concern with PAWS in several areas. The test is perceived to be too long. Also, scores are not given to educators in a timely manner.

Many educators indicate that they want to use test results to help place students, but because results are not received until three months into the semester PAWS has limited use in guiding classroom instruction.

There are also concerns with the ability of the vendor to properly administer the test because of problems experienced in previous

years related to technology and scoring. Educators also question the validity and reliability of PAWS, but often associate that assumption with different criticism of PAWS (e.g. test administration, lack of a growth model within ESEA, etc.).

According to the *Standards for Educational and Psychological Testing*, validity refers to the degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test. Reliability is defined as the degree to which test scores are consistent over repeated applications of a measurement procedure. Through the federal Department of Education's peer review process, PAWS is deemed to be valid and reliable for purposes of NCLB.

There is no doubt that NCS Pearson, Inc., encountered problems and performed poorly with respect to the administration of PAWS during the 2009-2010 school year. An apology from NCS Pearson, Inc., to Wyoming residents in July 2010, as part of ongoing settlement negotiations, was disturbingly similar to earlier apologies to Florida and Minnesota for similar problems. In Wyoming's apology, the company stated it was "*dismayed to have let you down,*" but promised to "*take additional steps to regain your confidence.*"

As a result of the problems experienced during its last PAWS administration, many policy makers and state officials have lost complete confidence in the vendor. The result on students' achievement is not as clear cut. It should be noted however, that WDE is still in settlement negotiations with NCS Pearson, Inc.

According to WDE officials, the federal peer review process has been laborious, has impacted the instructionally supportive component of PAWS, and has been

inconsistently applied to all states. Education experts in the field concur, and one fully blames the bureaucracy within the federal Department of Education (Department) for the inability of PAWS to be used to its fullest potential.

As a result, Wyoming's inability to fully utilize its integrated statewide assessment as discussed in Chapter 2 may be partially a result of stressed resources at the WDE level. According to the Department of Education, 31 states have received full peer review approval.

According to sources in the press, as well as various state and federal officials, ESEA will more than likely be reauthorized by the end of 2011. Generally, there is a consensus that an overemphasis exists on standardized testing for accountability purposes, which may drastically change the current law.

Wyoming, like every other state in the nation, is experiencing uncertainty with respect to its statewide assessments. States are pondering changes beyond ESEA's most recent reauthorization, NCLB. As a result, states are in a holding pattern with respect to their statewide assessments, even though they anticipate substantive changes in the next reauthorization.

States use a variety of statewide assessments to meet NCLB requirements. As long as states successfully illustrate to the federal government that statewide assessments align with their content and performance standards, there is flexibility as to the development of their assessments. The primary requirements however, are that statewide assessments should be summative in nature, as well as testing the appropriate grades and topic areas.

Principal Findings

There are six significant finding areas and related recommendations discussed in more detail in this report:

1. PAWS as an integrated statewide assessment not fully utilized (Chapter 2);
2. Local districts criticisms of PAWS (Chapter 3);
3. Vendor performance during the SY 2009-2010 PAWS administration (Chapter 4);
4. Federal Department of Education's peer review process frustrates WDE and other states (Chapter 5);
5. Elementary and Secondary School Act (ESEA) reauthorization will impact all states (Chapter 6); and
6. Similar requirements are applied to all states related to NCLB (Chapter 7).

Agency Comments

WDE agrees fully with all of the report's recommendations. In addition, Chapter 4, Recommendation #2, is "*already in place.*"

WDE also stated it "*is committed to the importance of assessment and its appropriate place in the standards, instruction and assessment cycle. The department will continue to provide professional development in the areas of understanding how to use PAWS assessment results in conjunction with other district and classroom assessment results for school program and instructional improvement, as well as an understanding of the related current and future federal and state assessment requirements.*"

Copies of the full report are available from the Wyoming Legislative Service Office. If you would like to receive the full report, please fill out the enclosed response card or phone 307-777-7881. The report is also available on the Wyoming Legislature's website at <http://legisweb.state.wy.us>

Recommendation Locator

Chapter Number	Recommendation Number	Recommendation Summary	Page Number	Party Addressed	Agency Response
2	1	WDE should take action to determine how often districts use the “traffic-signal” reports and its instructional suggestions offered as part of an instructionally supportive statewide assessment. In addition, WDE should consider developing a reporting mechanism for districts to use in order to report this information, as well as identify ways that districts actually use the reports.	41	WDE	Agree
2	2	WDE should continue to work with districts to tie in existing use of formative assessments, more specifically NWEA MAP, with PAWS. These assessments throughout the year can provide important information to districts and schools with respect to instructional changes.	42	WDE	Agree
2	3	At a minimum, WDE should track the cost-effectiveness of PAWS by comparing the dollars expended for professional services against the use of instructional components of the integrated statewide assessment.	42	WDE	Agree
2	4	Pending reauthorization of ESEA, WDE should work with the Wyoming Board of Education to determine if the statewide assessment and related reports should be used for other high-stakes purposes such as graduation requirements, college scholarship criteria, or pay-for-performance for teachers and administrators. Based on the conclusions, criteria should then be factored into the next RFP.	54	WDE	Agree

Chapter Number	Recommendation Number	Recommendation Summary	Page Number	Party Addressed	Agency Response
3	5	The WDE should develop and administer training on NCLB requirements, as well as pending ESEA reauthorization with respect to differences between various assessments as they apply to NCLB requirements. This should be applied to additional training on the PAWS program from RFP to contract.	66	WDE	Agree
4	6	WDE should continue seeking the maximum amount of liquated and default damages based on problems encountered by NCS Pearson, Inc. as it carried out contractual obligations.	77	WDE	Agree
4	7	WDE should continue working with the Department with respect to its AYP waiver. It should also ensure that the recommendations from NCIEA's report are fully implemented.	78	WDE	Already in Place
4	8	WDE should continue to address concerns with respect to the writing portion of PAWS. It should also determine whether eliminating the writing portion in the future, would diminish the importance of writing within the districts. It should work with the local school districts to obtain their feedback as well on this option, as well as modifications to remove the "drafting" portion of the component.	80	WDE	Agree
5	9	The Management Audit Committee may wish to consider requesting that the U.S. Senate Committee on Health, Education, Labor, and Pensions request that the Government Accountability Office (GAO) conduct an audit of the federal peer review process pending reauthorization of ESEA.	90	Legislature	N/A
6	10	WDE should work with the current vendor to correct discrepancies with respect to PAWS administration during Spring 2010. In addition, it should not make drastic changes to the	103	WDE	Agree

Chapter Number	Recommendation Number	Recommendation Summary	Page Number	Party Addressed	Agency Response
		current assessment, until ESEA has been reauthorized.			
6	11	WDE should study the activities at the federal, state, and local levels with respect to the direction of ESEA reauthorization. Information related to the blueprint, RTT funding, as well as the development and funding of various consortium can be used to position itself for eventual ESEA reauthorization.	103	WDE	Agree



TABLE OF CONTENTS

Proficiency Assessments for Wyoming Students (PAWS)

Introduction: Objective and Scope, Methodology	1
Chapter 1: Background	9
Chapter 2: PAWS is not being used fully as an Integrated Statewide Assessment	21
Chapter 3: Districts are critical of PAWS, but also appear to speak from frustration.	55
Chapter 4: School Year 2009-2010 PAWS administration failed the citizens of Wyoming..	67
Chapter 5: U.S. Department of Education peer review process appears inconsistent across all states	81
Chapter 6: Elementary and Secondary Education Act (ESEA) impact on PAWS.....	93
Chapter 7: Other States	105
Chapter 8: Conclusion	117
Agency Response	119
<i>Appendices</i>	
(A) <i>District Superintendent Survey Comments</i>	A-1
(B) <i>Selected Statutes</i>	B-1
(C) <i>Alleged Liquidated and Default Damages</i>	C-1
(D) <i>Health, Education, Labor, and Pensions Committee Letter</i>	D-1



List of Acronyms and Definitions

Proficiency Assessments for Wyoming Students (PAWS)

AIMSweb: Assessment and data management for RTI. It is a benchmark and progress monitoring system based on direct, frequent and continuous student assessment.

ARRA: American and Recovery Reinvestment Act (ARRA) of 2009, which includes funding for Race to the Top (RTT) grant funds.

AYP: Adequate Yearly Progress, as mandated by the No Child Left Behind Act, 2001.

BOE: District Body of Evidence, which has been studied by the Wyoming Board of Education, with recommendations for change to W.S. 21-2-304(a) (iv).

CCSSI: Common Core State Standards Initiative.

CCSSO: Council of Chief State School Officers.

Commission: Commission on Instructionally Supportive Assessment, which released a report in 2001, used by the Wyoming Statewide Task Force on Student Assessment and Educational Accountability (Task Force) to make recommendations in 2003.

Common Core of Knowledge and Skills: Wyoming standards pursuant to W.S. 21-9-101 (b), which focus on reading, writing, and mathematics for 1st through 8th grades.

CRT: Criterion Referenced Test.

Department: U.S. Department of Education.

DIBELS: Dynamic Indicators of Basic Early Literacy Skills.

ELL: English Language Learner.

Embedded Formative Assessment Component: A component of a summative assessment designed to support instructional improvement through embedded or computerized assessments.

ESEA: Elementary and Secondary Education Act (1965).

- Title I-Improving the Academic Achievement of the Disadvantaged;
- Title II-Preparing, Training, and Recruiting High Quality Teachers and Principals;
- Title III-Language Instruction for Limited English Proficient and Immigrant Students;

- Title IV-21st Century Schools;
- Title V-Promoting Informed Parental Choice and Innovative Programs;
- Title VI-Flexibility and Accountability;
- Title VII-Indian, Native Hawaiian, and Alaska Native Education; and
- Title VIII-Impact Aid Program.

Formative Assessment: An assessment that provides feedback to teachers for the purpose of improving instruction, which is used for student improvement rather than for making final decisions or for accountability. Such assessments can also be called intermediate assessments.

Harcourt Education International: The first vendor to develop and administer PAWS.

IASA: Improving American Schools Act (1994).

IBARS: Internet Budgeting Analysis and Reporting System.

Integrated Statewide Assessment: Includes an instructionally supportive component and an embedded formative assessment component.

Instructionally Supportive Component: A component of a summative assessment designed to measure a student's knowledge of content and performance standards, as well as assisting to improve teaching and learning.

NAEP: National Assessment of Academic Progress-the only nationally representative assessment of how students test against various subject areas. States receiving Title I funds are required to participate in the program.

NCS Pearson, Inc.: The second vendor to develop and administer PAWS. This vendor in May, 2007, purchased Harcourt Education International.

NGA Center: National Governors Association Center for Best Practices.

NRT: Norm Referenced Test.

NWEA MAP: Northwest Evaluation Association-Measures of Academic Progress.

PAWS: Proficiency Assessment for Wyoming Students.

PAWS-ALT: Proficiency Assessment for Wyoming Students-Alternative.

Partnership for Assessment of Readiness for College and Careers (Partnership): A collection of 26 states that have competed successfully for RTT grants to develop common assessments aligned to the CCSSI.

RTI: Response to Intervention.

Smarter Balanced Consortium (SBAC): A collection of 31 states that have competed successfully for RTT grants to develop common assessments aligned to CCSS.

Student Content and Performance Standards: Wyoming standards pursuant to W.S. 21-2-304 (a) (iii) are used to measure against the Common core of Knowledge and Skills.

Summative Assessment: An assessment intended to represent systematically described content and performance standards.

TAC: Technical Advisory Committee required by NCLB to assist states with psychometric concerns, etc., related to federal requirements.

Task Force: The Wyoming Statewide Task Force on Student Assessment and Educational Accountability (Task Force), created under 2003 Wyoming Session Laws, Chapter 208, Section 201.

Traffic Signal Process: WDE instructionally supportive process for reporting based on the skill-reporting category level by student and district. There are 22 cognitive skills that are measured as part of PAWS to determine if students are proficient against Wyoming's content and performance standards. Results are illustrated either by a red light, yellow light, or green light. The process also includes instructional suggestions for districts and teachers, based on various reports.

WDE: Wyoming Department of Education.

WOLFS: Wyoming Online Financial System.

WyCAS: Wyoming Comprehensive Assessment System. WyCAS is most notable for its "matrix design", where multiple forms are used to measure against content and grade.

2004 RFP: RFP # 0343-L originally requesting bids for an integrated statewide assessment with an instructionally supportive component as well as an embedded formative assessment component.

2008 RFP: RFP # 0056-S modified requesting bids for an integrated statewide assessment with an instructionally supportive component, but not the embedded formative assessment.



INTRODUCTION

Objective, Scope, and Methodology

Introduction

The Management Audit Committee requested an audit of the Proficiency Assessments for Wyoming Students (PAWS). PAWS is used to assess students' proficiency in reading, math, writing, and science (as of Spring 2008).

W.S. 21-2-304 (a) (v) requires the State Board of Education, through the Superintendent and in consultation and coordination of local school districts, to implement a statewide assessment to measure individual student performance in grades three (3) through eight (8) and grade eleven (11).

In Wyoming, student proficiency is assessed against Wyoming's "*common core of knowledge*" as annotated in W.S. 21-9-101 (b). They focus primarily on reading, writing, and mathematics for grades one (1) through eight (8). It should be noted, that pursuant to W.S. 21-2-304 (a) (iii) "*student content and performance standards*" are also required in order to measure against the "*common core of knowledge*" and "*common core of skills*" annotated in W.S. 21-9-101 (b) and WDE Chapter 31 Rules (Graduation Requirements), Section 8.

Wyoming has 48 independent school districts, which are required by W.S. 21-3-110 (a) (xxiv) to develop student assessment systems to measure student performance against student content and performance standards, relative to the common core of knowledge and skills prescribed under W.S. 21-9-101 (b). Their systems should be integrated with the statewide assessment system (e.g. PAWS).

Wyoming statutes, as discussed above, are generally consistent with the No Child Left Behind (NCLB) Act of 2002. More specifically, 20 USC 6311, Section 1111 (b) (3) requires a single statewide assessment for students in grades 3 through 8 in

mathematics, reading, language arts and science (no later than SY 2007-2008). NCLB also requires that assessments measure against challenging state academic content and achievement standards (Section 111 (b) (2)).

Legislative Concern(s) The Committee's request was based on concerns that PAWS was too expensive, took too long to administer, and was not administered in a manner that provides timely scores to school districts. There were also concerns about the subjectivity of scoring certain portions of PAWS, such as the writing component. It also questioned whether different assessments already being used in Wyoming's school districts could be used in place of PAWS.

The Committee inquired about the impact of pending reauthorization of the Elementary and Secondary Education Act (ESEA), originally passed in 1965. The most recent reauthorization of the Act is No Child Left Behind (NCLB), passed in 2001. Pending reauthorization will more than likely change federal requirements related to Title I (Improving the Academic Achievement of Disadvantaged Students) and other federal funds.

Objective

W.S. 28-8-107(b) authorizes the Legislative Service Office to conduct program evaluations, performance audits, and analyses of policy alternatives. Generally, the purpose of such research is to provide a base of knowledge from which policymakers can make informed decisions.

On February 4, 2010, the Management Audit Committee directed staff to conduct an audit of the Proficiency Assessment of Wyoming Students (PAWS). The overall objective of this audit is to examine and comment on the costs associated with PAWS, and to explore other alternatives for the future. More specifically, the report addresses the following questions:

1. Given that PAWS is an outcome measure used primarily for measuring Adequate Yearly Progress (AYP), has Wyoming expended \$63 million (includes BYF 11 Budget Request figures) wisely?
 - Can more be accomplished with respect to aligning PAWS with the ongoing screening and diagnostic assessments occurring through the State on a continual basis?
 - Should PAWS test scores be tied somehow to other education requirements such as graduation, ACT testing, Hathaway scholarships, or even performance-based pay for teachers?
 - Can expenditure analysis indicate areas in which funds were expended, that may have been expended more wisely in other areas? Such an analysis could be beneficial with respect to the pending reauthorization of NCLB.
2. Is the whole notion of AYP an internal bias for schools that administer PAWs to “*teach to the test*”, as opposed to concentrating more on other screening and diagnostic assessments?
 - What is the sentiment of school districts regarding NCLB requirements, AYP, and other assessments? What types of ideas might they have with respect to tying in accountability measures related to PAWS, with ongoing efforts to help students become proficient?
3. Are there issues with respect to the United States Department of Education’s peer review of PAWS-ALT and the Science portion of PAWS re: compliance with the Elementary and Secondary Education Act, 1965 that need additional review?
4. What are the major changes that will occur with PAWS administration with respect to the contract with the newly selected vendor? In other words, how has PAWS, as well as its administration changed since 2006?
 - What can Wyoming’s parents and students expect in the future related to the use of PAWS?

5. Explore other assessments that could possibly be used in the future such as AIMSWeb, National Assessment of Educational Progress (NAEP), Northwest Evaluation Association Measures of Academic Progress (MAP), Dynamic Indicators of Basic Early Literacy Skills (DIBELS), and the 11th grade ACT.
6. Research possible changes to PAWS based on Elementary and Secondary Education Act (ESEA) reauthorization towards a Growth Model.
7. Examine the flexibility of the current PAWS contract to modify PAWS if needed, as a result of federal changes.
8. Examine proposals with respect to the current writing portion of PAWS.

Scope Explanation It should be noted that states are also required by NCLB (Title IX, Part C, Section 9302) to complete and submit to the federal Department of Education a *Consolidated State Application Accountability Workbook*. Currently, all states have approved accountability plans (as of June 10, 2003), which are different than statewide assessments required by NCLB (Title I, Part A, Section 1111) to measure Adequate Yearly Progress (AYP).

The deadline for State submissions was January 31, 2003. The *Workbook* is different than the development of the statewide assessment to meet AYP, essentially covering broad areas of a state's education accountability system. Generally, the accountability system is required to address the following: 1) all schools; 2) all students; 3) method of AYP determinations; 4) annual decisions; 5) subgroup accountability; 6) based on academic assessments; 7) additional indicators; 8) separate decisions for reading/language arts and mathematics; 9) system validity and reliability; and 10) participation rate.

Our audit does not discuss the process by which state accountability plans are developed, submitted, and approved. Rather, we focus primarily on the development, submission, and approval of Wyoming's statewide assessment (PAWS). According to the federal Department of Education, only 31 states

have received full peer review approval for their statewide assessments.

Body of Evidence (BOE) as an area for further study by the Management Audit Committee

During our audit, we heard concern from districts with respect to the Body of Evidence (BOE) peer review process conducted by WDE. We also understand that a study group has made a recommendation to the Wyoming Board of Education (WBOE) has made a recommendation to discontinue the BOE in lieu of a more standard graduation exit examination. More specifically, *“to amend Wyoming Statute 21-2-304(a) (iv) which requires a district BOE system to determine graduation to include a state-developed statewide graduation assessment system to measure student performance in regard to the state content and performance standards.”*

W.S. 21-2-304 (a) (iv) and WDE Chapter 31 (Graduation Requirements) rules require districts to submit graduation requirements for peer review approval. More specifically, to require the State Board of Education to work with districts to *“establish requirements for students to earn a high school diploma as measured by each district’s body of evidence assessment system prescribed by rule and regulation of the state board and required under W.S. 21-3-110 (a) (xxiv).”* It should also be noted that Boards of Trustees are required to include a body of evidence as a component of its district assessment system.

WDE Chapter 31 Rules-Graduation Requirements, Section 10 (d) requires a districts’ Body of Evidence to undergo a peer review process. The peer review group is comprised of Wyoming educators *“who have successfully completed peer review training conducted by WDE.”* This group makes recommendations to the Superintendent of Public Instruction.

The rules also state that evidence shall include the following components: 1) district assessment plans; 2) evidence of alignment among standards, curriculum, and assessments; 3) sample assessments; 4) evidence of consistency, documentation of the standard setting methods; 5) evidence supporting the fairness of the assessment system; 6) documentation supporting the comparability of the assessment system across schools and

years; and 7) other documentation that the district chooses to submit to support the technical quality of the body of evidence assessment system.

If the Management Audit Committee so desired, it could request the Legislative Service Office (LSO) to conduct a *Scoping Paper* to determine BOE is a suitable audit topic.

Scope and Methodology

This evaluation was conducted according to statutory requirements and professional standards and methods for governmental audits. The research was performed from June 2010 through November 2010. The general time frame for which we included information for this report is BFY 2005 through BFY 2011 (unless otherwise noted).

Our research methods included:

Interviews

1. Interviewed Wyoming Department of Education (WDE) officials and personnel responsible for programmatic and financial aspects of PAWS.
2. Interviewed past WDE officials, Technical Advisory Committee (TAC) members, and other experts in the field of psychometrics and statewide assessments.
3. Interviewed officials from the U.S. Department of Education.
4. Interviewed staff from the U.S. Senate Committee on Health, Education, Labor, and Pensions.

Survey Instrument

5. Developed a *Confidential District Superintendent Survey* that was sent to 48 school districts, 43 of which completed and returned the survey; data was fully analyzed.

Research and Analysis

6. Reviewed existing Legislative Service Office (Education Finance Section) publications related to work conducted by the School Finance Recalibration Committee.
7. Reviewed Wyoming Statutes Title 21 and related rules and

- the No Child Left Behind Act (2001) and related Code of Federal Regulations (CFR).
8. Reviewed professional publications and literature from Center for Education Assessment, National Assessment of Educational Progress, National Conference of State Legislatures, Council of Chief State School Officers, Wyoming Department of Education, U.S. Department of Education, and Northwest Evaluation Association.
 9. Reviewed information on various consortia: New England Compact; Smarter Balanced Assessment Consortium; Partnership for the Assessments of College and Career Readiness; 14-State Collaboration (e.g. end-of-course exam in Algebra 2); World Class Instructional Design and Assessment; National Center on Education and the Economy; State Collaborative on Assessment and Student Standards on Career Technical Education; and the State Collaborative on Assessment and Student Standards.
 10. Conducted Internet research on other states' practices; we also interviewed and surveyed other states officials of selected states for clarification.
 11. Reviewed the Obama Administration's Blueprint for Success.
 12. Reviewed various publications related to "growth modeling" as an indicator of student progress towards proficiency.
 13. Reviewed "growth modeling" as it relates to approved U.S. Department of Education programs as part of NCLB.
 14. Reviewed WDE staff qualifications related to psychometric backgrounds or education.
 15. Reviewed Wyoming Title 21 with respect to various reporting and auditing authority for the Superintendent and WDE.
 16. Reviewed AYP data for Wyoming school districts.
 17. Reviewed Data Advisory Committee minutes for discussion of the Program; W.S. 21-2-203 (d) creates the Committee.
 18. Reviewed and analyzed PAWS and WyCAS item types for various school years.
 19. Reviewed PAWS and NAEP scores for general analysis as compared to other states.
 20. Reviewed and analyzed various audits from the Government Accounting Office, U.S. Department of Education, and other states.

21. Reviewed and analyzed American Recovery and Reinvestment Act (ARRA) funds related to assessment activities, including the Race to the Top Program.
22. Reviewed and analyzed RFP No. 0343-L (2004) and RFP No. 0056-S (2008) and vendor responses.
23. Reviewed and analyzed Proficiency Assessment for Wyoming Students (PAWS) data for SY 2008 (PAWS administered in Spring 2009).
24. Reviewed and analyzed various WDE surveys to districts with respect to PAWS administration, paper and pencil options, etc.
25. Reviewed and analyzed district Body of Evidence (BOE) folders for indications that PAWS was used for criteria related to graduation requirements.
26. Reviewed and analyzed WDE “traffic-signal” reports and instructional suggestions to serve as the instructionally supportive component of PAWS.
27. Reviewed and analyzed various formative assessments to determine if they could be used in place of PAWS: AIMSWeb; National Assessment of Educational Progress (NAEP); Northwest Evaluation Association Measures of Academic Progress (MAP); Dynamic Indicators of Basic Early Literacy Skills (DIBELS); and the 11th grade ACT.
28. Conducted reconciliation(s) with respect to IBARS appropriations and WOLFS expenditures and encumbrances.
29. Reviewed and analyzed U.S. Department of Education peer review notes and responses by WDE.

Acknowledgements

The Legislative Service Office expresses appreciation to the Wyoming Department of Education for their continued cooperation throughout this project.

We would also like to thank the school district superintendents who completed and returned our *Confidential District Superintendent Survey* as well as officials from other states who assisted our research by providing additional interviews and documentation.

CHAPTER 1

Background

Wyoming's response to the No Child Left Behind Act of 2001 (NCLB)

*Elementary and
Secondary Education
Act (1965)*

The Elementary and Secondary Education Act (Pub. L. 89-10, 79 Stat. 27, 20 U.S.C. Ch. 70) was originally enacted on April 11, 1965. The purpose of the Act is to fund primary and secondary education in the United States. The Act is comprised of seven titles setting forth the major provisions that include the following: professional development; instructional materials; resources to support educational programs; and parental involvement.

The Act was originally authorized from 1965 through 1970 and has been reauthorized every five years until 2001. The most recent reauthorization is the No Child Left Behind Act of 2001 (NCLB), which requires children to become proficient in reading and math by 2014. NCLB places focus on setting high standards and creating measurable goals that will improve education. It also requires states to create assessments to be administered in certain grades that measure against state content and achievement standards.

According to sources in the press, as well as various state and federal officials, ESEA will more than likely be reauthorized by the end of 2011. Generally, there is a consensus that an overemphasis exists on standardized testing for accountability purposes, which may drastically change the current law.

From what we understand from various education experts, it is highly likely that the core of ESEA will remain intact re: requirements for content standards, academic achievement standards, as well as grade-level summative assessments in reading and mathematics that disaggregate student data. There will more than likely be some type of measurement for Adequately Yearly Progress (AYP), although it is unclear whether it will focus strictly on standardized test scores to

illustrate proficiency, as opposed to broad growth indicators that show gradual proficiency towards college or career readiness. Whatever form AYP takes however, it will probably be based on some type of requirement for a summative assessment(s) measuring against shared common or state standards.

For more specific details of the ESEA reauthorizations, including NCLB, see Chapter 6 of this report.

*State law mirrors
NCLB, but goes
beyond*

W.S. 21-2-304 (a) (v) requires the State Board of Education (Board), through the Superintendent and in consultation and coordination of local school districts, to implement a statewide assessment to measure individual student performance in grades three (3) through eight (8) and grade eleven (11). PAWS is used to assess students' proficiency in reading, math, writing, and science (as of Spring 2008).

Wyoming statutes, as discussed above, are generally consistent with NCLB. More specifically, 20 USC 6311, Section 1111 (b) (3) requires a single statewide assessment for students in grades 3 through 8 in mathematics, reading, language arts and science (no later than SY 2007-2008). NCLB also requires that assessments measure against challenging state academic content and achievement standards (Section 111 (b) (2)).

In Wyoming, student proficiency is assessed against Wyoming's "*common core of knowledge*" as annotated in W.S. 21-9-101 (b). They focus primarily on reading, writing, and mathematics for grades one (1) through eight (8). It should be noted, that pursuant to W.S. 21-2-304 (a) (iii) "*student content and performance standards*" are also required in order to measure against the "*common core of knowledge*" and "*common core of skills*" annotated in W.S. 21-9-101 (b) and WDE Chapter 31 Rules (Graduation Requirements), Section 8.

The standards as discussed above were adopted by the Wyoming State Board of Education on November 19, 2008. They are listed on WDE's webpage at the following address: <http://www.k12.wy.us/SA/standards.asp>.

In accordance with W.S. 21-2-304 (a) (iii) the Board in consultation and coordination with local school districts will

“prescribe uniform student content and performance standards for the common core of knowledge and the common core of skills specified under W.S. 21-9-101(b). . . .” Under W.S. 21-2-304(c) the Board is directed to evaluate and review the uniformity and quality of the educational program standards at least every five (5) years. This was last done in 2003 and again in 2008.

According to WDE’s website, it is *“currently developing a process, beginning summer 2008, to conduct a thorough review and revision of the Standards in each content area, to be completed by 2013. The process will provide for broad involvement of Wyoming citizens, including those from school districts, Wyoming Community Colleges, the University of Wyoming, parents, and the wider Wyoming community.”*

It should also be noted, that on June 16, 2010 the Board voted to adopt the common core standards drafted by the Common Core State Standards Initiative (CCSSI). More specifically, the Board approved the following motion:

“The Board approves the adoption of 100% of the common core state standards in English Language Arts and Mathematics in the next revision of the Wyoming Content Performance Standards to be completed by December, 2011. The common core state standards will comprise at least 85% of the revised Wyoming Content Performance Standards in each of these two content areas.”

See Chapter 6 for additional information on Wyoming’s action with respect to the common core standards and other areas concerning ESEA reauthorization.

Wyoming Dismantles its Current Statewide Assessment

Wyoming Legislature creates the Task Force on Student Assessment and Education

In 2003, Wyoming began the process of moving from the Wyoming Comprehensive Assessment System (WyCAS), the statewide assessment that satisfied the requirements of the 1994 Improving America’s School Act (IASA). The requirements of the subsequent 2001 authorization of ESEA, NCLB, were much

Accountability more stringent than the previous authorization.

The Wyoming Legislature created the Task Force on Student Assessment and Education Accountability (Task Force) to issue findings and make recommendations to the Joint Education Interim Committee (JEIC). It was charged with two major tasks: **1)** recommend modifications, if necessary, to Wyoming's statewide assessment system to improve teaching and learning and foster school improvement; and **2)** recommend an accountability system with consequences that will assist in meeting NCLB's accountability requirements while maintaining uniformity and quality of state standards.

On August 28, 2003, the Task Force presented its findings and recommendations to the JEIC, which provided feedback to the Task Force. That feedback served as the foundation for the preparation of its final report. More specifically, the findings and recommendations were intended to be the basis for recommendations made to the Legislature during the 2004 budget session.

In October 2003, the Task Force released its final report and recommendations, which became the basis for an initial Request for Proposal (RFP), which was opened June 2, 2004. The recommendations from the Task Force also became the basis for changes to Wyoming Statutes, Title 21, Chapter 2 (The Administration of the State System of Education at the State Level), as discussed above.

It should be noted that its recommendations were primarily based on a report issued in 2001 by the Commission on Instructionally Supportive Assessment (Commission). The Commission, chaired by Dr. James Popham (also chair of Wyoming's Technical Advisory Committee), issued *Building Tests to Support Instruction and Accountability: A Guide for Policymakers*, which offers advice to state policy makers to assist them in the development of tests that are instructionally supportive.

The advice is contained within nine (9) requirements, which according to the Commission's report, "*must be met to ensure a responsible educational assessment system for the improvement*

of learning.”

The following table provides additional information on the Commission’s earlier recommendations.

Table 1.1
Commission Recommendations

Requirements	Detail
Requirement 1	<p>A state’s content standards must be prioritized to support effective instruction and assessment.</p> <ul style="list-style-type: none"> • Small number of content standards used for large-scale assessment and represent the most important skills students need to learn. • Several test questions must be asked to assess student performance on those standards, which is one reason standards must be prioritized. Too many standards to be assessed will make tests too long.
Requirement 2	<p>High-priority content standards must be clearly and thoroughly described so knowledge and skills students need to demonstrate competence are evident.</p> <ul style="list-style-type: none"> • If educators clearly understand what to focus on and test-developers, what to test, then better teaching and testing will result.
Requirement 3	<p>Results of a state’s assessment of high-priority content standards should be reported standard-by-standard for each student, school, and district.</p> <ul style="list-style-type: none"> • Stakeholders need information about which content standards students are or are not achieving so improvement can be addressed.
Requirement 4	<p>Must provide educators with optional classroom assessment procedures that can measure students’ progress on content standards not assessed by state tests.</p>

Requirements	Detail
Requirement 5	Must monitor the breadth of the curriculum to ensure that instructional attention is given to all content standards and subject areas, including those that are not assessed by state tests.
Requirement 6	<p>Must ensure that all students have the opportunity to demonstrate their achievement of state standards; consequently, it must provide well-designed assessments appropriate for a broad range of students, with accommodations and alternate methods of assessment available for students who need them.</p> <ul style="list-style-type: none"> • This is also a federal requirement
Requirement 7	A state must generally allow test developers a minimum of three years to produce statewide tests that satisfy <i>Standards for Educational and Psychological Testing</i> and similar test-quality guidelines.
Requirement 8	<p>A state must ensure that educators receive professional development focused on how to optimize children's learning based on the results of instructionally supportive assessments.</p> <ul style="list-style-type: none"> • Policymakers should provide educators with professional development opportunities to increase their success in using the type of system (a new approach to assessment systems) the Commission advocates.
Requirement 9	<p>A state should secure evidence that supports the ongoing improvement of its state assessments to ensure those assessments are (a) appropriate for the accountability purposes for which they are used, (b) appropriate for determining whether students have attained state standards, (c) appropriate for enhancing instruction, and (d) not the cause of negative consequences.</p> <ul style="list-style-type: none"> • Assessments need to be continuously evaluated and improved upon.

Source: Legislative Service Office from information provided by WDE.

Wyoming develops an integrated statewide assessment

In June 2004, rather than attempting a modification of WyCAS, WDE released a Request for Proposal (RFP) for the development of a new statewide assessment that would not only address the federal accountability requirements, but would also contain an instructionally supportive component, as well as a formative assessment component. Interestingly, the formative assessment component was removed from the RFP released in 2008 for re-bid of PAWS, which caused Wyoming to lose \$387,770.93 during the first contract period for two online assessments (*Stanford Learning First & Bookette-Skill Writer*) that it has never used. These online assessments were part of a portal (WEdGate) that included various educational tools for teachers, which according to WDE were not routinely used by the districts. WDE expended \$1.15 million on this contract through June 30, 2009.

It should be noted that according to WDE, the term “formative assessment” has evolved since the 2004 RFP. The 2004 RFP was written using the “commercial” definition, which caused assessments like *Stanford Learning First & Bookette-Skill Writer* to be used. Currently, according to WDE, the definition of formative assessment has evolved into a process that is used continually as part of instruction to diagnose where students are in their learning and subsequently make changes to their instruction. For purposes of this audit however, we will use the earlier definition, given the exact wording in the initial RFP.

Part of the instructionally supportive framework was the “*traffic-signal-project*”, which provides teachers feedback on their incoming students. If a student’s report has a “*red-light*” in a certain skill area, the report indicates to the teacher that student needs more instruction in that area. Following this logic, a “*yellow-light*” indicates the student may need more instruction in an area, and a “*green-light*” indicates a student does not need additional instruction in an area.

This manner of reporting is part of an overall instructionally supportive assessment. In other words, these reporting categories are used to report student performance related to 23 cognitive skills (see Table 1.3 below), based on Wyoming Content and Performance Standards. Each skill also includes an assessment description, which briefly depicts the nature of the

skill to assist teachers with instructional decisions.

Reporting students' results through this process not only includes the different color traffic signals, but also includes raw scores and scale scores, as well as performance levels. According to WDE, traffic signal participants are trained in reviewing the assessment descriptions, as well as how to develop suitable guidance to their peers in the form of instructional suggestions. WDE also provides overall training to teachers and district administrators throughout the year in the areas of test administration, test security and ethics, as well as statistical analysis of item-level performance.

An important part of an instructionally supportive assessment is the proper mix of item types that test each skill area. Both the 2004 and 2008 RFPs contain language specific to the types of items to be included in the assessment. Generally, the RFPs called for a rich mix of multiple measures and item types including grade appropriate multiple choice and open-ended testing such as constructed-response, extended-response and performance-based tasks.

Such a mix of item types assists with ensuring alignment to the statewide student content and performance standards. See the following table for additional information.

Table 1.2
Specific Examples of Various Item Types in the 2004 and 2008 RFPs

Year	Page No.	Item Type Requirement Language
2004	16	Section 7.1 (E): Include multiple measures and item types including grade appropriate <i>multiple choice and open-ended testing such as constructed-response, extended-response and performance-based tasks</i> , to ensure alignment to the statewide student content and performance standards. [emphasis added]
	23	Section 9.5 (h)(i): The end-of-year summative reading, and mathematics tests will include items that are <i>dominantly selected-response items, and long constructed-response items not to exceed 10 minutes</i> (which also may vary by grade level). [emphasis added]
	24	Section 9.5 (m): Although the State envisions having a unique

Year	Page No.	Item Type Requirement Language
		assessment system that <i>includes only items that are aligned to Wyoming Content and Performance Standards</i> , the State will consider the use of items described below so long as they are also aligned with Wyoming's Content and Performance Standards, described more fully herein. [emphasis added]
2008	25	Section 4.1.1 (v) (E): Include multiple measures and item types including grade appropriate <i>multiple choice and open-ended testing such as constructed-response, extended-response and performance-based tasks</i> , to ensure alignment to the statewide student content and performance standards. [emphasis added]
	45	Section 5.2.1 (A)(1)(b): Reading and mathematics (Grades 3-8) and science (Grades 4, 8, and 11), each had two (2) sessions, an <i>online session composed of multiple-choice (MC) items only</i> and a paper-and-pencil session composed of <i>short and extended constructed-response (CR) items only</i> . [emphasis added]
	45	Section 5.2.1(A) (1) (c): In 2008, <i>reading constructed-responses and writing prompts for Grades 10-11 were administered online</i> . For all other grades, the writing assessment was administered only in a paper-and-pencil mode. [emphasis added]
	45	Section 5.2.1(A) (1) (d): In 2009, <i>reading CR [constructed response] items and writing prompts in Grades 6-8 will be administered online</i> . [emphasis added]
	50	Section B(1)(a): The PAWS reading and science tests include items that are <i>predominantly multiple-choice</i> ; in addition, <i>a few (3-4) short constructed-response items and 1 or 2 extended constructed-response items not to exceed 10 minutes</i> (which also may vary by grade level) are part of the test. [emphasis added]
	50	Section B (1) (b): The mathematics assessment includes <i>multiple-choice and short-response items only</i> . [emphasis added]
	55	Section 5.2.5(A) (1). ...This mixed mode [online and pencil and paper] administration has <i>all multiple-choice (MC) items delivered online for Grades 3-8, 10 and 11; the constructed-response (CR) items for reading and writing are also delivered online for Grades 10-11</i> . In 2009, plans are to administer the CR items for reading and writing online for Grades 6-8. CR items for reading and writing prompts for Grades 3-5, as well as all CR items for mathematics and science, will continue to be administered on paper. [emphasis added]

Source: Legislative Service Office from information provided by WDE.

As discussed earlier, PAWS was designed to test against 23 cognitive skills. Of those 23 skills, eight (8) are in reading, four (4) in writing, and eleven (11) in mathematics. Some of the item types are weighted more heavily than others. The *traffic-signal-reports* developed by WDE, show how well students perform against these skills.

The following tables illustrate additional information on the cognitive skills as well as the number of item types for PAWS.

Table 1.3
PAWS 23 Cognitive Skills

Subject	Skill Descriptions
Reading	<p><i>Functional Texts</i></p> <ol style="list-style-type: none"> 1) Determine information's relevance and importance. 2) Select and apply information for a task. <p><i>Expository Texts</i></p> <ol style="list-style-type: none"> 1) Understand main points and supporting details. 2) Recognize expository organization and its use. 3) See relationship of text's content to broader issues/topics. <p><i>Narrative Texts</i></p> <ol style="list-style-type: none"> 1) Identify the development of basic story elements. 2) Understand a story's plot development. 3) Identify a story's theme(s) and its (their) development.
Writing	<p><i>Idea Development</i></p> <ol style="list-style-type: none"> 1) The writer develops the content of the message through the use of details. <p><i>Organization</i></p> <ol style="list-style-type: none"> 1) The writer builds the organizational structure to support the purpose and effectiveness of the writing. <p><i>Voice</i></p> <ol style="list-style-type: none"> 1) The writer uses effective language to communicate to the audience in a way that is individual, compelling, and engaging. <p><i>Conventions</i></p> <ol style="list-style-type: none"> 1) The writer develops the mechanical correctness of the piece including spelling, capitalization, punctuation, and grammar.
Mathematics	<i>Number Operations and Concepts</i>

Subject	Skill Descriptions
	<p>1) Understand ways to represent numbers, relationships among numbers, and number systems. 2) Develop the connection between conceptual understanding and computational proficiency.</p> <p><i>Geometry</i></p> <p>1) Specify locations and describe spatial relationships using coordinate geometry and other representational systems. 2) Analyze characteristics and properties of two and three dimensional geometric shapes. 3) Apply transformations and use symmetry to analyze mathematical situations.</p> <p><i>Measurement</i></p> <p>1) Understand measurable attributes of objects and the units, systems, and processes of measurement. 2) Apply appropriate techniques, tools, and formulas to determine perimeter, area, or volume.</p> <p><i>Algebra</i></p> <p>1) Understand patterns, relations, and functions. 2) Use mathematical models to represent and understand quantitative relationships.</p> <p><i>Data Analysis and Probability</i></p> <p>1) Collect, organize, and display relevant data to answer questions and use appropriate statistical methods to analyze the data. 2) Develop and evaluate inferences and predictions based on data.</p>

Source: Legislative Service Office from information provided by WDE.

Table 1.4
PAWS Item Types and Score Percentages

Grade	Test	Multiple Choice (MC)		Constructed Response (CR)	
		No. of items	% of score	No. of items	% of score
3	Reading	43	78%	5	22%
3	Math	50	83%	5	17%
7	Reading	40	71%	6	29%
7	Math	56	85%	5	15%
11	Reading	40	71%	6	29%
11	Math	63	86%	5	14%

Source: Legislative Service Office from information provided by WDE.

Tremendous cost associated with PAWS There is no doubt that PAWS is expensive. Wyoming will have expended \$78,561,068 on PAWS through BFY 2011 for an integrated statewide assessment. In addition, of the total amount expended from BFY2005 through BFY 2009 (\$43,489,397.98) for PAWS, 85% or \$36,964,537.62 was expended for consulting and production services.

Also, of the total amount of encumbered monies from BFY2009 (\$6,242,942.35) 92% or \$5,734,442.95 is encumbered for consulting and production services. Finally, the entire unobligated amount from BFY 2009 (\$3,329,930.16) will be used for consulting and production services.

See Chapter 2, Tables 2.1 through 2.5 for a detailed cost analysis of PAWS.

School Finance Recalibration Committee considers using PAWS for accountability purposes During the 2010 Wyoming Budget Session, HB 1, Section 334 created the School Finance Recalibration Committee (Committee). As part of the Committee's overall charge to study recalibration of the education resource block grant model, it is also required to "*review methods of assessment of student achievement under the current model and recommend student assessment policies and procedures in the recalibrated system of education.*"

Although the Committee did not specifically study PAWS, it identified it as an assessment that can be used to measure how well districts are doing as part of the statewide accountability system it is studying. With the caveat however, that it may need to develop a subcommittee to review PAWS as needed. This of course could include adjustments, etc., with respect to pending ESEA reauthorization.

CHAPTER 2

PAWS is not being used fully as an Integrated Statewide Assessment

Finding 1: PAWS is not being used as originally envisioned. For example, it is not being fully used as an integrated statewide assessment that is instructionally supportive, includes multiple summative assessments and includes a formative assessment component. The integrated approach is also extremely costly.

The Proficiency Assessments for Wyoming Students (PAWS) is not being fully administered as an integrated statewide assessment. The original intent was to measure twice per year what students know by using multiple summative statewide assessments that measure proficiency against student content and performance standards. It is also used for federal accountability purposes.

In addition, formative assessments were envisioned to supplement the summative assessments in order to gauge a student's progress throughout the school year. In other words, to make daily instructional decisions about students based on multiple assessments.

It should be noted however, that No Child Left Behind (NCLB) does not require states to create statewide assessments that are instructionally supportive or to include formative assessments. Stated another way, Wyoming chose to go above and beyond what was required by NCLB.

Although envisioned as such, PAWS is not fully instructionally supportive, nor does it include a formative assessment component. More specifically, WDE does not know the extent to which schools and districts use various instructionally supportive reports and suggestions based on PAWS scores, nor are formative assessments used as part of PAWS, to supplement the summative assessment administered each year.

As a result, districts may not be using PAWS to support instructional decisions, which calls into question the \$78,561,068 Wyoming will spend on PAWS through BFY 2011 for an integrated statewide assessment.

For example, of the total amount expended from BFY2005 through BFY 2009 (\$43,489,397.98) for PAWS, 85% or \$36,964,537.62 was expended for consulting and production services. Also, of the total amount of encumbered monies from BFY2009 (\$6,242,942.35) 92% or \$5,734,442.95 is encumbered for consulting and production services. Finally, the entire unobligated amount from BFY 2009 (\$3,329,930.16) will be used for consulting and production services.

It should also be noted that Wyoming has expended \$387,770.93 during the first contract period for two online assessments (*Stanford Learning First & Bookette-Skill Writer*) that it has never used, resulting in a loss to Wyoming with no benefit. This loss resulted in the decision to not use the formative assessment components during the first and second contract periods.

These online assessments were part of a portal (WEdGate) that included various educational tools for teachers, which according to WDE were not routinely used by the districts. WDE expended \$1.15 million on this contract through June 30, 2009.

Wyoming Legislative Intent

Wyoming statutes generally mirror NCLB requirements

W.S. 21-2-304 (a) (v) requires the State Board of Education, through the Superintendent and in consultation and coordination of local school districts, to implement a statewide assessment to measure individual student performance in grades three (3) through eight (8) and grade eleven (11). PAWS is used to assess students' proficiency in reading, math, writing, and science (as

of Spring 2008).

Wyoming statutes, as discussed above, are generally consistent with NCLB. More specifically, 20 USC 6311, Section 1111 (b) (3) requires a single statewide assessment for students in grades 3 through 8 in mathematics, reading, language arts and science (no later than SY 2007-2008). It also requires that assessments measure against challenging state academic content and achievement standards (Section 111 (b) (2)).

In Wyoming, student proficiency is assessed against Wyoming's "common core of knowledge" as annotated in W.S. 21-9-101 (b). This focuses primarily on reading, writing, and mathematics for grades one (1) through eight (8). It should be noted, that pursuant to W.S. 21-2-304 (a) (iii) "student content and performance standards" are also required in order to measure against the "common core of knowledge" and "common core of skills" annotated in W.S. 21-9-101 (b) and WDE Chapter 31 Rules (Graduation Requirements), Section 8. The standards as discussed above were adopted by the Wyoming State Board of Education on November 19, 2008.

The standards are listed on WDE's webpage at the following address: <http://www.k12.wy.us/SA/standards.asp>.

Wyoming statutes require statewide assessments to be instructionally supportive and to include a formative assessment component

Although Wyoming statutes generally mirror the language of NCLB as illustrated above, they also go further by requiring that Wyoming's statewide assessment integrate an instructionally supportive component, as well as a formative assessment component into an integrated statewide assessment. According to W.S. 21-2-304 (a) (v), the statewide assessment system is

"...comprised of a coherent systems of measures that when combined, provide a reliable and valid measure of individual student achievement for each public school and school district within the state, and the performance of the state as a whole. Improvement of teaching and learning in schools and fostering school program improvement shall be the primary purposes of statewide assessment of student performance in Wyoming."

As discussed earlier and as noted above, the primary purpose of Wyoming's statewide assessment is to improve *“teaching and learning,”* as opposed to basic testing of students. For example, prior to statutory changes in 2004 (Session Laws of Wyoming, 2004, Ch. 109, Section 1), the focus was primarily on measuring students: *“...implement a statewide assessment system for measuring student progress based upon uniform educational program and student content and performance standards imposed by law and by board rule and regulation.”*

After 2004 however, that focus shifted to the use of multiple summative and formative assessments to assist teachers in making instructional decisions based on students' performance against established content and performance standards.

For example, W.S. 21-2-304 (v) (B) states

“...The structure and design of the assessment system shall allow for the comprehensive measurement of student performance through assessments that are administered each school year simultaneously on a statewide basis and through assessments administered periodically over the course of the school year which are designed to provide a more comprehensive and in-dept measurement of subject areas aligned to the state content and performance standards.”

As part of the overall statewide assessment, the results from formative assessments were envisioned for use throughout the year by teachers, principals, and administrators to make daily instructional decisions. For example, in WDE's PAWS BFY 2007 budget request PAWS is described: *“The new system increases the number of grade levels to be assessed from three to seven and PAWS will be an instructionally supportive, on-line, formative/summative state assessment system, which is a totally new approach to the state test.”*

As stated previously, NCLB does not require that statewide assessments be instructionally supportive, nor does it require that statewide assessments include a formative assessment component.

Wyoming Emphasized Instructional Support and Formative Assessment in the Original RFP

The Wyoming Statewide Task Force on Student Assessment and Educational Accountability mirrored a 2001 report by the Commission on Instructionally Supportive Assessment

The Wyoming Statewide Task Force on Student Assessment and Educational Accountability (Task Force) was created under 2003 Wyoming Session Laws, Chapter 208, Section 201, to “*review the statewide student assessment system and to develop recommendations on modifications to the system as necessary to meet federal requirements under the NCLB and to ensure year-to-year changes in individual student performance and school performance.*”

In October 2003, the Task Force released its final report and recommendations, which became the basis for an initial Request for Proposal (RFP), which was opened June 2, 2004. The recommendations from the Task Force also became the basis for changes to Wyoming Statutes, Title 21, Chapter 2 (The Administration of the State System of Education at the State Level), as discussed above.

The Task Force’s recommendations were primarily based on a report issued in 2001 by the Commission on Instructionally Supportive Assessment (Commission). The Commission, chaired by Dr. James Popham (also chair of Wyoming’s Technical Advisory Committee), issued *Building Tests to Support Instruction and Accountability: A Guide for Policymakers*, which offers advice to state policy makers to assist them in the development of tests that are supportive of instruction and accountability purposes.

The advice is contained within nine (9) requirements, which according to the Commission’s report, “*must be met to ensure a responsible educational assessment system for the improvement of learning.*” See Chapter 1 (background) for more detailed information about the nine (9) requirements.

According to WDE officials, PAWS was developed to be an instructionally supportive assessment. First, it provides information on how well students are performing related to Wyoming content and Performance Standards in reading, writing, and mathematics. It also provides professional development opportunities to educators and district

administrators related to the mechanics of PAWS and how to use assessment results. Finally, it provides additional skill-level reporting categories to assist teachers in interpreting and addressing the specific needs of students.

Although some education experts state that PAWS is the “*best in the nation*”, there are some valid criticisms from other education experts that an instructionally supportive standardized assessment is not the best route for states. According to FairTest, The National Center for Fair and Open Testing, in an August 27, 2007 reply to the Commission’s report, there are some concerns with such an approach.

The following are excerpts from FairTest’s reply.

- *“The Commission makes a fundamental error in maintaining that large-scale, state exams can play a primary role in providing instructionally supportive assessment. The Commission’s recommendations are therefore not adequate for the goal of developing an assessment system that can support high-quality assessment and accountability-though the recommendations are valuable for any standardized testing that might be included within a broader assessment system.”*
- *“The report assumes not only the existence of, but also the need for and value of having state-centralized, standardized testing as the primary, largely controlling, measure of student learning and even, it would appear, the primary means of formative assessment (that is, of assessment intended to shape subsequent instruction for the particular student(s)). However, because of the serious limits inherent in any feasible large scale, standardized testing, such tests should not play the primary role in measuring learning outcomes and cannot adequately play much of a role at all in formative assessment-which is the more important assessment task.”*

Despite the criticisms of a state-centralized and standardized test as the primary component of student learning, Wyoming chose to move forward as annotated in the Commission’s and Task Force’s reports. This is reflected not only in the statutes

discussed above, but as well in initial and subsequent RFP language.

The 2004 RFP envisioned use of multiple summative assessments, as well as formative assessments to aid instruction

Language in the 2004 RFP (0343-L) is consistent with Wyoming statute, but also goes beyond what NCLB requires. For example, illustrative language in Section 7.1 (H) of the RFP states that the statewide assessment system shall

“Provide a measure of accountability to enhance teaching and learning in Wyoming and in combination with other measures and information, assist school districts in determining individual student progress. The assessment results shall be reported to students, parents, schools, school districts and the public in an accurate, complete and timely manner and shall be used in conjunction with a school district’s annual assessment to design educational strategies for improvement and enhancement of student performance.”

Section 8.1 goes on to state

“Because the purpose of the new assessments will be to both contribute to a newly established accountability system and, at the same time, to support instructional improvement, the new assessments can accurately be characterized as instructionally supportive accountability tests. The statewide assessment system will be comprised of summative and formative assessments in reading, writing, and mathematics, which will be administered in grades three (3) through eight (8) and in grade eleven (11).”

WDE Reemphasized Instructional Support in the 2008 RFP, but Eliminated the Formative Assessment Component

The 2008 RFP maintained the importance of instructional support but eliminated the

According to WDE officials, changes to the 2008 RFP (RFP No. 0056-S) retained the original requirement of an instructionally supportive assessment, but eliminated the requirement of formative assessments as part of an integrated statewide assessment.

requirement of formative assessment For example, Section 4.2 (Rationale for and purpose of the Wyoming Instructionally Supportive Accountability Tests) describes what specifically makes PAWS an instructionally supportive assessment.

“Annually, after the assessment window closes, the Agency empanels groups of experienced Wyoming teachers to engage in a systematic review of the item level performance within each skill reporting category. The purpose of this review is to formulate “peer to peer” instructional suggestions related to the tested content standards. The “Traffic Signal Report” provides guidance at the skill reporting category and is simplified to three (3) color-coded levels: Green, no additional instruction on this skill category seems needed; Yellow, additional instruction on this skill category may be needed; and Red, additional instruction on this skill category definitely seems needed.”

Finally, Section 5.2.14 (A) (3) clarifies the removal of the formative assessment component as follows:

“Initially, the PAWS system was conceptualized as being comprised of a summative and formative assessment (including some classroom-embedded assessments) that would provide valid and reliable data to inform performance evaluations at all levels. Currently, it is implemented as a summative assessment only.”

According to WDE officials, the decision to deemphasize the formative component of the integrated statewide system was made since formative assessments are not being administered as a formal part of the integrated statewide assessment. Although this rationale appears to be a logical reason to make adjustments to the integrated statewide assessment, WDE is technically out of compliance with state statute (W.S. 21-2-304 (a) (v) and W.S. 21-2-304 (v) (B)) for not implementing the statewide assessment as statutorily envisioned.

Instructionally Supportive and Formative Assessment Components not Fully Utilized

WDE encourages all districts to use the traffic-signal process; however, this instructionally supportive component is not tracked by WDE

According to WDE officials, an aspect of the instructionally supportive portion of PAWS is made available to Wyoming districts through its *traffic-signal process*. The process is formally described in the PAWS Technical Manual (Section 7.5). However, WDE does not track how often the process is used by districts, which disallows WDE from gauging the effectiveness of PAWS for instruction purposes.

More specifically, WDE *“provides guidance at the skill-reporting category level when reporting PAWS results.”* For example, PAWS measures 23 cognitive skills (i.e. eight (8) in reading, four (4) in writing, and eleven (11) in mathematics), which students need to achieve proficiency of Wyoming’s student content and performance standards.

WDE officials stated this report is intended to be interpreted independently of the proficiency and scale scores of PAWS, which indicate levels of proficiency. Rather, it is *“intended to apply only to the specific set of student performance in need of instructional attention.”*

The process occurs at the end of each test administration, according to WDE officials. It is a peer review process in which panelists of Wyoming educators are selected to participate. Panelists meet for two days to review items and assessment descriptions, in order to generate instructional suggestions for their peers.

Also, teachers and district officials receive reports with respect to the number of points a student needs to earn within each skill category to receive a *“green light”, “yellow light”, or “red light.”*

However, WDE does not track overall district participation in this process, nor does it track the use of instructional reports by teachers, schools, and districts. Rather, it makes reports available on its website for all districts to use. Since 2006, the reports have been made available to teachers, schools, and districts in the summer following each PAWS administration,

according to WDE officials. Reports are available at the student level, the school level, and the district level. Reports can also be used by current school year teachers or teachers with incoming students from the previous years and can provide demographic reports broken out by gender, race, special education, etc.

Also, “instructional suggestions”, based on students’ PAWS scores, are available on WDE’s website to assist with instructional improvement and change. However, according to WDE officials, much like the reports discussed above, WDE does not insist that teachers, schools, or districts use instructional suggestions. Nor, does it track the use of instructional suggestions.

By not requiring and/or tracking this type of information however, WDE is missing an opportunity to gather, analyze, and report vital information to the Legislature on the effectiveness of Wyoming’s integrated statewide assessment. Given that Wyoming’s statewide assessment is distinctly different than what other states use, it is problematic not to somehow track the effectiveness of its innovative approach, at least through the usage of its instructionally based reports and suggestions.

WDE officials stated however, that they will begin to discuss ways in which WDE can monitor school and district use of reports and instructional suggestions in order to provide this type of information to the Legislature; for example recording the number of “hits” on its webpage, etc.

Multiple summative assessments and formative assessments have not been used as envisioned by the Legislature

According to WDE officials, the original intent of PAWS was to include multiple summative and formative assessments throughout the school year. The original RFP stated that several dominant elements “will distinguish the new state assessment system,” one of which is year-round summative and formative assessments. However, PAWS is not being administered more than once per year, nor does it include formative assessments throughout the year.

The rationale was that students would be able to take a summative assessment (early PAWS) in January of a school year, and “bank” those scores for AYP purposes that Spring. If however, the student did better during the Spring administration

of PAWS (official PAWS), the student could use the better scores that Spring.

According to the initial RFP, *“The use of year-round summative assessments by districts might provide an alternative route to the demonstration of proficiency on the single end-of-year summative assessment. For instance, with respect to individual students, if a student demonstrated proficiency on year-round summative, in-depth, focus area assessments, s/he would not be required to take the single on-demand end-of-year summative assessment.”*

WDE officials stated that 2007 was the first year to “bank” student scores, which according to WDE, was essentially a field test. However, as a result of not having enough time to score multiple exams and to determine which students should retake the exam, the decision was made to do away with dual test administration. This decision was made partly based on information provided by the PAWS Redesign Task Force, which met in April, 2007.

Some of the problems encountered during the administration of both assessments included wasted instructional time, no time limits on tests, missed reporting deadlines, as well as taking too much time to administer the tests. The “banking” of student scores was eliminated as an option for school districts after 2007.

WDE expended \$387,770.93 for formative assessments that were never used Formative assessments are also not being used. As a result of not using formative assessments as originally envisioned, Wyoming expended \$5,462.93 for *Stanford Learning First* and \$382,308 for *Bookette-Skillwriter* for no benefit. According to WDE, districts did not routinely use these assessments, opting instead for Northwest Evaluation Association’s Measures of Academic Performance (MAP).

Formative assessments in reading, writing, and mathematics were originally envisioned to supplement the summative assessments. According to the initial RFP, *“these tools would support instruction and be aligned with the end-of-year summative and year-round summative assessments. The formative assessment tools should be periodic items or assessments that could be administered at multiple times during the school year and possibly scored locally.”*

According to WDE officials, the intent of the 2004 RFP (NO. 0343-L) was to provide districts with a palette of on-line, Wyoming State Standards-aligned, multi-grade level assessment programs. The state offered *Standard Learning First* and *Bookette Skill Writer* assessments to districts. *Standard Learning First* was dropped by the state, as it was not fully aligned with the Wyoming Content and Performance Standards. *Bookette* was dropped, as it was not utilized by the districts.

This action is consistent with what WDE stated in its BFY 2009 budget request, where PAWS was described as an instructionally supportive assessment, but without mention of formative assessments: “PAWS is an instructionally supportive, on-line and booklet form, summative state assessment system.”

Wyoming Expends Tremendous Resources

Through BFY 2011, Wyoming will expend \$78,651,068 on the development and administration of PAWS

Wyoming has expended a tremendous amount of federal and state money on the development and administration of an integrated statewide assessment system. However, because Wyoming has not fully implemented an integrated PAWS, Wyoming may not be receiving its money’s worth since the decision was made to develop a new statewide assessment.

In addition, WDE does not conduct any type of cost benefit analysis to assess the cost effectiveness of PAWS as an integrated statewide assessment. Such analysis is necessary in order to determine how well an integrated assessment works for Wyoming.

Analysis is also needed to determine whether the tremendous amount of money spent to date was expended wisely. This type of information could also be used with respect to changes after the reauthorization of the Elementary and Secondary Education Act.

The following tables provide additional budget and appropriation information. More specifically, Table 2.1 provides information on biennial budget requests from the Wyoming Department of Education (Agency 005) and Education-School Finance (205) related to PAWS development and implementation.

Table 2.2 provides expenditure information for these two agencies with respect to PAWS development and implementation. Table 2.3 illustrates legislative appropriations related to 900 series (consulting and production services), while Table 2.4 illustrates actual expenditures from a series of appropriation units for both agencies.

Finally, Table 2.5 provides additional information from WDE with respect to encumbered and unobligated amounts.

Table 2.1
PAWS Assessment Budget Requests
***Agency 205 (Education-School Finance) & Agency 005 (WDE)**

**Budget Fiscal Year (BFY)	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
2005-06	N/A	N/A	\$6,817,945.00	\$6,817,945.00
2007-08	\$9,403,286.00	\$1,398,928.00	\$6,838,165.00	\$17,640,379.00
2009-10	\$19,556,307.00	\$1,291,571.00	\$6,769,678.00	\$27,617,556.00
2011-12	\$18,606,307.00	\$1,109,203.00	\$6,769,678.00	\$26,485,188.00
Total	\$47,565,900.00	\$3,799,702.00	\$27,195,466.00	\$78,561,068.00

Source: Legislative Service Office from information provided by WDE and Education-School Finance budgets and IBARS.

*Monies from 4601 are appropriated from the School Foundation Program Account (Fund 009); monies from 6105 are appropriated from the General Fund (001); and monies from 6373 are 100% from a U.S. Department of Education grant re: the Elementary and Secondary Education Act (ESEA), Title I, State Program Implementation.

**\$5,151,270 in BFY 2005 for budget units 4601 and 6105 was requested for WyCAS.

Table 2.2
PAWS Expenditures
***Agency 205 (Education-School Finance) & Agency 005 (WDE)**
9/28/10

Budget Fiscal Year (BFY)	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
2005-06	\$2,588,708.00	N/A	\$3,383,324.00 (630) \$3,382,612.00 (E32)	\$9,354,644.00
2007-08	\$9,393,179.50	\$1,402,754.53	\$3,382,612.00 (630) \$3,391,566.00 (E32)	\$17,570,112.03

Budget Fiscal Year (BFY)	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
2009-10	**\$13,635,208.94	***\$1,007,998.62	+\$1,592,095.68 (630) ++\$329,338.71 (E32)	\$16,564,641.95
2011-12	N/A	N/A	N/A	N/A
Total	\$25,617,096.44	\$2,410,753.15	\$15,461,548.39	\$43,489,397.98

Source: Legislative Service Office from information provided by WDE and Education Finance budgets.

*Monies from 4601 are appropriated from the School Foundation Program Account (Fund 009); monies from 6105 are appropriated from the General Fund (001); and monies from 6373 are 100% from a U.S. Department of Education grant re: the Elementary and Secondary Education Act (ESEA), Title I, State Program Implementation.

**\$4,550,158.06 is encumbered under Object Code 0901. An additional \$231,117.71 was expended for Object Code 0901 under 4603; that unit includes a \$165,822.29 encumbrance. An additional \$340,179.89 was expended for Object Code 0901 under 4604; that unit includes a \$96,947.57 encumbrance. An additional \$20,287.44 was expended for Object Code 0901 under 4608; that unit includes a \$245,729.54 encumbrance.

***\$63,854.33 is encumbered under Object Code 0901.

+ \$1,030,430.56 is encumbered under Object Code 0901. \$758,884.76 is unobligated under Object Code 0901.

+ \$90,000.00 is encumbered under Object Code 0901. \$2,983,965.29 is unobligated under E32.

Table 2.3
***PAWS Appropriations Figures**
Consulting and Production Services (0900 Series)
September 16, 2010

0900 Series Appropriations	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
**2005-06				
0901	N/A	\$350,000		
0902	N/A	\$0	\$6,437,483.00	\$6,437,483.00
0905	N/A	\$90,000	\$0	
Subtotal	N/A	\$440,000	\$6,437,483.00	\$6,877,483.00
2007-08				
0901	\$4,002,346.00	\$350,000.00	\$5,412,787.00	\$9,765,133.00
0902	\$4,900,000.00	\$0	\$0	\$4,900,000.00
0905	\$0	\$90,000.00	\$0	\$90,000.00
Subtotal	\$8,902,346.00	\$440,000.00	\$5,412,787.00	\$14,755,133.00
2009-10				
0901	\$4,002,346.00	\$234,101.00	\$5,412,787.00	\$9,649,234.00
0902	\$15,053,021.00	\$90,000.00	\$0	\$15,143,021.00
0905	\$0	\$0	\$0	\$0
Subtotal	\$19,055,367.00	\$324,101.00	\$5,412,787.00	\$24,792,255.00
2011-12				
0901	\$3,052,346.00	\$209,101.00	\$5,412,787.00	\$8,674,234.00

0900 Series Appropriations	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
0902	\$15,053,021.00	\$0	\$0	\$15,053,021.00
0905	\$0	\$90,000.00	\$0	\$90,000.00
Subtotal	\$18,105,367.00	\$299,101.00	\$5,412,787.00	\$23,817,255.00
Grand Total	\$46,063,080.00	\$1,503,202.00	\$22,675,844.00	\$70,242,126.00

Source: Legislative Service Office from information provided in IBARS.

*Monies from 4601 are appropriated from the School Foundation Program Account (Fund 009); monies from 6105 are appropriated from the General Fund (001); and monies from 6373 are 100% from a U.S. Department of Education grant re: the Elementary and Secondary Education Act (ESEA), Title I, State Program Implementation.

**\$4,442,346 in BFY 2005 for budget units 4601 and 6105 was appropriated for WyCAS.

Table 2.4
***PAWS Expenditures**
Consulting and Production Services (0900 Series)

0900 Series WOLFS	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
2005-06				
0901	\$2,588,708.00	N/A	\$3,141,268.23 (630) \$1,948,702.76 (E32)	\$7,678,678.99
0902	N/A	N/A	\$0	\$0
0905	N/A	N/A	\$3,023.37 (E32)	\$3,023.37
Subtotal	\$2,588,708.00	N/A	\$ 5,092,994.36	\$7,681,702.36
2007-08				
0901	**\$9,393,179.50	***\$580,630.55	\$2,529,560.59 (630) \$2,294,473.00 (E32)	\$14,797,843.64
0902	\$0	\$0	\$0	\$0
0905	\$0	\$0	\$0	\$0
Subtotal	\$9,393,179.50	\$580,630.55	\$4,824,033.59	\$14,797,843.64
2009-10				
0901	****\$13,635,208.94	+\$173,074.83	++\$676,594.25 (630) +++ \$113.60 (E32)	\$14,484,991.62
0902	\$0	\$0	\$0	\$0
0903	\$0	\$0	\$0	\$0
Subtotal	\$13,635,208.94	\$173,074.83	\$676,707.85	\$14,484,991.62
2011-12				
0901	\$0	\$0	\$0	\$0
0902	\$0	\$0	\$0	\$0
0905	\$0	\$0	\$0	\$0
Subtotal	\$0	\$0	\$0	\$0

0900 Series WOLFS	Education-School Finance Assessment (4601)	WDE GF Assessment Administration (6105)	WDE Federal Funds Assessment (6373)	Total
Grand Total	\$25,617,096.44	\$753,705.38	\$10,593,735.80	\$36,964,537.62

Source: Legislative Service Office from information provided in WOLFS.

*Monies from 4601 are appropriated from the School Foundation Program Account (Fund 009); monies from 6105 are appropriated from the General Fund (001); and monies from 6373 are 100% from a U.S. Department of Education grant re: the Elementary and Secondary Education Act (ESEA), Title I, State Program Implementation.

**WDE set up an expenditure authority for \$4,889,893.50 for Object Code 0902, which it did not use.

***WDE set up an expenditure authority in WOLFS under Object Code 0905 for \$90,000, which it did not use.

****\$4,550,158.06 is encumbered under Object Code 0901. An additional \$231,117.71 was expended for Object Code 0901 under 4603; that unit includes a \$165,822.29 encumbrance. An additional \$340,179.89 was expended for Object Code 0901 under 4604; that unit includes a \$96,947.57 encumbrance (WEN). An additional \$20,287.44 was expended for Object Code 0901 under 4608; that unit includes a \$245,729.54 encumbrance (WEN).

+ \$63,854.33 is encumbered under Object Code 0901.

++ \$1,030,430.56 is encumbered under Object Code 0901. \$758,884.76 is unobligated under Object Code 0901.

+++ \$90,000.00 is encumbered under Object Code 0901. \$3,329,930.16 is unobligated under Object Code 0901.

Table 2.5
Explanation of Encumbrances and Unobligated Funds from Table 2.4

BFY	Budget Unit	Amount	WDE Explanation
2005	4601	\$4,889,893.50 Unused Expenditure Authority	0901 is set up for professional fees and 0902 as consulting services. Looking at the expenses, it seems that these payments were not made to consulting groups, but rather to professional groups or individuals. We work, in collaboration with SAO, to ensure that payments are coded to the object code most closely related to the actual service.
2007	6105	\$90,000 Unused Expenditure Authority	We follow SAO guidelines, from their website, which state, "When travel expenses are authorized and included as part of the contract, the payment is to be recorded <u>under the same object/sub-object code you are using to pay for the contracted service.</u> This also applies to copy charges, fax charges and other miscellaneous expenses made by (and billed by) contractors. This is necessary to comply with IRS guidelines." Sending WOLFS screenshot of series levels to show that the \$90,000 was spent in 0901, rather than 0905.
2009	4601	\$4,550,158.06 Encumbrance	This is the current contract with NCS Pearson from 12/2/08 through 12/2/12 for

BFY	Budget Unit	Amount	WDE Explanation
			deliverables specified in the contract. The \$4,500,158.06 encumbrance for appropriation unit 4601 under Object Code 0901 will be paid to NCS Pearson for the invoicing from March 2010, June 2010, and September 2010 pertaining to the PAWS contract. Funds are being withheld from NCS Pearson, Inc. pending settlement negotiations. The encumbrance will be expended from the 2009-2010 biennium funds, once settlement negotiations are completed and I receive approval from the Superintendent to pay Pearson.
2009	4303	\$231,117.71 Expense \$165,822.29 Encumbrance	Yes, this is the Early Learning Assessment budget, a breakout from 4601. The Early Learning Assessment Program uses the Instruction Foundations for Kindergarten (IF-K). This assessment is completed by all Kindergarten teachers each fall for each Kindergarten student. It assesses nine foundation areas based on the Wyoming Early Childhood Readiness Standards. Teachers and schools receive online reports of school readiness levels. The IF-K is a component of the Statewide Assessment System, though it is not Pearson product.
2009	4604	\$340,179.89 Expense \$96,947.57 Encumbrance	Wyoming Equality Network (WEN) Infrastructure Budget Unit
2009	4608	\$20,287.44 Expense \$245,729.54 Encumbrance	Wyoming Equality Network (WEN) Distance Education Budget Unit
2009	6105	\$63,854.33 Encumbrance	This is also the current contract with NCS Pearson from 12/2/08 through 12/2/12. Again, for more information, please contact Alan Moore. The \$63,854.33 encumbrance for appropriation unit 6105 under Object Code 0901 will be paid to NCS Pearson for the invoicing from March 2010, June 2010, and September 2010 pertaining to the PAWS contract. Funds are being withheld from NCS Pearson, Inc. pending settlement negotiations. The encumbrance will be expended from the 2009-2010 biennium funds, once settlement negotiations are completed and I receive approval from the Superintendent to pay Pearson.

BFY	Budget Unit	Amount	WDE Explanation
2009	6373	\$1,030,430.56 Encumbrance \$758,884.76 Unobligated	<p>The 0901 encumbrances in this federal budget are for NCS Pearson 12/2/08 – 12/2/12 and to individuals, assessment groups and the University of Wisconsin. I do not have the information as to the purpose. Please contact Alan Moore for this information. The nature of the contracts under appropriation 6373 (630)/object code 0901 is to pay for services of members of the Technical Advisory Committee for the PAWS assessments, members of the Standards Review Steering Committee, and Content Review Committee leaders of the current review and revision of the Wyoming Content and Performance Standards required by the State Board of Education. In addition a contract with R&J Technology pays for text to speech technology for the PAWS administration. Finally, the contract with the University of Wisconsin is a contract through the World-Class Instruction Design and Assessment (WIDA) Consortium (affiliated with the University of Wisconsin) which now provides the Access for ELLs assessment of English language proficiency for English language learners (ELL) required under NCLB.</p> <p>There are no unobligated funds in this budget, currently. All funds are expended or encumbered. This funding ended 9/30/10 and we have until 12/31/10 to finalize all expenses and draw down reimbursement from the U.S. Department of Education. The unobligated balance of \$758,884.76 on 9/28/10 is acceptable to the feds, as long as the unobligated balance is zero on 9/30/10, as it was. I would say that a budget the size of this one is difficult to pin down to the penny until the very end of an obligation period. There is a concerted effort during those last few weeks, on the part of the program managers and division directors, to communicate with vendors and contractors with outstanding invoices, asking them to</p>

BFY	Budget Unit	Amount	WDE Explanation
			submit. Then there is another challenge of entering these expenses into WOLFS and waiting for processing.
2009	6373	\$90,000 Encumbrance \$3,329,930.16 Unobligated	<p>WDE, when doing strategic planning, has established a fund for those projects or initiative that surfaced after the biennial budget planning. These projects are presented from the Superintendent and/or his Deputy and discussed in the department leadership team. Funding for these initiatives are mass encumbered and upon completion and invoiced, it is again discussed as to the most appropriate funding for that initiative. For federal funds, the department uses extra caution to assure that the federal guidelines are followed for the appropriate use of those funds. The \$90,000 mass encumbered will be used for a project that supports the federal Assessment award purpose and guidance.</p> <p>The unobligated \$3,329,930.16 in 6373-E32-2009, Object Code 0901 will be used to pay for future contracts similar to those in the past, consistent with the guidelines of this Federal budget including, but not limited to a) standards review committee work; b) catering services and meeting facilities for assessment and standards review projects; c) meetings and workshops in support of the Body of Evidence system; d) professional development and training for WIDA and Access for ELLs assessment; and the Pearson contract.</p>

*Legislative Service Office from information provided by WDE.

As seen from the above tables, the majority of money expended on PAWS development and administration is expended in the 900 series (consulting services). For example, of the total amount expended from BFY2005 through BFY2009 (\$43,489,397.98) for PAWS, 85% or \$36,964,537.62 was expended for consulting services. Also, of the total amount of encumbered monies from BFY2009 (\$6,242,942.35) 92% or \$5,734,442.95 is encumbered for consulting and production

services. Finally, the entire unobligated amount from BFY 2009 (\$3,329,930.16) will be used for consulting and production services.

See Chapter 5 for a further breakdown of consulting services required to implement an integrated statewide assessment and to pass the federal peer review by the U.S. Department of Education.

Districts use Measures of Academic Performance (MAP)

Wyoming districts use Northwest Evaluation Association's Measures of Academic Performance (MAP)

Although formative assessments are not being currently used as part of the PAWS program, districts in Wyoming appear to use the Northwest Evaluation Association's Measures of Academic Performance (MAP) assessment for their Response to Intervention (RTI). According to *Wyoming's State Performance Plan for Special Education*, the state has an RTI framework for schools to follow.

RTI is a collaborative effort in Wyoming, and is incorporated by WDE into its statewide education goals. It is designed to identify students with learning disabilities through collaboration among federal and state efforts (e.g. NCLB Reading First, Title I, Individuals with Disabilities Act (IDEA-2004), and the Reading Assessment and Intervention Program (W.S. 21-3-401)). RTI includes general, remedial, and special education tiers. The tiers include strategies or interventions to monitor student progress to determine ultimately which students are most in need of intensive services.

Also, districts appear to be using MAP in their student assessment systems. Pursuant to W.S. 21-3-110 (a) (xxiv) districts are required to develop student assessment systems to measure student performance against student content and performance standards, relative to the common core of knowledge and skills prescribed under W.S. 21-9-101 (b). Their systems should be integrated with the statewide assessment system (i.e. PAWS).

It should also be noted that all districts participating in the Bridges Summer School Program are required to assess all

students from the Spring of 2nd grade through the Fall of 9th grade (Spring and Fall semesters) using MAP (i.e. Board of Education Rules-Chapter 33, Section 7 (e)). According to WDE officials, only two districts (Sheridan #3 and Sublette #9) do not participate in the Bridges Summer School Program.

It is interesting to note, that according to WDE officials, NWEA was one of the respondents to the original RFP. Since the selection committee did not award the statewide assessment contract to NWEA however, MAP was not considered as part of the PAWS Program. Officials went on to state that WDE has always and continues to promote assessments such as MAP to districts, as one of multiple sources of information, also including PAWS, when considering instructional improvements and learning goals.

Recommendation: WDE should take action to determine how often districts use the “traffic-signal” reports and its instructional suggestions offered as part of an instructionally supportive statewide assessment. In addition, WDE should consider developing a reporting mechanism for districts to use in order to report this information, as well as identify ways that districts actually use the reports.

Wyoming has expended a tremendous amount of resources and money to go beyond the minimum requirements of NCLB by creating a statewide assessment that is instructionally supportive. The instructionally supportive nature of PAWS is required by Wyoming statute (pages 23-24). By not tracking how often districts use the “traffic-signal” reports and instructional suggestions, WDE is missing an opportunity to provide important information to the Legislature prior to the reauthorization of ESEA. The Legislature would benefit by knowing if the instructionally supportive component of PAWS has been successful and cost beneficial.

Recommendation: WDE should continue to work with districts to tie in existing use of formative assessments, more specifically NWEA MAP, with PAWS. These assessments throughout the year can provide important information to districts and schools with respect to instructional changes.

Wyoming's original intent, as evidenced by the initial RFP, was to include some sort of formative assessment component as part of an integrated statewide assessment. Even though online formative assessments have not occurred as originally envisioned by WDE, districts are still using NWEA MAP as part of their Response to Intervention (RTI), as well as their district student assessment systems. WDE should take the opportunity to analyze and review MAP scores at the district level, school, and student levels, to possibly make future decisions with respect to ESEA reauthorization.

Recommendation: At a minimum, WDE should track the cost-effectiveness of PAWS by comparing the dollars expended for professional services against the use of instructional components of the integrated statewide assessment.

As seen from the above, Wyoming has expended a tremendous amount of money to consultants related to its integrated statewide assessment. However, for WDE to neglect making any type of effort to measure whether it is benefiting from these expenditures is to deprive the Legislature of vital information. Especially, given that ESEA will likely be reauthorized in the next year or two. The Legislature will most certainly need solid cost data to determine whether an instructionally based statewide assessment should continue into the future.

Finding 2: The Wyoming Comprehensive Assessment System (WyCAS) could have been modified to meet the requirements of NCLB, according to past and current WDE officials.

WyCAS in its original form would not have met the requirements of NCLB. However, modifying the assessment to include 3rd through 8th and 11th grades and to include an additional content area (Science) to measure individual student performance, would not have been difficult, according to past and current WDE officials. WyCAS was originally administered in 4th, 8th, and 11th grades.

According to WDE officials, WyCAS set very high standards for proficiency relative to those of other states. From what we understand, it was developed as part of a major school reform effort launched in 1997. This reform brought Wyoming from ranking in the bottom half of the states participating in the National Assessment of Educational Progress (NAEP) in 1998 to a position among the top ten (10) states in both math and reading on the 2005 NAEP.

However, WyCAS was designed to track the performance of whole schools, rather than individual children. It may not have been easily adapted to meet the requirements of NCLB, which holds schools accountable for the performance of specific subgroups of students.

Although WyCAS data was aggregated and reported to the U.S. Department of Education (Department) aggregated by school, individual student reports were generated for parents. For example, parents received reports showing proficiency levels by subject. Modifications could have occurred to allow more disaggregated student reporting for NCLB purposes. As a result of Wyoming moving towards an integrated statewide assessment as opposed to a modified WyCAS to simply assess yearly outcomes, Wyoming has incurred additional costs.

For the most part, past and current WDE employees believe WyCAS could have been modified to meet NCLB requirements

For the most part, past and current WDE officials, as well as other experts in the education field, believe WyCAS could have been modified to meet NCLB requirements. However, in its original form, they also stated WyCAS would not have met NCLB requirements. In other words, modifications would have had to occur, but not necessarily the creation of the instructionally supportive and formative assessment components contained in PAWS.

For example, WyCAS in its original form was designed to **1)** aggregate student test data to illustrate whole school performance; **2)** use a minimum number of “items” for each subject area to illustrate whole school performance; **3)** test students in 4th, 8th, and 11th grades; **4)** test students using a “*matrix sample design*” where students take different tests under a common block of items, which is designed to measure school performance and assess instructional effectiveness; and **5)** align with the 1998 Wyoming Content Standards.

We received consensus among past and current WDE officials, as well as experts in the field, that the following modifications could have been undertaken to help WyCAS become NCLB compliant.

Utilize Existing Reports for Individual Students

According to the Task Force’s Recommendation #1, the new assessment had to **1)** include an annual *summative assessment* designed primarily to measure school, district, and state performance, which (with other measures) can provide information about individual student performance useful to teachers and parents; and **2)** include *embedded tools and assessments* (that lead to the summative assessment) designed to provide useful and timely information regarding individual student performance during the school year.

Although WyCAS reported at the federal level to illustrate whole school performance, we understand that WyCAS also generated individual student reports. These reports were sent to parents to summarize how well students did on their tests. This reporting component could have been modified in order to provide additional disaggregated student data. The embedded tools and

assessments recommended by the Task Force however, were not required by NCLB.

Item Development

According to the Task Force's Recommendation #4, the new assessment had to include, in relative terms, more multiple choice and short, open-ended questions, whereas the embedded assessments should include more extended response and performance-based tasks.

Although WyCAS used a minimum number of "items" for each subject area to illustrate whole school performance, additional items could have been developed and tested to ensure that student scores were comparable across all subject areas. As stated previously, WyCAS items were designed specifically to measure student proficiency at the school level, as opposed to subject proficiency at the student level.

Additional Grades and Subject Area

According to the Task Force's Recommendation #3, the new assessment had to measure student performance and progress against (and be aligned with) state content and performance standards in grades 3-8, and at least once in grades 10-12, in the subjects of language arts and math, effective in the 2005-06 school year (and at least once in grades 3-5, 6-9 and 10-12 in the subject of science, effective in the 2007-08 school year).

Although WyCAS was designed to test at the 4th, 8th, and 11th grades, it could have been modified to add 3rd, 5th, 6th, and 7th grades. In addition, it could have been modified to add the additional subject area (Science) for grades 4th, 8th, and 11th.

Matrix Design

Although WyCAS was designed to administer different tests under a common block of items to the student population, it could have been modified to meet NCLB. The options would have been to scrap the "matrix design", or work within NCLB guidelines for assessment built on a "matrix design."

For example, according to the Department's *A User's Guide to Preparing Submissions for the NCLB Standards and Assessments Peer Review*, states are allowed to use a matrix design, as long as the multiple forms within a content area and grade are 1) aligned with the content and achievement standards, 2) equivalent in terms of content coverage, difficulty and quality, and 3) yield comparable results.

1998 Wyoming Content Standards

Finally, although WyCAS was aligned to the 1998 Wyoming Content Standards, it could have been modified to align to Wyoming's newly developed standards in 2006. It would have been no different than what WDE encountered when studying alignment for PAWS.

As seen from the above, there is consensus from current and previous WDE employees and experts in the field, that WyCAS could have been modified to meet NCLB requirements. Saying that however, because of the direction that Wyoming chose with respect to developing an integrated statewide assessment, WyCAS may not have been an appropriate choice.

As stated above, Wyoming chose to make its new assessment instructionally supportive, as well as including a formative assessment component, two things that WyCAS did not have. Therefore, it may be academic to discuss in great depth how WyCAS could have been modified, as opposed to discussing the notion of spending a tremendous amount of money to make PAWS instructionally supportive and to include a formative assessment component; two things not required by NCLB.

Finding 3: Wyoming does not require that PAWS scores be used as high-stakes criteria related to graduation requirements, college entrance requirements, college scholarships, or pay for performance of teachers or school administrators.

PAWS scores are not being used in the state as high-stakes criteria. More specifically, districts are not required to use PAWS scores as part of their “body-of-evidence” discussed in W.S. 21-2-304 (a) (iv) and WDE Rules: Ch. 31 (Graduation Requirements).

As a result, districts are missing out on an opportunity to use an established summative assessment process as part of their graduation requirements, as well as providing incentive for students to take the test itself. Also, because PAWS is not being used in other areas such as criteria for scholarships, pay for performance, college entrance considerations, etc., it may not be viewed seriously by students who take the exam or by teachers who administer the exam.

District Body of Evidence

W.S. 21-2-304 (a) (iv) and WDE Chapter 31 (Graduation Requirements) rules require districts to submit graduation requirements for peer review approval. More specifically, they require the State Board of Education to work with districts to “*establish requirements for students to earn a high school diploma as measured by each district’s body of evidence assessment system prescribed by rule and regulation of the state board and required under W.S. 21-3-110 (a) (xxiv).*” It should also be noted that Boards of Trustees are required to include a body of evidence as a component of their district assessment system.

WDE Chapter 31 Rules-Graduation Requirements, Section 10 (d) requires district Body of Evidence to undergo a peer review process. The peer review group is comprised of Wyoming

educators “*who have successfully completed peer review training conducted by WDE.*” This group makes recommendations to the Superintendent of Public Instruction.

The rules also state that evidence shall include the following components: **1)** district assessment plans; **2)** evidence of alignment among standards, curriculum, and assessments; **3)** sample assessments; **4)** evidence of consistency, documentation of the standard setting methods; **5)** evidence supporting the fairness of the assessment system; **6)** documentation supporting the comparability of the assessment system across schools and years; and **7)** other documentation that the district chooses to submit to support the technical quality of the body of evidence assessment system.

WDE does not collect data which indicates the number of districts using PAWS data as criteria for proficiency on state standards in their Body of Evidence (BOE)

According to WDE officials, it does not collect data to illustrate how often districts use PAWS data as criteria in their Body of Evidence. The *WDE-663 Student Proficiency On Body of Evidence*, which districts use to report information regarding student proficiency, relates only to proficiency in each of the nine content areas (e.g. career vocational, fine performing arts, foreign language, health, language arts, math, physical education, science, and social studies). Therefore, the WDE-633 does not collect data on whether districts use PAWS scores as Body of Evidence (BOE) criteria.

However, LSO reviewed the Body of Evidence Folders for all 48 districts, because it is possible to search through the manual folders to determine which districts use PAWS data as Body of Evidence criteria. It is important to note however, according to WDE officials, that districts are required only to submit two of the nine content areas for peer review.

WDE also requires that of those two submitted, only one is required to be from one of the four core content areas (mathematics, social studies, language arts, or science) that are tested as part of PAWS. So, if a district chooses for example career vocational, fine performing arts, foreign language, health, social studies or history, or physical education as its core content area for peer review, PAWS data would not be available for that content area.

In our review of district BOE folders, we found that all districts submitted at least one of the four core content areas as part of their BOE. We also found that 38 districts or 79% allowed the use of PAWS scores as some type of criteria to help determine proficiency. The following table provides additional information.

Table 2.6
District Body of Evidence (BOE) Content Information

District	BOE Includes Two (2) of None (9) Content Areas		BOE Includes One (1) of Two (2) PAWS Content Areas		PAWS Scores Used for BOE Criteria	
	Yes	No	Yes	No	Yes	No
Albany #1	X		X		X	
Big Horn #1	X		X		X	
Big Horn #2	X		X		X	
Big Horn #3	X		X		X	
Big Horn #4	X		X		X	
Campbell #1	X		X		X	
Carbon #1	X		X		X	
Carbon #2	X		X		X	
Converse #1	X		X		X	
Converse #2	X		X		X	
Crook #1	X		X		X	
Fremont #1	X		X		X	
Fremont #2	X		X		X	
Fremont #6	X		X		X	
Fremont #14	X		X			X
Fremont #21	X		X		X	
Fremont #24	X		X			X
Fremont #25	X		X		X	
Fremont #38	X		X			X
Goshen #1	X		X		X	
Hot Springs #1	X		X		X	
Johnson #1	X		X		X	
Laramie #1	X		X			X
Laramie #2	X		X		X	
Lincoln #1	X		X			X
Lincoln #2	X		X		X	
Natrona #1	X		X		X	
Niobrara #1	X		X		X	
Park #1	X		X		X	
Park #6	X		X		X	

District	BOE Includes Two (2) of None (9) Content Areas		BOE Includes One (1) of Two (2) PAWS Content Areas		PAWS Scores Used for BOE Criteria	
	Yes	No	Yes	No	Yes	No
Park #16	X		X		X	
Platte #1	X		X		X	
Platte #2	X		X		X	
Sheridan #1	X		X		X	
Sheridan #2	X		X			X
Sheridan #3	X		X		X	
Sublette #1	X		X		X	
Sublette #9	X		X		X	
Sweetwater #1	X		X			X
Sweetwater #2	X		X		X	
Teton #1	X		X		X	
Uinta #1	X		X			X
Uinta #4	X		X		X	
Uinta #6	X		X		X	
Washakie #1	X		X			X
Washakie #2	X		X		X	
Weston #1	X		X			X
Weston #7	X		X		X	
Total	48		48		38	10

Source: Legislative Service Office from information provided by WDE.

Body of Evidence Studied by the State Board of Education The following table includes verbiage of WDE officials discussing a State Board of Education request to study the BOE. This study was proposed because of district concerns with the BOE requirements and WDE peer review process.

The proposal was presented to the State Board of Education on September 21, 2010.

**Table 2.7
Body of Evidence Proposal to State Board of Education
September 21, 2010**

State Board of Education Proposal Body of Evidence (BOE) Study Committee
At the request of the Wyoming State Board of Education (SBE), a Body of Evidence (BOE) Study Committee was established to review and make a recommendation regarding the current legislation, rules and regulations, and Wyoming Department of Education (WDE) policies that determine graduation using the Body of Evidence System.
This request was the direct result of concerns presented to the SBE at the November 18, 2009

**State Board of Education Proposal
Body of Evidence (BOE) Study Committee**

SBE work session meeting in Casper, Wyoming. Following the meeting in Casper, WDE enlisted representation from stakeholders statewide including WDE staff, superintendents, curriculum directors, high school principals, state board members, local board members, teachers, and citizens. A charter for the committee was agreed upon with the following goal: to review the Wyoming accountability system (BOE) in state statute and rules and regulations and provide consultative recommendations to the State Superintendent of Public Instruction and the Wyoming State Board of Education for improving the BOE system.

The timeline for this committee to complete the review and provide a recommendation began with a meeting in February 2010 and continued with a meeting once each month in order to provide recommendations in August 2010. The completed activities of the committee include a review of the history of the BOE accountability system (see appendix A), a review of the current statutory and rule and regulation guidance, a review of the BOE technical requirements, research other states' accountability systems, and develop recommendations for improving the BOE accountability system. The BOE study committee conducted two surveys on perceptions about the Body of Evidence system: one that included statewide stakeholders, and a second that included only superintendents of schools. The general perception is that stakeholders see a need for a change or updating of the current system.

BOE is a district assessment system used to determine student proficiency on the state Standards in order for districts to certify that students are ready to graduate from high school. During the ten years of using and implementing the current BOE system, the complexities of the system have placed substantial burdens upon many districts. The past four years of WDE focus on assisting districts in updating their respective systems has created a greater awareness of the complexities of this system, particularly the five design criteria of alignment, consistency, fairness, standard setting and comparability necessary in each individual district BOE system.

The consensus of the BOE Study Committee regarding these design criteria is:

- Alignment – is necessary and has been attainable in large part by all districts
- Consistency – reliability of an assessment as well as the scoring of the assessment is very difficult for districts of any size to accomplish
- Fairness – the capacity of a district to ensure fairness and/or even identifying bias is difficult under the best of circumstances
- Standard Setting – many districts have difficulty in setting cut-scores, again due to the lack of capacity and expertise to provide appropriate time and training for staff
- Comparability – some districts may be able to establish comparability between classes, however across years and especially across districts, this has not been achieved

RECOMMENDATIONS

To this end, the SBE BOE State Study Committee recommends that:

The Wyoming State Board of Education should propose the following to the Wyoming State Legislature:

**State Board of Education Proposal
Body of Evidence (BOE) Study Committee**

Amend Wyoming Statute 21-2-304(a) (iv) which requires a district BOE system to determine graduation to include a state-developed statewide graduation assessment system to measure student performance in regard to the state content and performance standards.

Source: Legislative Service Office from information provided by WDE officials.

Criteria for scholarships and pay for performance According to WDE current and past employees, as well as others with education expertise, use of PAWS could be considered high-stakes (i.e. meaningful purpose), since it is instructionally supportive. There was also a general notion among those interviewed that the use of one standardized or summative test as a graduation exit examination is somewhat outmoded.

However, if scores from an assessment such as PAWS are used as a criterion for graduation, school accreditation, and scholarships amongst others, this would not necessarily be a negative thing. As seen from the above table, it does appear that PAWS is used by a majority of Wyoming's school districts as criteria for determining student proficiency as part of graduation.

However, we also received feedback that using a statewide assessment such as PAWS for making salary decisions for teachers and administrators may not be entirely fair. For example, if teachers are subject to assessment scores as criteria for their pay decisions only because they teach subjects or grades where PAWS is administered, this could be unfair if there is nothing comparable for their peers who teach grades or subjects where PAWS is not administered.

Legislation filed in 2009 to make PAWS scores a requirement for the Hathaway Scholarship As stated above, as long as PAWS scores would be used as one selected criterion for some high-stakes purposes, there was not disagreement among those we interviewed. This would also include the use of PAWS scores as criteria for scholarships, such as the Hathaway Scholarship. Such use of PAWS scores could increase the awareness of high school students related to the importance of statewide assessments.

During the 2009 General Session, SF 139 was introduced and referred to Committee on January 28, 2009. Partly, it proposed that successful students for the Hathaway Scholarship “*demonstrate proficiency at the eleventh grade level in reading,*

writing, and mathematics as measured by the statewide assessment implemented under W.S. 21-2-304 (a) (v).” Although it died in Committee, there were positive reactions by the WDE.

According to the fiscal note, WDE “*believed that students will be motivated by the requirements of this bill and will work to achieve proficiency in the statewide tests, thus maintaining their Hathaway eligibility.*” This notion was not different from positive comments from other people we interviewed as part of this audit.

*School Finance
Recalibration
Committee considers
using PAWS for
accountability
purposes*

During the 2010 Wyoming Budget Session, HB 1, Section 334 created the School Finance Recalibration Committee (Committee). As part of the Committee’s overall charge to study recalibration of the education resource block grant model, it is also required to “*review methods of assessment of student achievement under the current model and recommend student assessment policies and procedures in the recalibrated system of education.*”

Although the Committee did not specifically study PAWS, it identified it as an assessment that can be used to measure how well districts are doing as part of the statewide accountability system it is studying. With the caveat however, that it may need to develop a subcommittee to review PAWS. This of course could include adjustments, etc., with respect to pending ESEA reauthorization.

Perhaps this is a productive occurrence, given that PAWS scores are not used for other high-stake purposes in the state. Using the statewide assessment, especially given associated costs, to measure how well districts are doing as part of a statewide accountability system, appears an appropriate use of the student test scores. Saying that however, it is still an unusual predicament that Wyoming finds itself, on the one hand identifying problems with PAWS and its administration, but also using it to measure district performance.

Recommendation: Pending reauthorization of ESEA, WDE should work with the Wyoming Board of Education to determine if the statewide assessment and related reports should be used for other high-stakes purposes such as graduation requirements, college scholarship criteria, or pay-for-performance for teachers and administrators. Based on the conclusions, criteria should then be factored into the next RFP.

An argument could be made that because PAWS is not used for meaningful purposes other than to measure AYP, students may not take the test seriously. Also, an argument could be made that to expend so much money on an integrated statewide assessment that results of the statewide assessment should be used for other purposes. Otherwise, it may be a better direction for the state to do the minimum, upon reauthorization of ESEA. In other words, modify the current statewide assessment to meet the minimum requirements for the reauthorized ESEA.

CHAPTER 3

Districts are Critical of PAWS, but also Appear to Speak from Frustration

Finding 1: Local districts are critical of PAWS, but may be lashing out because of test administration problems.

We developed and sent a 12-page survey instrument to all district superintendents, as part of our audit. More specifically, we sent 48 surveys, with a response rate of 98% (47 surveys). We did not receive a response from Fremont #38.

The survey included 19 questions, with an opportunity for comments for each question asked. All respondents included written comments, the majority of whom expressed frustrations with PAWS. However, it was also clear by the respondents' comments, that some did not fully understand requirements related to No Child Left Behind (NCLB). More specifically, this includes comments with respect to the use of summative assessments to measure Adequate Yearly Progress (AYP).

The following information illustrates the results of each yes and no question, which is then followed by a summary of comments. However, see Appendix A for more detailed respondent comments and concerns expressed by all responding school districts.

Illustration 3.1

District Superintendent Survey Instrument Yes/No Question Section

Question 2a	What is your opinion of PAWS as a valid and reliable instrument to meet the mandates of No Child Left Behind (NCLB)?			
	Valid	Somewhat Valid	Not Valid	Total Survey Responses
2a	21.28% (10)	55.32% (26)	23.40% (11)	100%

Question 2b	What is your opinion of PAWS as a valid and reliable instrument to meet the mandates of No Child Left Behind (NCLB)?				
	Reliable	Somewhat Reliable	Not Reliable	No Response	Total Survey Responses
2b	19.15% (9)	40.43% (19)	36.17% (17)	4.26% (2)	100%

Question 3	How satisfied are you with the process used to make recommendations for developing a new statewide assessment (PAWS) after the passage of No Child Left Behind (NCLB)? More specifically, the process used by the Wyoming Statewide Task Force on Student Assessment and Educational Accountability and the Commission on Instructionally Supportive Assessment, which ultimately served as the basis for an initial RFP?				
	Satisfied	Somewhat Satisfied	Not Satisfied	No Response	Total Survey Responses
3	10.64% (5)	48.94% (23)	31.91% (15)	8.51% (4)	100%

Question 4a	Are you familiar with the process used by NCS Pearson, Inc., to administer PAWS each year?			
	Yes	No	No Response	Total Survey Responses
4a	89.36% (42)	10.64% (5)	0% (0)	100%

Question 4b	If yes, how satisfied are you with the PAWS administration process from 2005 to present?				
	Satisfied	Somewhat Satisfied	Not Satisfied	No Response	Total Survey Responses
4b	2.13% (1)	42.55% (20)	40.43% (19)	14.89% (7)	100%

Question 5a	Are you familiar with the process used by NCS Pearson, Inc. to score PAWS each year?			
	Yes	No	No Response	Total Survey Responses
5a	72.34% (34)	23.40% (11)	4.26% (2)	100%

Question 5b	If yes, how satisfied are you with the PAWS scoring process from 2005 to present?				
	Satisfied	Somewhat Satisfied	Not Satisfied	No Response	Total Survey Responses
5b	6.38% (3)	34.04% (16)	40.43% (19)	19.15% (9)	100%

Question 6a	In your opinion, how do students in your district perceive taking the PAWS assessment? 4 th			
	Seriously	Somewhat Seriously	Not Seriously	Total Survey Responses
6a	65.96% (31)	31.91% (15)	2.13% (1)	100%

Question 6b	In your opinion, how do students in your district perceive taking the PAWS assessment? 8 th			
	Seriously	Somewhat Seriously	Not Seriously	Total Survey Responses
6b	36.17% (17)	44.68% (21)	19.15% (9)	100%

Question 6c	In your opinion, how do students in your district perceive taking the PAWS assessment? 11 th			
	Seriously	Somewhat Seriously	Not Seriously	Total Survey Responses
6c	14.89% (7)	61.70% (29)	23.40% (11)	100%

Question 7a	In your opinion, should PAWS, as the statewide assessment to meet NCLB requirements, also be used for other purposes as follows? : • Exit Exam for Graduation			
	Yes	No	No Response	Total Survey Responses
7a	6.38% (3)	93.62% (44)	0% (0)	100%

Question 7b	In your opinion, should PAWS, as the statewide assessment to meet NCLB requirements, also be used for other purposes as follows? : • Criteria for Hathaway or other scholarships			
	Yes	No	No Response	Total Survey Responses
7b	29.79% (14)	68.09% (32)	2.13% (1)	100%

Question 7c	In your opinion, should PAWS, as the statewide assessment to meet NCLB requirements, also be used for other purposes as follows? : • Criteria for making salary decisions for teachers, principles, and superintendents			
	Yes	No	No Response	Total Survey Responses
7c	12.77% (6)	87.23% (41)	0% (0)	100%

Question 7d	In your opinion, should PAWS, as the statewide assessment to meet NCLB requirements, also be used for other purposes as follows? : • Criteria for making funding decisions and other decisions at the district and school levels			
	Yes	No	No Response	Total Survey Responses
7d	21.28% (10)	76.60% (36)	2.13% (1)	100%

Question 7e	In your opinion, should PAWS, as the statewide assessment to meet NCLB requirements, also be used for other purposes as follows? : • Criteria for determining Wyoming's "Body of Evidence"			
	Yes	No	No Response	Total Survey Responses
7e	27.66% (13)	70.21% (33)	2.13% (1)	100%

Question 9	Does your school district use PAWS scores as part of its "Body of Evidence" with respect to graduation and mastery of the "Wyoming Content Standards"?			
	Yes	No	No Response	Total Survey Responses
9	40.43% (19)	57.45% (27)	2.13% (1)	100%

Question 10a	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your "yes" and "no" answers.: • Northwest Evaluation Association MAP (Measures of Academic Progress)			
	Yes	No	No Response	Total Survey Responses
10a	61.70% (29)	29.79% (14)	8.51% (4)	100%

Question 10b	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your "yes" and "no" answers.: • AIMESweb			
	Yes	No	No Response	Total Survey Responses
10b	8.51% (4)	70.21% (33)	21.28% (10)	100%

Question 10c	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your "yes" and "no" answers.: • ACT for 11th graders			
	Yes	No	No Response	Total Survey Responses
10c	36.17% (17)	59.57% (28)	4.26% (2)	100%

Question 10d	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your "yes" and "no" answers.: • DIBELS (Dynamic Indicators of Basic Early Literacy Skills)			
	Yes	No	No Response	Total Survey Responses
10d	23.40% (11)	76.60% (36)	0% (0)	100%

Question 10e	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your “yes” and “no” answers.: • NAEP (National Assessment of Educational Progress)			
	Yes	No	No Response	Total Survey Responses
10e	27.66% (13)	68.09% (32)	4.26% (2)	100%

Question 12	In your opinion, should Harcourt Assessments, Inc., have been monetarily penalized for problems in 2006 and 2007 for problems with reporting and scoring?			
	Yes	No	No Response	Total Survey Responses
12	63.83% (30)	12.77% (6)	23.40% (11)	100%

Summary of Comments Related to the Above Questions

General opinion of PAWS

General consensus from the districts surveyed express concern with PAWS in several areas. The test is perceived to be too long, and scores are not given to educators in a timely manner. Many responses in the survey indicate that educators want to use test results to help place students, but because results are sometimes not received until three months into the semester PAWS has limited use in guiding classroom instruction. There are also concerns with the ability of the vendor to properly administer the test because of problems experienced in previous years related to technology and scoring.

It also appears that educators have a general concern with the writing portion of PAWS because of the time that it takes a student to complete. They are also concerned with the subjective nature of the scoring. The writing portion of PAWS is not timed. The WDE has recommended time frames for completion, however the time students can take for completion can far exceed that which is necessary and reasonable. Related to scoring, subjectivity plays a large role and educators feel there is little guidance to help prepare students for this portion, which adds to the confusion and frustration on the part of all stakeholders.

Reliability and validity with respect to NCLB

There is a large misconception about PAWS on the part of districts where validity and reliability of PAWS is associated with NCS Pearson’s administration of the test. Validity and reliability for a testing instrument are necessary for a state’s assessment to meet requirements set forth by NCLB. Validity refers to a measure to the extent which a concept, conclusion or

measurement is well-founded and corresponds accurately to the real world. Reliability is defined as the consistency of a set of measurements or of a measuring instrument. Both concepts are necessary for a well grounded assessment.

Confusion exists within the definition of these concepts where PAWS is viewed as less valid and reliable. Seventy eight-percent (78%) or 36 district superintendents deem PAWS somewhat valid or not valid and 76% or 35 deem the test somewhat reliable or not reliable. This is in contrast to psychometric data which forms the basis of recommendations by the PAWS Technical Advisory Committee, which deem PAWS to demonstrate the highest level of validity and reliability. PAWS has also passed the U.S. Department of Education (Department) peer review, which assesses validity and reliability.

In addition, confusion exists within the realm of PAWS as a valid and reliable indicator of student achievement because it does not include a growth model. Districts believe a growth model component is needed for validation to show individual student progress. However, this component was not a consideration in constructing PAWS or for meeting AYP at the present time. It may be necessary in the future to add this component to the state assessment, however doing so before reauthorization of ESEA may be presumptive. In other words, district officials are confusing expectations with respect to ESEA reauthorization, with criticism of PAWS.

Those who use the traffic signal process speak favorably

Limited input from districts about the traffic signal process indicates it is used for instruction decisions even though it is sometimes received late in the school year. The traffic signal process is meant to be an instructionally supportive element of PAWS as an indicator of student achievement/proficiency level.

The purpose is to place students and document their proficiency level as they move through each grade level. It is constructed based on students' scores and details areas where a student does not need additional instruction (green-light), may need additional instruction (yellow-light), or more than likely needs additional instruction (red-light). The light system is based on a percentage of correct answers on PAWS where scores for the various levels (cut scores) are determined by a panel/committee of statewide

educators.

Five respondents commented that they do use the traffic signal process, meaning they use the reports and instructional suggestions. However, there were also concerns that the information was not made available to districts in a timely manner.

PAWS may not be taken seriously, but change is approached cautiously

There appears to be a major public relations problem with PAWS in terms of acceptability on the part of district administrators, teachers, and students. The problems with accepting PAWS relate to educators seeing the exam as a federal requirement hurdle that takes away from classroom instruction. This perception permeates downward to students who actually take the exam. This is especially ironic, given that PAWS was designed to be instructionally supportive.

There appears to be a downward trend in students who take the exam seriously as age of the student increases. This mirrors district administrator sentiments that as age increases students are increasingly distracted by other obligations and tend to disregard the test starting in 8th grade. An example of this trend is that students see more value in taking the ACT because it has implications for their future as opposed to PAWS which has little consequence based on the outcome. In other words when prioritizing tests, PAWS falls to the bottom.

Responses to our survey also indicate that districts are well informed about the process used by NCS Pearson, Inc. to administer and score the PAWS exams. There is an equally high dissatisfaction rate with the process used by NCS Pearson. This is mainly attributable to the platform problems of the previous year and problems associated with scoring going back several years. This may be why 63% or 29 are in favor of monetarily penalizing NCS Pearson, Inc. for the company's difficulty in properly administering PAWS in the previous years.

Finally, districts seem to want progression to a new assessment system, yet they are cautious of immediately changing. These feelings are based on the uncertainty surrounding reauthorization of ESEA at the federal level and the changes that might occur if Wyoming commits to a consortium. To this end there are also

concerns related to losing the longitudinal data from another exam. The feeling seems to ruminate that educators will have more unusable student data going forward, which does not sit well with districts.

Finding 2: There is a desire among some districts to use formative assessments in lieu of PAWS. However, such use would be contrary to NCLB requirements.

Misunderstanding with respect to using other types of formative assessments in place of PAWS

For the most part districts recognize that PAWS should be the only test used to meet NCLB requirements. There is a smaller percentage ranging from 10% to 30% that support the notion of using different types of formative assessments in lieu of PAWS. However, other types of suggested assessments are contrary to NCLB and would not pass the federal peer review process; primarily because they are not aligned to Wyoming state standards.

In addition, NCLB requires that a state assessment be summative in nature rather than formative. The distinction is that a formative assessment rates a student by his or her current abilities in a given subject at a point in time, whereas a summative test is comprehensive in nature showing results of knowledge gained by a student over a given time period.

PAWS originally was designed with an embedded formative assessment component

As discussed in Chapter 2, PAWS was originally intended to have summative and formative components that would provide districts with the ability to monitor progress year round. However, both the Stanford Learning First and Benchmark-Skillwriter have been dropped from PAWS because of non-use. This leaves a gap for educators to fill to be able to assess students quickly and efficiently throughout the year.

We also discussed earlier that all school districts use NWEA-MAP as a formative assessment, which helps support and guide instruction for individual students and in the classroom at the district level. It should be noted that all districts participating in the Bridges Summer School Program are required to assess all

students from the Spring of 2nd grade through the Fall of 9th grade (Spring and Fall semesters) using MAP (e.g. Board of Education Rules-Chapter 33, Section 7 (e)).

Although MAP fits the needs expressed by districts to access and monitor students throughout the year as needed, it should not be mistakenly applied towards meeting the requirements of NCLB.

PAWS has received approval from the federal Department of Education

Through our research on the criteria needed to meet federal requirements, PAWS is the only test Wyoming has available to comply with NCLB. To meet federal requirements a test must be first and foremost aligned to state standards. It must also test grades 3rd through 8th and 11th and be a summative criterion referenced assessment with the intended purpose of measuring student achievement in areas relevant to specific grades. Subjects tested must include math, language arts, and science where student level data is recorded.

Finally, we found that federal peer reviewers interpret NCLB as requiring that all students receive identical assessments; formative assessments can be different for various student groups. Other assessments suggested for use by various district superintendents, do not meet these criteria, as described below.

Description and Analysis of Formative Assessments re: NCLB

American College Testing Program (ACT)

The ACT Plan, and Explore do not meet federal requirements because they are not aligned to state standards. The ACT is an IQ test intended to predict how a student will perform in college. With the ACT, each student receives a different test. The test can be given to 8th through 11th graders, but has no comparable test that is aligned to students in previous grades.

The proper subjects are tested to meet federal requirements; however the test has elements of a criterion referenced test, but is also norm referenced. The combination of the factors mentioned above make this test unsuitable to meet federal requirements.

Northwest Evaluation Association (NWEA)-

The MAP test could not be adapted to meet federal requirements because it is not aligned to state standards and because it is formative, not summative in nature. The MAP test is intended

Measures of Academic Progress (MAP) to test on a learning continuum which will not indicate if a student has mastery of a subject. In addition this test is designed for students in K-6th with no comparable test for 7th, 8th, and 11th grades. It does test the proper subjects, however it is a computer adaptive test which implies all students will not receive the same test.

In a July 15, 2010 correspondence from Ms. Ginger Hopkins, Vice President of Partner Relations, NWEA to Wyoming school districts, she discusses the positive use of MAP throughout Wyoming. However, she discusses this within the framework of a formative assessment.

“Throughout our nearly 20 years of service in Wyoming, partners using NWEA assessments have understood the value of MAP as a stable component in a comprehensive assessment system. MAP, as delivered to our district partners, is suited for many purposes, including interim or benchmark testing and reporting; it is not permitted for summative testing to satisfy NCLB accountability measures, another component of a comprehensive assessment system in today’s education climate.

While NWEA is working with state education departments for a variety of purposes, and we have the capacity to deliver assessments to meet summative requirements, such assessments are distinctively different from standard MAP. NWEA recognizes that NCLB summative assessments have specific blueprint and design requirements, so care should be taken to ensure the right assessments are used for the desired purposes.”

As can be seen from the above correspondence, Ms Hopkins refutes the use of a formative assessment, albeit a highly valued assessment in Wyoming, for NCLB purposes.

Other assessments The National Assessment of Academic Progress (NAEP) is the only nationally representative assessment of how students test against various subject areas. States receiving Title I funds are required to participate in the program.

Although NAEP is a test that is given across the nation, it is not aligned to any specific state standards. The test is only given to 4th and 8th graders, where a sample of the student population is chosen to take the test. Each student receives a different test in the subjects of math and reading, but not science. The test is aggregated only at the school level.

AIMSweb and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) are formative tests that are not aligned to state standards and are only available for grades K-8th and K-6th respectively. AIMSweb covers math, reading and written expression and DIBELS covers reading and written expression. Neither test has a science portion. The test can be web based, paper and pencil, or oral; the latter being problematic for passing the peer review because each student could potentially receive a different test.

The following table provides more detailed information on the assessments discussed above

Table 3.1
Assessment Information

Criteria	PAWS	NAEP	*ACT	MAP	AIMSweb	DIBELS
Aligned w/standards	Yes	No	No	No	No	No
Level of Data	Student	State	Student	Student	Student	Student
Grades Tested	3 rd -8 th & 11 th	4 th and 8 th	8 th – 12 th	K-6 th	K-8 th	K-6 th
Broad Type	Summative	Summative	Summative	Formative	Formative	Formative
Detail Type	CRT	CRT	NRT	Learning Continuum by Subject	Curriculum Based Measurement	**CRT
Subjects Tested	Math, Reading, Writing, and Science	Reading, Math, Science, and Writing	English, Math, Reading, and Science	Reading, Language Usage, Math, and Science	Reading, math, and Written Expression	Reading and Written Expression
All Students Receive Same Test	Yes	No	No	No	No	No
Delivery	Paper &	Paper &	Paper &	Computer	Web Based	Web Based

Criteria	PAWS	NAEP	*ACT	MAP	AIMSweb	DIBELS
	Pencil	Pencil	Pencil	(adaptive test)	Testing, Oral Testing, and Paper & Pencil	Testing, Oral Testing, and Paper & Pencil

Source: Legislative Service Office from Internet and other research.

*Includes Plan (9th and 10th grades) and Explore (8th grade). Currently, WDE is working with ACT, Inc. to “operationalize a Statewide 8th Grade EXPLORE test and Senior WorkKeys Pilot testing program, which is an extension of the testing program as mandated by the Legislature.”

**DIBELS are criterion-referenced because each measure has an empirically established goal (or benchmark) that changes across time to ensure students' skills are developing in a manner predictive of continued progress.

Recommendation: The WDE should develop and administer training on NCLB requirements, as well as pending ESEA reauthorization with respect to differences between various assessments as they apply to NCLB requirements. This should be applied to additional training on the PAWS program from RFP to contract.

WDE has a tremendous opportunity to educate districts with respect to current NCLB requirements and pending ESEA reauthorization at what appears to be a pivotal time. Because of movement with respect to the current administration’s blueprint, as well as interest in Race to the Top funding by various states and consortia, WDE can work closely with districts so they understand the current condition, as well as pending changes. WDE and the districts should work closely together after ESEA is reauthorized, to determine what type of changes should be made to the current statewide assessment. In addition, districts should fully understand the PAWS program (and process) from RFP to contract.

CHAPTER 4

School Year 2009-2010 PAWS Administration Failed the Citizens of Wyoming

Finding 1: NCS Pearson, Inc., encountered problems and performed poorly during the 2009-2010 school year as it carried out contractual obligations

There is no doubt that NCS Pearson, Inc., encountered serious problems and performed poorly with respect to the administration of PAWS during the 2009-2010 school year. An apology from NCS Pearson, Inc., to Wyoming residents in July 2010, as part of ongoing settlement negotiations, was disturbingly similar to earlier apologies to Florida and Minnesota for similar problems. In Wyoming's apology, the company stated it was "*dismayed to have let you down,*" but promised to "*take additional steps to regain your confidence.*"

As a result of the problems experienced during its last PAWS administration, policy makers and state officials have lost complete confidence in the vendor to perform throughout the remainder of the contract. The result on students' achievement however, is not as clear cut.

Rather than editorialize about the obvious fact that the company did not perform its duties and responsibilities in accordance with its contractual obligations, it is important to factually provide a summary of those problems. It is equally important to identify actions WDE has taken since those problems occurred, as well as other factors that might impact Wyoming into the future.

Summary of problems caused by NCS Pearson, Inc.

Generally, problems ranged from difficulties logging onto the system, delivery of late testing materials, crashes of the untested online testing platform, poor communication and lack of resolution from the vendor, and late test results. WDE, as a result, took the proper steps to stop payments to the vendor and

enter into negotiations for settlement. Section 8 (P) of the current contract allows for WDE to recoup liquidated and default damages.

According to correspondence dated April 26, 2010 from State Superintendent of Public Instruction Jim McBride to NCS Pearson, Inc., liquidated and default damages were initially figured at \$9,562,803. In the correspondence, the Superintendent formally imposes upon the vendor to pay that amount, as well as overtime worked by district teachers and administrators to resolve administration problems because of a “*complete default of the contract*,” which makes the test “*essentially unusable by the Department and the State of Wyoming as a whole.*”

The following tables provide a summary of alleged damages. For more specific information on alleged liquidated and alleged default damages, see Appendix C.

Table 4.1
Alleged Liquidated Damages (Test Materials and Administration)

Description of Events	Liquidated Damage
The <i>PAWS 2010 Directions for Administration</i> manual had to be split into two volumes, Paper/Pencil and Online due to the incompleteness of the online platform functionality. The <i>PAWS 2010 Online Directions for Administration</i> manual was posted as a PDF on Feb. 23, 2010. Printed copies were not delivered to schools until February 26, 2010. In addition, the Header Sheets, which provide pre-printed school information and used when shipping test materials back to Pearson, had to be re-printed due to school code errors. These too were shipped out the week of Feb. 23, 2010.	\$ 70,000 (35 days)
See explanation for Liquidated Damages in Table 4.1.	\$ 10,000 (5 days)

Source: Legislative Service Office from information provided by WDE.

Table 4.2
Alleged Liquidated Damages (Technology)

Description of Events	Liquidated Damage
Online platform failed to perform throughout the window when the tutorial was available; technical issues remained during the duration; problems were never completely rectified.	\$ 86,000 (43 days: Feb 15 – April 14, 2010)

Description of Events	Liquidated Damage
For purposes of calculating liquidated damages the Department utilized the tutorial and testing windows; arguably liquidated damages could continue to be imposed to date.	
<p>The Pearson TestNav 7.0 online assessment platform continued to malfunction throughout the test administration window. This is supported by Pearson's PAWS Communication #37 (April 5, 2010), "Please continue to report any technical challenges you may experience to Pearson's CSC, even though you have resolved the situation. Pearson is aware that some schools are still challenged by random incidents with the online testing platform. Pearson will notify WDE to introduce new software coding in the event we identify any specific root cause for these issues."</p> <p>Testing window had to be extended an additional three (3) days due to the technical difficulties of the on-line platform.</p> <p>For purposes of calculating liquidated damages the Department utilized the testing window; arguably liquidated damages could continue to be imposed to date.</p>	\$ 56,000 (28 days: March 8 – April 14, 2010)

Source: Legislative Service Office from information provided by WDE.

**Table 4.3
Alleged Liquidated Damages (Reporting Services)**

Description of Events	Liquidated Damage
On April 9, 2010, the WDE declined to accept the Real-Time (On-Demand), online multiple-choice individual student results due to multiple significant errors in appearance and accuracy. Satisfactory reports were delivered on April 23, 2010.	\$ 36,000 (18 days: March 31 - April 23, 2010).
Due to the deficiencies of the TestNav 7.0 Platform, analysis by an independent 3 rd party was necessary prior to processing of data and information and compilation of any reports; all reporting was delayed significantly as a result of the failure of the TestNav 7.0.	\$ 172,000 (86 days: May 31, 2010 – Sept. 27 2010).
See explanation for Liquidated Damages above.	\$ 190,000 (95 days: May 31, 2010 – October 8, 2010).
See explanation for Liquidated Damages above.	\$200,000 (100 days: May 31, 2010 – October 15, 2010).
See explanation for Liquidated Damages above.	\$ 104,000 (52 days: August 5, 2010 – Oct. 15, 2010).

Source: Legislative Service Office from information provided by WDE.

**Table 4.4
Alleged Liquidated Damages (Program Management)**

Description of Events	Liquidated Damage
Contractor failed to staff call center during the tutorial and test administration window with staff with sufficient technical expertise to assist districts and the Agency is resolving technical issues.	\$ 78,000 (39 days: Feb 19 – April 14, 2010).
The Agency was not accurately informed of status of the program and the issues with TestNav which affected many of the deliverables. Quarterly reports lacked detail and did not include call center information. Agency repeatedly asked for call center reports and the Contractor failed to provide it in a timely manner.	\$ 128,000 (64 days: Dec. 1 – Feb. 28). \$ 132,000 (66 days: Mar. 1 – June 1, 2010) \$ 14,000 (7 days: June 1 – June 9, 2010) \$ 50,000 (25 days: Sept. 1 – October 5, 2010)
The Agency was not provided with the quarterly invoice in a timely manner; when quarterly reports were submitted, lacked critical information.	\$ 16,000 (8 days: Sept. 1 – Sept. 9, 2010)

Source: Legislative Service Office from information provided by WDE.

**Table 4.5
Alleged Default Damages (Test Development and Materials)**

Description of Events	Default Damage
Acceptable and accurate test forms for the 2010 TestNav Platform were not provided by the Contractor to the Agency until February 26, 2009.	64 days: Dec. 1 2009 – Feb. 26, 2010
See explanation for Liquidated Damages #1) above.	64 days: Dec. 1 2009 – Feb. 26, 2010

Source: Legislative Service Office from information provided by WDE.

**Table 4.6
Alleged Default Damages (Events and Committees)**

Description of Events	Default Damage
<p>Pearson's Proposal included a detailed description of the creation of a Language Arts score for Adequate Yearly Progress (AYP) which in turn promised to stabilize the PAWS writing results by relying on a combination of links to multiple-choice Language Arts items, changing the writing rubrics, changing the number of score points possible and placing the PAWS reading and writing scores on a single scale. To assist with the creation of the vertical scale from grades 3 – 8, students responded to an "off-grade" writing prompts at grades 3 – 8. The contract called for the creation of approximately 100 multiple-choice Language Arts items to be reviewed, edited, field tested, scored and taken through a data review process in the 2010 assessment cycle. These tasks were completed. However, upon the failure of the online TestNav platform to function, several steps were taken to report PAWS writing scale scores as the entire test had changed. When Pearson's lead psychometrician attempted to implement the writing equating model as proposed, it was discovered to be flawed. As a result, all of the multiple-choice language arts items have been removed from any further use on a future PAWS assessment and the time and resources spent on their development now represent a loss to the Department. Finally, all student responses and subsequent hand scores of the "off-grade" writing prompts used at grades 3-8 were unusable due to the failure of the equating model.</p>	<p>N/A: acceptable standard setting for ELA cannot be completed by Contractor.</p>

Source: Legislative Service Office from information provided by WDE.

**Table 4.7
Alleged Default Damages (Data Management Service)**

Description of Events	Default Damage
<p>See explanation for Liquidated Damages in Table 4.6 above.</p>	<p>92 days: May 31, 2010 – October 5, 2010</p>
<p>See explanation for Liquidated Damages in Table 4.6 above.</p>	<p>49 days: August 5, 2010 – October 5, 2010</p>

Source: Legislative Service Office from information provided by WDE.

**Table 4.8
Alleged Default Damages (Psychometric Services)**

Description of Events	Default Damage
Due to the deficiencies of the TestNav Platform, Technical Manuals could not be prepared.	20 days: Sept. 15, 2010 – October 5, 2010

Source: Legislative Service Office from information provided by WDE.

**Table 4.9
Alleged Default Damages (Program Management)**

Description of Events	Default Damage
On-line portion of Spring 2010 PAWS Administration was not adequately staffed; indicated 6 full-time staff employed by Pearson for <i>all</i> on-line assessments nationwide contracted by Pearson; insufficient staffing as evidenced by TestNav 7.0 failure. Customer Service Center was inadequately staffed as well; evidenced by experience of State and District Staff.	N/A
Communication from Contractor to the State and for communication to the District routinely has numerous spelling, grammatical and typographical errors; basic proof-reading is not taking place prior to publication of communication.	N/A

Source: Legislative Service Office from information provided by WDE.

**NCS Pearson, Inc., Experienced Similar Problems
in other States; Wyoming was Aware of Problems**

NCS Pearson, Inc. has had problems in other states

There is no debate that NCS Pearson, Inc. has had similar problems in other states. What is unclear however, is the monetary impact on those states.

The following table provides information from a Washington Post article that annotates an inventory of problems experienced by the vendor.

Table 4.10
NCS Pearson, Inc. Reported Problems in Other States

Year	State	Problem
1998	California	Test score delivery was delayed.
1999-2000	Arizona	12,000 tests misgraded due to flawed answer key.
2000	Florida	Test score delivery delayed resulting in \$4 million fine.
2000	Minnesota	45,739 graduation tests misgraded. Lawsuit followed with \$11 million settlement. A judge identified quality control issues and emphasis on "profitability and cost-cutting".
2000	Washington	204,000 Washington Assessment of Student Learning writing exams rescored.
2005	Michigan	Scores delayed and fines levied per contract.
2005	Virginia	Computerized test was misgraded resulting in \$5,000 scholarships being awarded to five students.
2005-2006	Not stated	4,400 SAT college admissions test scored wrong. \$3 million settlement.
2008	South Carolina	School report cards were delayed due to scoring errors.
2008-2009	Arkansas	Actual test used as practice; first graders forced to retake the exam.
2010	Florida	Test score delivery delayed over a month resulting in the imposition of fines. Score accuracy questioned by superintendents across the state.
2010	Minnesota	Online science test results taken by 180,000 students delayed due to scoring error.

Source: Legislative Service Office from information provided in a Washington Post article entitled "*New test scoring Problems for Pearson*".

WDE was also aware of the vendor's past problems, as well as its predecessor, Harcourt Assessment and Harcourt International, which it acquired in 2007 for \$950 million. For example, from 2005 through 2008 WDE experienced various scoring and printing errors. However, at that time, for various reasons related to contract language and modifications, WDE did not consider financial penalties. Had the contracts been properly updated as changes were made, though, WDE's ability to impose penalties on Harcourt would have been a possibility.

After the acquisition however, WDE officials did not expect an adverse impact to PAWS, as expressed in various newspaper quotes.

- "*We have weathered a number of storms with regard to PAWS,*" McBride said. "*Our earliest experience with Harcourt, Inc. could even be called rocky. Through that*

process we have made modifications to the assessment. “We have improved the use of technology in delivering the test,” he said. “Many, many processes have been improved and refined at the state and local district levels.”

- *“An important development occurred when Harcourt, Inc. was purchased by Pearson Education. The new leadership at Pearson Education was able to keep the best of what Wyoming had identified in the Harcourt process and brought in very qualified assistance where needed,” McBride said. “When the time came to award the new contract, Pearson Education had put its stamp – that of an international leader in education technologies and assessment – on the Wyoming bid.”*

NCS Pearson, Inc. discloses past performance issues

As an RFP requirement in 2008, vendors were required to identify the problems or incidents they have had in the past with their assessment administrations and how they resolved the issues. NCS Pearson, Inc. identified 21 incidents related to administration of statewide assessments. With the caveat that NCS Pearson, Inc. is a very large company, the other vendors listed significantly fewer numbers of incidents: Data Recognition Corporation (DRC) reported 9 incidents; CTB/McGraw-Hill reported only one (1); and Measured Progress reported no past incidents.

Although WDE understood the past associated with Harcourt and NCS Pearson, Inc., it chose to move forward with the 2008 contract with NCS Pearson, Inc. It should be stated however, that WDE also ensured, through modifications to contractual language, that it would be able to hold the new vendor accountable if similar problems occurred; an occurrence which can be viewed as somewhat prophetic.

WDE Takes Additional Steps

WDE changes PAWS testing format

On May 17, 2010, the WDE announced changes to the 2011 PAWS administration. More specifically, it announced that the test will be administered solely through paper and pencil, and that because students will not have to compete for limited computers, less instructional time will be used. However, the amount of time it takes to take the test (e.g. eight (8) or nine hours (9)) will remain the same.

The decision to change the format resulted from feedback received from the districts, to discontinue the use of the online portion. According to the May 17, 2010 press release *“two out of every three respondents preferred to not continue with an online portion of the assessment.”* Generally, 53% of the respondents believe the 2011 PAWS administration should be conducted as a fully pencil and paper test. Thirty-seven percent (37%) disagreed.

Some respondents spoke to problems with the online portion; for example *“Let’s go to what we can control. It would also be an incredible help with scheduling in the building. Trying to work 280 kids around computer time is a nightmare in itself.”*

There were also respondents in favor of continuing the online portions; for example *“Results take too long with paper & pencil tests. I don’t want to pack up that many more boxes and have more booklets to deal with.”*

WDE applies for AYP waiver

Perhaps one of the more curious actions taken by WDE, was to formally request a waiver from the U.S. Department of Education (Department), prior to the release of a report by the National Center for the Improvement of Educational Assessment, Inc. (NCIEA), located in Dover, New Hampshire. WDE hired NCIEA to conduct a study to **1)** determine likely impact of administration problems and **2)** make recommendations relative to reporting of results: At what level(s) should reports be produced, and with what caveats?

However, before the study was completed, the WDE formally requested that the Department *“grant a waiver in using the results of the 2010 PAWS to make accountability and AYP determinations.”* The letter also outlines problems encountered during the 2010 administration of PAWS.

In a May 26, 2010 response from the Department, it is clear that Wyoming’s request will be considered, but is also clear that WDE is held primarily responsible for problems encountered with the vendor.

“As a preliminary matter, WDE continues to bear responsibility for ensuring compliance with the

standards, assessment, and accountability requirements of the ESEA. Although a state assessment contractor may serve as the WDE's agent, it is the WDE's obligation to ensure that its contractor's failure to perform will not prevent it from administering assessments and making timely AYP determinations based on current achievement data. Assurance from a contractor that a test will be administered properly are not sufficient to alleviate a State educational agency of its statutory and regulatory obligations."

The letter ends by requesting that WDE should submit a corrective action plan *"detailing the actions it will take to prevent this problem from arising in the future in the event of another malfunction of its online assessment systems,"* once a waiver is granted.

According to WDE, the AYP waiver request was separate from the request to NCIEA to study the impact that test administration potentially had on the PAWS results. Saying that however, WDE may have prematurely made its AYP waiver request prior to the actual reporting of student scores.

NCIEA report results may not warrant a waiver request

According to NCIEA's report released on August 27, 2010, technical problems with the 2010 PAWS administration had limited impact on *"its scope and student achievement."* An excerpt from the study provides additional detail.

"While scores at some grades were down, the average change was positive-even with the administration problems, students scored higher, on average, in 2010 than they had in 2009. So neither study provided evidence that the administration problems had an impact on student performance."

It should be noted however, a potential bias may be present in NCIEA's methodology. For example, it appears that NCIEA did not validate the original sample submitted by NCS Pearson, Inc. NCS Pearson, Inc. pulled the sample from its toll-free center. They determined that 400 of 489 calls that were logged onto "tickets", identified students who were affected by administration problems.

Also, it appears that NCS Pearson, Inc. conducted the statistical analysis related to the population of 400 students. For example, on page three (3) of the report, NCIEA states “*Pearson computed the deviation of these students’ scaled scores (divided by the standards deviation of student scaled scores, so the results would be reported in a standardized form) from the state average in 2009 and 2010.*” It is not clear however from the report, the type of statistical verification conducted by NCIEA, which may have occurred.

Nonetheless, NCIEA provided sound recommendations for the future. Given the pending waiver request, it may be advantageous for Wyoming to follow-up on the recommendations prior to accepting a waiver from the Department.

The following table provides additional information.

Table 4.11
NCIEA Report Recommendations

Recommendation Number	Recommendation Detail
#1	All reports that were originally planned should be produced and distributed. Without evidence to the contrary, it should be assumed that the reports provide a valid estimate of student achievement.
#2	If it is known that a student was affected by administration problems, and the achievement of the student on PAWS was inconsistent with other information about the student, the PAWS result likely should be discarded. Note, however, that this recommendation is consistent with all good testing practice; any time an individual test results is not consistent with other known information about a student’s achievement level, the other information should take higher priority in judging the student.
#3	WDE and Pearson should make an offer to any district that feels it can identify subgroups that were clearly affected and clearly not affected to conduct the kind of impact study we were unable to do under the time constraints provided by this contract.

Source: Legislative Service Office from information in NCIEA report entitled *Report on Wyoming’s Testing Irregularities*.

Recommendation: WDE should continue seeking the maximum amount of liquated and default damages based on problems encountered by NCS Pearson, Inc. as it carried out contractual obligations.

WDE has already taken crucial steps against the current vendor with respect to collecting liquidated and default damages. Contractual language allows them to move forward in an aggressive and prudent manner to ensure the interests of Wyoming citizens and their children receive full recompense.

Recommendation: WDE should continue working with the Department with respect to its AYP waiver. It should also ensure that the recommendations from NCIEA's report are fully implemented.

Although WDE may be in somewhat of an awkward position with respect to its requested waiver, it should continue working with the Department for the best interests of Wyoming. It should also ensure that NCIEA's recommendations are fully implemented, to determine impact that the testing administration had on student scores.

Finding 2: The PAWS writing component is not required by NCLB

Even though NCLB does not require a writing portion for AYP purposes, Wyoming Statute 21-2-304(a) (v) (B) requires writing as part of the statewide assessment: "...*effective school year 2005-2006, and each school year thereafter, assessing student performance in reading, writing and mathematics at grades three (3) through eight (8) and at grade eleven (11).*" [emphasis added] As a result, the state may be expending an additional \$3.5 million per year for this requirement, which is not required by NCLB.

Few changes have been made to the writing component, but some still have concerns

The writing portion has only gone through one change since its inception. In the 2010 PAWS, the scoring guide was improved, which prioritized six traits of writing in the previous scoring guide into four clear curricular aims for instructional purposes in the classroom.

However, educators appear to have a general concern about the writing portion of PAWS because of the time it takes a student to complete, as well as the subjective nature of the scoring.

Originally, the writing portion included six traits on which to be tested, but that has been reduced to four traits. Even though there are a reduced number of traits, educators still claim that this portion of the test takes too long to administer and maintains subjectivity in the scoring.

Another complaint about the writing portion is that the prompts may not apply to all children. For example, a prompt may ask a student to write about their pets or the games they play with their families. This is problematic for a child if he or she does not own a pet or whose family does not play games together. Some educators also believe that the “drafting” portion of the test should be eliminated to reflect more “on demand” writing.

Given some concerns expressed by educators, there are still educators who believe it is important to include a writing component within the statewide assessments. Without such a component, writing may be diminished in terms of importance when compared to other subject areas. Both concerns appear to be supportive in what we found in other states, where four of the five states we reviewed, included a writing component in their statewide assessment.

Finally, given the manner in which writing is tested, there appears to be comparability of scores within a current school year. Although not comparable across multiple years for subjects like mathematics, the writing component may still be able to serve an instructional purpose. It should also be noted that according to WDE, Wyoming educators who score the writing portion of PAWS undergo regional training in order to understand the scoring process and rubrics.

*Legislation proposed
to eliminate the writing
component*

During the 2010 Budget Session, House Bill 120 was introduced to initially modify the writing component, but then to eliminate it all together. More specifically, it called for the discontinuation of the use of essay writing, as well as constructed-response or open-response items to assess writing skills. Rather, essay writing would have been replaced with a review of students’ writing portfolios drafted throughout the year; also with the possibility of multiple choice items. Ultimately however, the bill sought to drop the writing portion altogether.

WDE estimated it would cost an additional \$7 million per year to develop a portfolio system for the PAWS writing portion. If the writing portion were eliminated, WDE estimated a \$3.5 million per year savings, which would not offset the additional costs to develop a new system. In addition, WDE stated there would be an undetermined cost related to the re-design of PAWS, given NCLB requirements.

HB 120 was placed on General File on February 25, 2010, but was not considered by Committee of the Whole (CoW). According to one legislator, the general consensus was that there may not have been enough time to continue debating on the bill given its complexion, and that changes to the statewide assessment should come from within WDE.

Recommendation: WDE should continue to address concerns with respect to the writing portion of PAWS. It should also determine whether eliminating the writing portion in the future, would diminish the importance of writing within the districts. It should work with the local school districts to obtain their feedback as well on this option, as well as modifications to remove the “drafting” portion of the component.

At this point, there is legitimate concern about the writing portion of PAWS. However, there also appears to be legitimate reasons to include writing as part of Wyoming’s statewide assessment. Without specific cost figures to show the impact of modifying the writing portion and/or eliminating that portion altogether however, the discussion appears academic. There appears to be merit though, to discuss various options with districts to determine if modifications should occur in the future.

CHAPTER 5

U.S. Department of Education (Department) peer review process appears inconsistent across all states

Finding 1: The federal peer review process has frustrated WDE, and appears to be inconsistently applied across all states

According to WDE officials, the federal peer review process has been laborious, has impacted the instructionally supportive component of PAWS, and has been inconsistently applied to all states. Education experts in the field concur, and one fully blames the bureaucracy within the U.S. Department of Education (Department) for the inability of PAWS to be used to its fullest potential. As a result, Wyoming’s inability to fully utilize its integrated statewide assessment as discussed in Chapter 2 may be partially a result of stressed resources at the WDE level.

Broad federal law establishes the peer review process According to NCLB (Title I, Part A, Section 1111, (e)), the Secretary of Education is required to establish a peer review process. More specifically, it requires the Secretary to do the following:

**Table 5.1
NCLB Peer Review Requirements**

NCLB Peer Review Requirements as Directed to the Secretary of the Department of Education
(A) establish a peer-review process to assist in the review of State plans; (B) appoint individuals to the peer-review process who are representative of parents, teachers, State educational agencies, and local educational agencies, and who are familiar with educational standards, assessments, accountability, the needs of low-performing schools, and other educational needs of students; (C) approve a State plan within 120 days of its submission unless the Secretary determines that the plan does not meet the requirements of this section; (D) if the Secretary determines that the State plan does not meet the requirements of subsection (a), (b), or (c), immediately notify the State of such determination and the reasons for such determination; (E) not decline to approve a State's plan before—

NCLB Peer Review Requirements as Directed to the Secretary of the Department of Education
<p>(i) offering the State an opportunity to revise its plan;</p> <p>(ii) providing technical assistance in order to assist the State to meet the requirements of subsections (a), (b), and (c); and</p> <p>(iii) providing a hearing; and</p> <p>(F) have the authority to disapprove a State plan for not meeting the requirements of this part, but shall not have the authority to require a State, as a condition of approval of the State plan, to include in, or delete from, such plan one or more specific elements of the State's academic content standards or to use specific academic assessment instruments or items.</p>

Source: Legislative Service Office from information from the NCLB Act, 201.

In addition, NCLB allows for monetary penalties against states that fail to meet NCLB requirements. More specifically, it allows the Secretary to withhold funds for State administration until the Secretary determines states have complied with NCLB. According to the Department, 31 states have fully passed the federal peer review.

No Code of Federal Regulations (CFRs) appear to exist detailing the peer review process

Although broad federal law establishes the peer review process, there does not appear to be any specific CFR, policies, operating procedures, etc., with respect to the actual peer review process that is employed by the Department. There is however, guidance provided to states that discusses seven integrated components of the peer review process. The guidance is entitled *Standards and Assessments Peer Review Guidance: Information and Examples for Meeting Requirements of the No Child Left Behind Act of 2001*.

More specifically, the guidance sets forth criteria which are required by the Department to meet NCLB standards. The guidance lays out the regulatory requirements as well as parameters for state assessments. It also generally defines the peer review process, the roles of peer reviewers, the review teams, and states' roles in this process. More importantly, the guidance describes critical aspects that state assessments must include with respect to the seven integrated components, which are illustrated in the following table.

Table 5.2
NCLB Peer Review Components

Reference to NCLB Legislation	Reference to CFR	Content Requirements
Sec.1111(b) (1)	Sec. 200.1	A single statewide system of challenging academic content standards applied to all public schools and LEAS.
Sec.1111(b) (1)	Sec. 200.1	A single statewide system of challenging academic achievement standards applied to all public schools and LEAS.
Sec.1111(b) (3)	Sec. 200.2, 200.3, 200.5, 200.6	A single statewide system of annual high-quality assessments.
Sec.1111(b) (3)	Sec. 200.2	A system of assessments with high technical quality.
Sec.1111(b) (1,3)	Sec. 200.2, 200.3	Alignment of academic content standards, academic achievement standards, and assessments.
Sec.1111(b) (1,3)	Sec. 200.1, 200.2, 200.6	Inclusion of all students in the assessment system
Sec.1111(b) (3)	Sec. 200.8	An effective system of assessment reports.

Source: Legislative Service Office from information provided by the U.S. Department of Education.

U.S. Department of Education provides broad understanding of the process

During our discussions with Wyoming's contact with the Department, we developed a broad understanding of how the process works. However, we were not provided internal policies, training materials for reviewers, operating procedures, etc. It is our understanding from our discussions that the process essentially consists of the following components.

- Peer review team is selected from experts in the areas of assessment and testing. Teams consist of three reviewers that include at least one individual who works with special populations (e.g. students with disabilities, ELL students, etc.). **Note:** Team members' identities are kept confidential and are not disclosed to reviewed states. Therefore, state education officials are not able to correspond directly with officials personally.
- Peer review team members review the assessment materials from selected states independently. Then, they participate in consensus meetings in Washington, D.C. for further discussion.
- The team drafts a written report for the Department, based

on its combined ratings re: the seven integrated components discussed in Table 5.2. The Assistant Secretary for Elementary and Secondary Education is responsible for making a final decision, based on the team's report.

- The Department sends a letter to the selected state designating either full approval, full approval with recommendations, approval expected, approval pending, and non-approval.
- After receiving the Department's letter, the state has the option of sending in additional information to allay concerns with respect to the seven integrated components.

Peer Review Process Becomes Convolved in Practice

States receive broad criteria for meeting the seven (7) integrated components

As discussed previously, in its guidance to the states, the Department provides broad criteria with respect to each of the seven (7) integrated components. According to WDE officials, the criteria can be so broad, that it is difficult to understand exactly what will satisfy a particular requirement.

For example, related to the 1st requirement for challenging academic content standards, possible evidence consists of a *“detailed description of the process the State used in developing its standards to review their rigor, such as its participation in a process to benchmark them to nationally recognized standards. This process should include substantive input from relevant stakeholders and individuals or organizations with expertise in standards development.”* Although this may appear straightforward, there may be discussion with respect to how one defines *“detailed description”* and *“substantive input from relevant stakeholders...”*

Related to Wyoming's experience, a comment from one WDE official is consistent with the challenges raised above. *“Several concerns raised by the peer reviewers, or their misunderstanding of Wyoming legislation, which took months to clarify due to the scheduling and processes set by the USDE, could have been clarified in a phone call. Through peer notes and conversations with USDE, the request was made for evidence that the State Board of Education had approved the Academic Achievement Standards. The WDE provided evidence that our State Superintendent of Public Instruction has the authority to approve*

the Academic Achievement Standards. Each time this request was made, WDE had to resubmit this evidence.”

The process may be more convoluted and disjointed since there is no standard operating procedure within the CFR for peer reviewers to follow. In other words, peer review teams have no cohesive and consistent manner in which they evaluate all states. Because of this, each state could potentially receive a different review based on differences in interpretation by peer reviewers of the guidance.

Subsequently, the peer review process becomes more confusing and frustrating to states because they are not allowed to talk to the actual reviewers assigned to their submission. As a result, feedback comes in the form of notes taken by reviewers during the review session, which may be open to interpretation. From what we understand, states are also subject to a rotating group of peer reviews and liaisons from year to year, which may cause inconsistency in the way states are treated in their review.

Our review of Department letters point towards inconsistency across states

Our review of Department approval letters confirmed that states were treated differently in the peer review process. For example, states which included Alaska, Arizona, Alabama, Florida, and Georgia received approval for the math and reading portion of their assessment. However, when these states submitted the science portion, the peer review teams re-evaluated all subjects and then revoked approval status for all content areas subject to submission of more evidence.

Another example comes from Arkansas and Georgia, which were placed on mandatory oversight after their first submission with the possibility of Title 1 funds being withheld. This is in contrast to Nebraska, which received numerous warnings before being placed on mandatory oversight.

Our review of approval letters also showed that some reviewed states in 2006 and early 2007 passed through the peer review process more seamlessly than others who were reviewed after mid 2007. The process for those states who applied early may have been expedited by the expectation in the guidance that articulates a 120 day or less timeframe for approval.

These inconsistencies have several possible explanations. According to a Department official, states that passed quickly usually had an adaptable state assessment in place, which already fit with NCLB requirements. For example, Delaware, Indiana, Ohio, Utah, Tennessee passed review in early 2006/2007 with only minor deficiencies.

According to a federal education expert, the peer review process has changed over time, which has led to inconsistent evaluation of states. It is possible that both statements in conjunction with each other form one conclusion. The conclusion: that initially the process was simple for states that had assessments closely resembling NCLB standards and as time went on the process became more complex and more difficult for states that were still not approved.

*WDE officials
expressed frustrations
with the process*

After reviewing information submitted and peer review notes related to Wyoming's experience it appears that the peer review process was not transparent and lacked consistency. We reviewed submissions and review notes related to PAWS in technical quality, assessment reports, alignment, content standards, academic achievement standards, and alternate assessments. Over four submission periods spanning three years it was evident that review notes were not consistent from one submission to the next.

In addition to inconsistent evaluation it appears the peer reviewers stepped outside of their regulatory bounds by requiring additional information on instructionally supportive material which had no impact on AYP. This conclusion was supported by WDE officials and the chair of Wyoming's Technical Advisory Committee. The basic premise being that Wyoming's assessment through the *traffic-signal-process* as an instructionally supportive measure went above and beyond federal requirements. This process was heavily scrutinized by federal reviewers even though it appeared to be beyond the bounds of federal requirements.

Eventually, this component of PAWS was removed from Wyoming's submission and more attention was focused on portraying the exam as being aligned to state standards which resulted in approval of the math and reading portion of PAWS.

However, it should be noted that the WDE did receive guidance from past peer reviewers to help build and review their submission. Even though these extraordinary measures were taken, the third and fourth submission still did not receive approval. Currently, Wyoming is still entangled in this process as it is not approved in the area of alternate assessment or science.

The following table provides specific information on peer review comments related to instructionally supportive items (*traffic-signal-process*) and dates in between reviews.

Table 5.3
Wyoming NCLB Peer Review Comments

Review Date	Months In- Between Reviews	Comments
February 2006 (1 st Review)	N/A	WY references a Popham comment (3.4.4) regarding instructionally supportive purposes of assessments, but that paper provides no evidence about WY.
February 2006 (1 st Review)	N/A	Document 7.3.1 is the initial effort to provide evidence about student achievement. There is a color coded schema that is designed to be tied to student achievement levels as stated in document 7.3.2. (Tryout of the PAWS traffic signal reporting system page 2 of the appendix). The color coding schema as described in 7.3.2 is not consistent with NCLB requirements due to lack of connection to standards.
January 2007 (3 rd Review)	11	Given that the main focus of the PAWS is to be instructionally supportive, the lack of clear alignment between test design, scoring and reporting, limits the meaningfulness of the traffic light reporting system. It is not clear whether the subtests/clusters are equated from year to year. If the clusters are not equated from year to year, this potentially limits the continuity and meaning, which could lead to erratic interpretation of results.

Review Date	Months In- Between Reviews	Comments
January 2007 (3 rd Review)	11	<p>How does the traffic light feedback relate to the student's performance level?</p> <p>What criteria are used to make the traffic light color distinctions?</p> <p>It appears that the traffic light color determinations are made based upon a limited number of items. For example, according to the baseline report (#85), 8th grade measurement system includes three items.</p> <p>How does the student's performance on these three items translate into red, yellow, and green performance feedback?</p> <p>Is there validation evidence to support the green feedback statement that "no further instruction is needed?"</p>
April 2007 (4 th Review)	3	<p>There is no evidence to support the claim that the assessment system allows for instructionally relevant feedback to be provided.</p> <p>Clarification regarding the relationship between the traffic light reporting system and the PAWS achievement standards.</p> <p><i>Response from WDE</i> - The traffic light displayed was removed from the parent and student reports but maintained on the classroom roster.</p>
April 2007 (4 th Review)	3	<p>This is especially true for a system that is conceived as being "Instructionally Supportive". This process should be clearly described in the technical manual.</p>

Source: Legislative Service Office from information provided by WDE.

Peer review process has impacted WDE resources

According to WDE officials, the back and forth peer review process has not only impacted WDE financially, but also its staff resources. For example, from BFY 2005 through BFY 2009 \$592,717.77 has been expended for services provided by the Wyoming Technical Advisory Committee (TAC), who provided assistance to WDE with respect to federal peer review and other questions. In addition, \$263,851.50 was expended for other

consulting services related to the federal peer review process for various alignment studies.

This notion is supported by the Chair of the Technical Advisory Committee in his book *Unlearned Lessons, Six Stumbling Blocks to our Schools' Success*, where he discusses the small size of the WDE assessment team. Dr. Popham states *"The WDE staff was quite literally crippled for about two years by a straight-jacketing federal review process."* He also spoke to the underpinning of the slow process *"But, frankly, the PAWS tests were evaluated by traditional peers who brought traditional perspectives to the peer review of an untraditional assessment program."*

In addition, he provides context by detailing how for two years WDE's assessment team worked for federal peer review approval. *"And so it was that for almost two years, the WDE assessment staff was obliged to spend an enormous amount of time trying to obtain peer-review approval of the PAWS program. Members of the Wyoming TAC looked on with dismay at the lack of attention being given to PAWS-related instructional initiatives."*

According to WDE officials, after states expressed frustration to the Department, it offered the opportunity to use a fourth reviewer to serve as liaison between the team and the state. This would provide an opportunity for state officials to actually sit down with the fourth reviewer and discuss the team's notes. According to WDE officials, Wyoming made this request to the Department related to the April 2008 Science peer review. Although other states were taking advantage of the fourth peer review, Wyoming's verbal and written requests were not granted, according to WDE officials.

Another example of proposed assistance by the Department was provided in correspondence from the Department to WDE dated June 30, 2006. In the letter the Department states, *"We recognize the challenges that you face as you work to implement valid and reliable standards and assessment systems. As such, we would like to offer a wide range of technical assistance, particularly in those areas that have proven to be especially challenging."*

According to WDE officials however, it received no assistance until one year later from a technical advisor from the Department. Even with that individual's assistance as well as his confidence that Wyoming's PAWS-ALT would pass peer review, Wyoming did not receive approval. As of the date of this report, Wyoming has still not passed the peer review process. However, it has secured peer review approval for general reading/language arts and mathematics (as of November 13, 2007).

Previous vendor may not have had the expertise to move an instructionally supportive component forward

Dr. Popham in his book also suggests that the previous vendor (Harcourt Assessment) may not have been the ideal vendor to contribute to the success of an integrated statewide assessment. Although Harcourt was selected out of nine (9) vendors that expressed interest in administering an instructionally supportive assessment for NCLB, he indicates the company may have been too traditional. *"It soon became apparent, however, that although Harcourt personnel were remarkably skilled in carrying out the many traditional operations called for as part of statewide testing, they had little to contribute on the instructional side of the equation."*

Given that Wyoming was experimenting with a new approach in its PAWS design, it may very well be that none of the interested vendors were appropriate for moving Wyoming's vision forward. What can be said however is that Dr. Popham, with collaboration from WDE officials, as well as a new approach to statewide assessment, was not necessarily welcomed with open arms by the Department. This caused additional resources to be utilized to pass the federal peer review.

Recommendation:

The Management Audit Committee may wish to consider requesting that the U.S. Senate Committee on Health, Education, Labor, and Pensions request that the Government Accountability Office (GAO) conduct an audit of the federal peer review process pending reauthorization of ESEA.

Given the frustrations of Wyoming, as well as other states with

the federal peer review process, it may be beneficial for the GAO to review the process prior to reauthorization of ESEA. Reviewing the process, as well as exploring various states experiences, may assist the Department in fine-tuning the process for the future.



CHAPTER 6

Elementary and Secondary Education Act (ESEA) Impact on PAWS

Finding 1: Elementary and Secondary Education Act (ESEA) will surely impact PAWS.

According to sources in the press, as well as various state and federal officials, ESEA will more than likely be reauthorized by the end of 2011. Generally, there is a consensus that an overemphasis exists on standardized testing for accountability purposes, which may drastically change the current law.

Wyoming, like every other state in the nation, is experiencing uncertainty with respect to their statewide assessments, as it ponders changes beyond No Child Left Behind (NCLB). As a result, states are in a holding pattern with respect to their statewide assessments, even though they anticipate substantive changes in the next reauthorization.

ESEA (Pub. L. 89-10, 79 Stat. 27, 20 U.S.C. Ch. 70) was originally enacted on April 11, 1965. The Act is extensive; its sole purpose is to fund primary and secondary education in the United States in the following areas: professional development; instructional materials; resources to support educational programs; and parental involvement. More specifically, the ACT includes the following major provisions:

- Title I-Improving the Academic Achievement of the Disadvantaged;
- Title II-Preparing, Training, and Recruiting High Quality Teachers and Principals;
- Title III-Language Instruction for Limited English Proficient and Immigrant Students;
- Title IV-21st Century Schools;
- Title V-Promoting Informed Parental Choice and Innovative Programs;
- Title VI-Flexibility and Accountability;

- Title VII-Indian, Native Hawaiian, and Alaska Native Education; and
- Title VIII-Impact Aid Program.

The Act, originally authorized through 1970, has been reauthorized every five years since its enactment, until 2001. The following are the most recent reauthorizations.

- Education Consolidation and Improvement Act of 1981
- Improving America's Schools Act of 1994
- No Child Left Behind Act of 2001

*NCLB Focuses on
Standardized Testing
for all States*

The most recent reauthorization of ESEA is the No Child Left Behind Act of 2001. The original bill received bipartisan support, and was signed into law on January 8, 2002. NCLB's focus is standards-based education reform, which focuses on setting high standards and establishing measurable goals to improve education. For example, NCLB requires that every child will become proficient at reading and math by 2014.

The Act requires all states to develop assessments in basic skills to be administered in certain grades. More specifically, 20 USC 6311, Section 1111 (b) (3) requires a single statewide assessment for students in grades 3 through 8 in mathematics, reading, language arts and science (no later than SY 2007-2008).

NCLB also requires that assessments measure against challenging state academic content and achievement standards (Section 111 (b) (2)). However, NCLB does not require national achievement standards as those standards are set by individual states. Also, as originally envisioned by ESEA, the Act continues to explicitly forbid the establishment of a national curriculum.

NCLB requires states to develop assessments for administration to all students in certain grades, as a contingency for federal funding. The Act includes a measurement to determine how public schools are doing with respect to standardized test results. The measure is called Adequate Yearly Progress (AYP), which essentially is a diagnostic tool to help schools determine if they are meeting state standards.

All schools are required to demonstrate AYP in the areas of reading/language arts, mathematics, as well as graduation rates and/or attendance rates for all grades. Generally, schools should meet the following requirements: 1) Students are assessed for language arts and mathematics; 2) Students should meet the minimum annual target for meeting or exceeding state standards for language arts and mathematics; and 3) students should meet the participation rate, AYP targets, and reading proficiency (elementary and middle schools) and meet participation rate, AYP targets, and graduation rate (high school).

Indications Exist on Direction of Reauthorization of ESEA

President Obama's Blueprint for Success ties to federal funding

On March 13, 2010 the Obama administration released its blueprint for revising the Elementary and Secondary Education Act (ESEA). Generally, it recommends incentives for states to adopt academic standards that prepare students to succeed in college and the workplace, and suggests creating accountability systems that measure student growth toward meeting the goal that all children graduate and succeed in college. More specifically, it touts the development of college and career ready standards, as well as developing assessments aligned to those standards for English language arts and mathematics.

More importantly for states, the blueprint is tied to funding related to the American and Recovery Reinvestment Act (ARRA) of 2009. It is tied directly to requirements and criteria for Race to the Top (RTT) funds (\$4.35 billion), which were announced in July, 2009. Much like RTT, the blueprint recommends that beginning in 2015, *“formula funds will be available only to states that are implementing assessments based on college and career ready standards that are common to a significant number of states.”*

In order to receive RTT funds, states are encouraged to adopt common standards for K-12. The Common Core State Standards Initiative (CCSSI) is coordinated by the National Governors Association Center for Best Practices (NGA Center) and the Council of Chief State School Officers (CCSSO). The standards were developed in collaboration with teachers, school administrators, and experts. They provide a framework for

measuring students' progress towards readiness for college and the workforce. In addition to RTT funds, an additional \$350 million in federal grants exists for the development of assessments aligned to the common standards.

RTT caused the majority of states (48) to adopt the common standards discussed above. Adoption of the standards appeared to accelerate after the federal government's August 1, 2010 deadline for adoption, after which states would not receive points towards round two (2) applications.

Although Wyoming applied for the first round of RTT funds, it was not successful. Only two states (Delaware and Tennessee) successfully received funds. Wyoming did not apply for the second round of funding. For that round, only ten (10) states were successful (e.g. District of Columbia, Florida, Georgia, Hawaii, Maryland, Massachusetts, New York, North Carolina, Ohio, and Rhode Island).

Smarter Balanced Assessment Consortium and the Partnership for Assessment of Readiness for College and Careers

The Smarter Balanced Consortium (SBAC) is a collection of states (31 as of June 17, 2010) that have been working collaboratively to develop a student assessment system that is aligned to the common standards discussed above. Since December 2009, the Consortium has been working on this project in order to apply for the \$350 million RTT grants to develop common assessments.

The Partnership for Assessment of Readiness for College and Careers (Partnership) is also a consortium of states that has been working to compete for RTT grants to develop common assessments. The Partnership consists of 26 states, which "*educate more than 60 percent of the K-12 students in the United States.*" On behalf of the Partnership, Florida submitted the application for RTT funds.

On September 2, 2010, the SBAC was awarded a four-year \$160 million RTT grant by the U.S. Department of Education (Department) to develop a statewide assessment system aligned to the common core of standards. The other grant in the amount of \$170 million was awarded to the Partnership.

The following table provides additional information on member

states of each consortium.

Table 6.1
Member States of Two Consortia

Consortium	Member States	RTT Grant Award
Smarter Balanced Assessment Consortium	Alabama, Connecticut, Georgia, Idaho, Kansas, Maine, Missouri, Nevada, New Jersey, North Carolina, Ohio, Oregon, South Carolina, Utah, Washington, Wisconsin, Colorado, Delaware, Hawaii, Iowa, Kentucky, Michigan, Montana, New Hampshire, New Mexico, North Dakota, Oklahoma, Pennsylvania, South Dakota, Vermont, and West Virginia.	\$160 million
Partnership for Assessment of Readiness for College and Careers	Arizona, District of Columbia, Florida, Illinois, Indiana, Louisiana, Maryland, Massachusetts, New York, Rhode Island, Tennessee, Alabama, Arkansas, California, Colorado, Delaware, Georgia, Kentucky, Mississippi, New Hampshire, New Jersey, North Dakota, Ohio, Oklahoma, Pennsylvania, and South Carolina.	\$170 million

Source: Legislative Service Office from information provided on the websites of the Smarter Balanced Consortium and the Partnership for Assessment of Readiness for College and Careers.

Wyoming adopts common core standards but is adopting a “wait and see” attitude on joining a consortium

On June 16, 2010 the Wyoming State Board of Education voted to adopt the common core standards. More specifically, the Board approved the following motion:

“The Board approves the adoption of 100% of the common core state standards in English Language Arts and Mathematics in the next revision of the Wyoming Content Performance Standards to be completed by December, 2011. The common core state standards will comprise at least 85% of the revised Wyoming Content Performance Standards in each of these two content areas.”

As a result of this motion, it appears Wyoming is positioning itself for future funding under a reauthorized ESEA. The

Board's decision appears to be a logical step, given the blueprint's pronouncement that by 2015, categorical grants will only go to states that have similar assessments to provide measurement towards college and career readiness standards.

By adopting the standards first, Wyoming is in a position to join one or both of the above consortia, if ESEA is reauthorized as envisioned by the blueprint. Joining a consortium has benefits obviously, such as being cost effective, benefiting from other states' expertise, and the ability to compare Wyoming students with other states using the same assessment. However, the obvious drawback is that member states lose control over the development process in the areas of test design, item develop, and item scoring. As a result, it may be best to adopt a "wait-and-see" attitude.

Although the blueprint clearly articulates direction, ESEA has yet to be reauthorized

At this point, it should be noted that although the blueprint annotates a clear direction for pending ESEA reauthorization, only time will tell with respect to how Congress will actually change the current law. Also, there are legitimate concerns shared by teacher unions and state officials (executive and legislative) that the blueprint and RTT criteria grants too much authority to the federal government, and that future funding will rest on adoption of shared standards and assessments.

These concerns are offset though, by others who opine that as long as states make their own decisions to adopt shared standards and assessments that are free from federal influence or requirements, sharing standards and assessments could be very effective. In other words, the ability to make independent decisions about which standards and assessments to use, even though they are shared, would not lead to de facto national standards. Both points are sure to be debated in Congress when ESEA is reauthorized.

From what we understand from various education experts, it is highly likely that the core of ESEA will remain intact re: requirements for content standards, academic achievement standards, as well as grade-level summative assessments in reading and mathematics that disaggregate student data. Of course, established consortia that have successfully developed assessments for their member states, may be accepted as part of

ESEA reauthorization.

There will more than likely be some type of measurement for AYP, although it is unclear whether it will focus strictly on standardized test scores to illustrate proficiency, as opposed to broad growth indicators that show gradual proficiency towards college or career readiness. Whatever form AYP takes however, it will probably be based on some type of requirement for a summative assessment(s) measuring against accepted common or state standards.

Current Options for Wyoming

Because of the concerns and problems identified in Chapters 3 and 4, we have heard various options expressed through the news media, during our interviews, and within our district superintendent survey responses for Wyoming that involve moving towards another vendor or statewide assessment. For purposes of this audit, we annotate these various options and provide pros and cons for each.

The following table provides additional information.

Table 6.2
Options for Wyoming State Legislature

+ Option	Pros	Cons
1. Use the current vendor to amend PAWS after ESEA reauthorization	* Current vendor is already in place and contract language allows for making changes to the assessment based on changes to federal or state law.	*Current vendor has failed to properly administer PAWS during Spring of 2010. This has led to a loss of confidence in the ability of the contractor to perform for Wyoming and its K-12 students.
2. Use the current vendor through the current contract period, or until ESEA is reauthorized.	* Current vendor is already in place and settlement negotiations are moving forward. * Riding out the current contract, will not likely cause additional costs to be incurred by Wyoming. * Wyoming is in a position to work with the vendor to make	* Current vendor has failed to properly administer PAWS during Spring of 2010. This has led to a loss of confidence in the ability of the contractor to perform for Wyoming and its K-12 students.

+ Option	Pros	Cons
	<p>improvements to the testing platform for future years.</p> <p>*The least disruption and change to districts. PAWS would be administered through 2013.</p>	
<p>3. Use the current vendor to study ESEA reauthorization re: the blueprint, RTT funding, development of consortium in anticipation of changes to PAWS.</p>	<p>* Current vendor is already in place. Because of past problems with administration and ongoing negotiations, Wyoming could influence the vendor to perform additional study and planning re: ESEA reauthorization.</p>	<p>* Current vendor has failed to properly administer PAWS during Spring of 2010. This has led to a loss of confidence in the ability of the contractor to perform for Wyoming and its K-12 students.</p> <p>* Granting additional funding to a vendor that has already let Wyoming's students down would not be politically or practically advantageous.</p>
<p>4. Canceling the current contract for cause and reissuing an RFP to continue administering PAWS to meet NCLB requirements.</p>	<p>*Selecting a new vendor could provide comfort to those who have completely lost faith in the current contractor.</p> <p>* Wyoming would be seen by the public as taking decisive action to correct a problem.</p>	<p>*++It is not likely that an RFP could be issued, responses analyzed, and a new assessment in place by 2011.</p>
<p>5. Canceling the current contract for cause and reissuing an RFP in anticipation of ESEA reauthorization.</p>	<p>* Selecting a new vendor could provide comfort to those who have completely lost faith in the current contractor.</p> <p>*Wyoming would be seen by the public as taking decisive action to correct a problem.</p>	<p>*Although a new vendor could study pending ESEA reauthorization, it is not likely that current activities (e.g. blueprint, RTT funding, development of consortium in anticipation of changes to PAWS) will be mirrored in the actual reauthorization. This could cause additional funds to be spent prematurely, without fully understanding federal requirements.</p>

Source: Legislative Service Office from information provided during interviews and district superintendent surveys.

+It should be noted that Section 8 (H) of the current contract, allows for changes in the contract based on changes to federal and state law.

+ +According to WDE, it could take up to four (4) years to develop a new statewide assessment. See Table 6.3 for specific timelines.

**Table 6.3
Detailed Timeline to Change the Statewide Assessment**

Description of Task	Date
1. Hold Wyoming Statewide Assessment Taskforce Meetings to determine the means by which academic constructs would be measured to meet federal and state assessment requirements	December 2010-March 2011
2. Assessment Technical Advisory Committee works with the WDE to develop specifics of test design by grade and content area to be included in the RFP	March 2011- June 2011
3. Pre-proposal Conference	May 2011
4. WDE and Technical Advisory Committee write the RFP	May-September 2011
5. State Issues Request for Proposal	September, 2011
6. Notice of Intent to Propose Deadline	October 2011
7. Vendors Written Questions Deadline	October 2011
8. State Responds to Vendors Written Questions	October 2011
9. Proposal Submission Deadline	November 2011
10. Agency Invitation to Contractors to provide Oral Presentations (if Agency determines need)	November 2011
11. Oral Presentations	December 2011
12. Agency Complete Evaluation and Calculation of Scores for Technical and Cost Proposals	January 2012
13. Contract Negotiations	January – March 2012
14. Contract Signature Deadline	March 2012
15. Contract Start Date	March 2012
16. Item development begins for initial 2013 field test (Extremely compressed timeline for initial field test items)	March 2012
17. Item Review	August 2012
18. Item reconciliation	September 2012
19. Item development begins for embedded 2014 field test	November 2012
20. Develop Scorer Protocols	September-October 2012
21. Current PAWS Assessment Administration (from an amended contract for an additional year of PAWS)	March 2013
22. Stand-Alone Field Test for new state assessment	April 2013
23. Item Review	July 2013
24. Data Review	July 2013
25. Rangefinding	October 2013
26. First operational administration of new state assessment	March 2014
27. Preliminary Standard Setting	May 2014

Description of Task	Date
28. Final Standard Setting	June 2014
29. Alignment Study	July 2014
30. Peer Review Submission to USED	November 2014

Source: Legislative Service Office from information provided by WDE.

Wyoming appears to be waiting for ESEA reauthorization before making changes to PAWS

During the 2010 Wyoming Budget Session, HB 1, Section 334 created the School Finance Recalibration Committee (Committee). As part of the Committee's overall charge to study recalibration of the education resource block grant model, it is also required to "*review methods of assessment of student achievement under the current model and recommend student assessment policies and procedures in the recalibrated system of education.*"

Although the Committee did not specifically study PAWS, it did discuss problems with PAWS, as well as its administration. There was also discussion as to whether PAWS should be replaced with another assessment, etc. In the end however, legislators and educators understood the impact ESEA reauthorization will more than likely have on states and their assessments.

The Committee also identified PAWS as an assessment that may be used to measure how well districts are doing as part of a proposed statewide accountability system. With the caveat however, that it may need to develop a subcommittee to review PAWS. This of course could include adjustments, etc., with respect to pending ESEA reauthorization.

Recommendation: WDE should work with the current vendor to correct discrepancies with respect to PAWS administration during Spring 2010. In addition, it should not make drastic changes to the current assessment, until ESEA has been reauthorized.

Wyoming, like most states in the nation, is at an important crossroads. Wyoming must continue to meet the requirements of NCLB, but must also look forward to ESEA reauthorization and what that means to Wyoming citizens and their children. Wyoming also has an added concern with respect to its current vendor NCS Pearson, Inc.

The natural reaction to PAWS administration problems in SY 2009-2010 is to take decisive action with respect to firing the current vendor and finding another company to administer PAWS. Although this is an option, it would surely cost the state additional time and resources to accomplish. It appears WDE, through its settlement negotiations with the vendor, as well as anticipating ESEA reauthorization through the adoption of common-core standards, and leaving the door open to join a consortium, is hedging its bets.

The lingering questions related to PAWS, vendor concerns, etc., could all be moot within the next year or two. After ESEA is reauthorized. Wyoming will then need to decide how it handles the rest of the current contract.

Recommendation: WDE should study the activities at the federal, state, and local levels with respect to the direction of ESEA reauthorization. Information related to the blueprint, RTT funding, as well as the development and funding of various consortia can be used to position itself for eventual ESEA reauthorization.

WDE should continue to study the activities at the federal level in order to place itself in a flexible position once ESEA is

reauthorized. Like other states, Wyoming can glean an idea of the direction of reauthorization, but will not know for sure until it actually happens.

CHAPTER 7

Other States

Finding 1: Wyoming and other states are not making changes to statewide assessments until after reauthorization of ESEA

Wyoming, like every state in the United States, is anticipating ESEA reauthorization. As discussed in Chapter 6, it appears there will be substantive changes as a result of reauthorization related to requirements for statewide assessments. The common perception is that ESEA reauthorization will reflect President Obama's blueprint for reauthorization, released on March 13, 2010.

Generally, the blueprint recommends incentives for states to adopt academic standards that prepare students to succeed in college and the workplace, and suggests creating accountability systems that measure student growth toward meeting the goal that all children graduate and succeed in college. More specifically, it touts the development of college and career ready standards, as well as developing assessments aligned to those standards for English language arts and mathematics.

However, it is highly likely that the core of ESEA will remain intact re: requirements for content standards, academic achievement standards, as well as grade-level summative assessments in reading and mathematics that disaggregate student data. Of course, established consortia that have successfully developed assessments for their member states may become models after ESEA reauthorization, with respect to potentially sharing standards and common assessments.

Finally, it is unclear whether Adequate Yearly Progress (AYP) requirements will focus strictly on standardized test scores to illustrate proficiency, as opposed to broad growth indicators that show gradual proficiency towards college or career readiness. It

is also likely that AYP at a minimum will be based on some type of requirement for a summative assessment(s) measuring against accepted common or state standards.

States use a variety of assessments to meet NCLB requirements

States use a variety of statewide assessments to meet NCLB requirements. As long as states successfully illustrate to the federal government that statewide assessments align to their content and performance standards, there is flexibility as to the development of their assessments. The primary requirements however, are that statewide assessments should be summative in nature, as well as testing the appropriate grades and topic areas.

Although the U.S. Department of Education does not have specific information on its website to annotate the assessments used by each state, we were able to obtain information from the Council of Chief State School Officers (CCSSO) to provide other states' context. The following information from CCSSO provides more specific information related to the array of assessments used by various states and whether states have received full peer review approval.

***Table 7.1
Other States
Statewide Assessments: SY 2009-10**

State	Program	**Full Approval by U.S. DOE
Alabama	Alabama Reading Test Alabama Mathematics Test Alabama High School Graduation Exam Alabama Science Assessment Test	Yes
Alaska	Standards Based Assessments Alaska Science Assessment	No
Arizona	Arizona Instrument to Measure Standards (AIMS)	No
Arkansas	Benchmark Exams End of Course Exams Literacy Exams Arkansas Science Exam (Under Development)	Yes
California	California Standard Tests (CSTs) Standards Based Tests-In Spanish California High School Exit Exam (CAHSEE)	No

State	Program	**Full Approval by U.S. DOE
Colorado	Colorado Student Assessment Program	Yes
Connecticut	Connecticut Mastery Test (CMT) Connecticut Academic Performance Test (CAPT)	Yes
DC	District of Columbia Comprehensive Assessment System District of Columbia Comprehensive Assessment System Biology	No
Delaware	Delaware Student Testing Program	Yes
Florida	Algebra I EOC Test Florida Comprehensive Assessment Test	Yes
Georgia	Criterion-Referenced Competency Tests (CRCT) End of Course Test (EOCT)	Yes
Hawaii	Hawaii State Reading Assessment; Terra Nova Hawaii State Mathematics Assessment; Terra Nova Hawaii State Science Assessment	No
Idaho	Idaho Reading Indicator Idaho State Achievement Test (ISAT)	Yes
Illinois	Illinois Standards Achievement Test (ISAT) Prairie State Achievement Examination (PSAE)	Yes
Indiana	Indiana Statewide Testing for Educational Progress Plus (ISTEP+) Graduation Qualifying Exam	Yes
Iowa	Iowa Tests of Basic Skills (ITBS) Iowa Test of Education Development	Yes
Kansas	Kansas State Assessment	Yes
Kentucky	Kentucky Core Content Test	Yes
Louisiana	Louisiana Educational Assessment Program (LEAP) Integrated Louisiana Educational Assessment Program (iLEAP) Graduation Exit Examination (GEE)	Yes
Maine	New England Common Assessment Program (Reading) PSAT New England Common Assessment Program (Mathematics)	No
Maryland	Maryland School Assessment (MSA)	Yes
Massachusetts	Massachusetts Comprehensive Assessment System (MCAS)	Yes
Michigan	Michigan Educational Assessment Program (MEAP) Michigan Merit Examination	No
Minnesota	Minnesota Comprehensive Assessments-Series II (MCA-II)	Yes
Mississippi	Mississippi Curriculum Test Subject Area Testing Program-English II, Algebra I, Biology (SATP)	No
Missouri	Missouri Assessment Program (MAP) EOC-English II EOC-Algebra I EOC-Biology	No
Montana	Montana Comprehensive Assessment System (MontCAS)	Yes

State	Program	**Full Approval by U.S. DOE
Nebraska	School-Based Teacher-Led Assessment and Reporting System (STARS)	No
Nevada	Criterion-Referenced Tests High School Proficiency Examination-Reading High School Proficiency Examination-Mathematics High School Proficiency Examination-Science	No
New Hampshire	New England Common Assessment Program-Writing & Reading New England Common Assessment Program-Mathematics New England Common Assessment Program (Science)	No
New Jersey	New Jersey Skills & Knowledge Assessment (NJASK) Biology Test	No
New Mexico	New Mexico School Assessment Program	Yes
New York	English Language Assessment Tests Mathematics Assessment Tests Science Examination	Yes
North Carolina	End of Course Exams-English I End of Course Exams-Algebra I, Algebra II, and Geometry End of Course Exams-Biology, Chemistry, Physics, Physical Science) End of Grade Assessment	No
North Dakota	North Dakota State Assessment	No
Ohio	Ohio Achievement Test Ohio Graduation Test	Yes
Oklahoma	Oklahoma Core Curriculum Tests	No
Oregon	Oregon Assessment of Knowledge and Skills Online-Reading Oregon Assessment of Knowledge and Skills Online-Mathematics Oregon Assessment of Knowledge and Skills Online-Science	Yes
Pennsylvania	Pennsylvania System of School Assessment (PSSA) Pennsylvania System of School Assessment (PSSA)-Science	No
Rhode Island	New England Common Assessment Program (NECAP) New England Common Assessment Program (NECAP)-Science	Yes
South Carolina	End-of-Course Examination Program (EOCEP) Palmetto Assessment of State Standards	
South Dakota	SD Test of Educational Process Dakota Assessment of Content Standards (DACS) Achievement Series Assessments	Yes
Tennessee	Tennessee Comprehensive Assessment Program Achievement Tests Tennessee Comprehensive Assessment Program End of Course	Yes
Texas	Texas Assessment of Knowledge and Skills (TAKS) End of Course (optional until the 2011-2012 school year)	Yes
Utah	Utah Core Curriculum Science Core Criterion Reference Tests	No

State	Program	**Full Approval by U.S. DOE
Vermont	New England Common Assessment Program (NECAP)	No
Virginia	Standards of Learning (SOL) Assessments End of Course Assessments	Yes
Washington	High School Proficiency Examination Measurement of Student Progress	Yes
West Virginia	West Virginia Educational Standards Tests (WESTEST)	Yes
Wisconsin	Wisconsin Knowledge & Concepts Examinations (WKCE)	Yes
Wyoming	Proficiency Assessments for Wyoming Schools (PAWS)	No

Source: Education Commission of the States webpage.

* CCSSO has not updated this list as of this report's writing.

**Full approval denotes approval of assessments used for mathematics, language arts, science, and alternative.

While there is variation with respect to the grades tested, all states listed above have assessment components for mathematics, reading, science, and alternative. Rather than conducting research on numerous states however, we focused on states with traits we thought could be applied to Wyoming, given the findings in our report.

Comparable states showed Wyoming expended more funds towards development of its statewide assessment

We surveyed five (5) comparable states to Wyoming, in order to obtain more specific information with respect to the development of other states' statewide assessments. Because Wyoming's population is lower than the majority of states in the United States (e.g. 544,270 as of the July 2009 U.S. Census) we focused on states with similar populations (i.e. Alaska, Montana, North Dakota, South Dakota, and Vermont).

The following table provides additional information from our surveys of selected states.

**Table 7.2
Selected States' Information**

Question	Alaska	Montana	N.D	S.D	Vermont
Population	698,473	974,989	646,844	812,383	621,760
Assessment Name	Standards Based Assessment	Montana Comprehensive Assessment System (MontCAS)	N.D. State Assessment	S.D. State Test of Educational Progress (Dakota STEP)	New England Common Assessment Program
Instructionally Supportive	**Yes	***Yes	No	****Yes	+Yes
Assessment Type	Summative	Summative	Summative	Summative	Summative
Average Time to Complete	One to three hours per subject	Three to six hours (all subjects)	Six to ten hours	Three to six hours for 3 rd , 4 th , 6 th , and 7 th grades- six to ten hours for 5 th , 8 th , and 11 th grades	Three to six hours for 3 rd , 4 th , 6 th , and 7 th grades-six to ten hours for 5 th , 8 th , and 11 th grades
*High Stakes	No	No	Did not provide information	No	++Yes
Writing Component	Yes	No	Yes	No	Yes (5 th , 8 th , and 11 th grades)
Vendor Name	CTB/McGraw-Hill	Measured Progress	CTB-McGraw Hill	Harcourt then NCS Pearson	Measured Progress
Satisfied With Vendor	Yes	Yes	Yes	Yes	Yes
Vendor Cost	Did not provide information	Did not provide information	Approximately \$3 million annually	\$850,000 per year for development and \$3.8 million for SY 2010-11.	\$2 million for development and \$2.25 million each year for administration
Peer Review Approval Status	Anticipates full approval	In process for science	In process for alternate assessment	Full approval	In process of alternate assessment
Opinion of Peer Review	Extremely burdensome and not consistent	Did not provide information	Burdensome	Extremely burdensome	Burdensome

Question	Alaska	Montana	N.D	S.D	Vermont
	across states				
Action(s) Re: ESEA Reauthorization	Not changing assessment	Did not provide information	Not changing assessment	Not changing assessment	Not changing assessment
State Moving Towards Growth Model	Uses one growth model for AYP and one for state accountability	Did not provide information	Uses one growth model	No	Reports growth statistics, but waiting for reauthorization before using for accountability

Source: Legislative Service Office from information provided by other states.

*High stakes is defined as either of the following: used for graduation or college scholarship criteria.

**Alaska uses online formative assessment system (ACFA) that utilizes items linked to their standards. It also conducts curriculum alignment institutes.

***Montana's Analysis and Reporting System allows schools and districts to access student scores by student, item, and standards. Districts and schools receive training from Montana's Office of Public Instruction.

****At the state level, a new initiative was set forth in reading.

+ Vermont provides a number of options for teachers, in excess of NCLB requirement, to make the assessment useful to teachers, including an online reporting tool which provides an array of data on individual students, released items, etc.

+ +Included on student transcripts in some, but not all schools (local option).

Common themes exist within selected states As seen from Table 7.2, there are some common themes that exist within Wyoming and our selected states. First, and the most obvious, all six states use summative assessments to meet NCLB requirements. Summative assessments are once a year assessments, designed to assess a student's mastery of content, knowledge, or skills for accountability purposes. They can either be norm referenced, measuring against their peers, or criterion referenced, measuring against a set of criteria. NCLB requires that summative assessments be used by states to measure AYP.

Instructionally Supportive

According to the surveys, we see that all but the state of North Dakota considers their statewide assessments to be instructionally supportive. It should be noted however, that Wyoming has more of a defined process (e.g. traffic-signal process) where reports and suggestions are made available to districts and schools. The states discussed above, although they do not have a defined process like Wyoming, do have some form of instructional component to their assessment (e.g. formative assessments aligned to test items and the ability to use a system to compare

and analyze student scores to selected items and standards).

Length of Time to Administer Assessment

Throughout our audit, we continually heard that nine (9) hours for a student to take the PAWS assessment was much too long. However, when compared to the selected states above, we see only two states (Alaska and Montana) where assessments always take less than nine hours. The remaining states (South Dakota, Vermont, and North Dakota) may take six (6) to ten (10) hours to administer their assessments, especially for 5th, 8th, and 11th grades.

High-Stakes Purposes

We also see from the table above, that all states do not require assessment scores for other high-stakes purposes such as graduation requirements or college scholarship requirements. Only one state (Vermont) allows local districts to include assessment scores on student transcripts. However, this is not required of the local districts.

Writing Component and Vendor Satisfaction

The majority of the states we selected, also include a writing component to their statewide assessment. The only state that does not include such a component is South Dakota. It is interesting to note that all states were satisfied with their vendors, even the state of South Dakota, which used Harcourt and then NCS Pearson, Inc. In fact, the South Dakota official who completed our survey stated the following about each vendor's performance as follows:

- *“Our program manager is very efficient and helps the SDDOE any way she can assist.” (Harcourt)*
- *“We have not had any problems with our scoring service.” (NCS Pearson, Inc.)*
- *“The SDDOE does not have psychometric services and relies on our vendor with backup from our TAC members.” (NCS Pearson, Inc.)*

Cost

Although two states did not respond to our question about cost, the three other states (North Dakota, South Dakota, and Vermont) all stated they expend less for their statewide assessments than Wyoming. For example, North Dakota stated it expends approximately \$3 million annually for administration of its assessment; South Dakota stated it expends approximately \$850,000 per year for development and \$3.8 million annually for administration; and Vermont expends \$2 million per year for development and \$2.25 million annually for administration. It should also be noted that South Dakota uses NCS Pearson, Inc., compared to Wyoming at roughly \$8 million per year.

Peer Review and Other

Finally, selected states we surveyed have either received full peer review approval from the U.S. Department of Education or are awaiting approval of their alternative or science components. All states with the exception of Montana, which did not provide information, stated the federal peer review process was burdensome and not consistent among all states. It should also be noted that all states with the exception of Montana, which did not provide information, are not making changes to their assessments until after ESEA reauthorization. Also, all states with the exception of Montana and South Dakota use some form of growth measure however, in anticipation of ESEA reauthorization.

Vermont belongs to the New England Common Assessment Program (NECAP) and uses shared assessments and standards

One state, Vermont, uses a shared assessment developed by the New England Common Assessment Program (NECAP). From what we understand, the genesis for the creation of NECAP was Measured Progress, as a result of concern from New England states about meeting the stringent requirements of NCLB.

One individual from that company worked with Rhode Island to coordinate meetings between interested states for regional collaboration. The original states included Rhode Island, Vermont, Maine, and New Hampshire. Eventually though, New Hampshire, Rhode Island, and Vermont agreed to develop a common assessment. Maine later rejoined the consortium after completion of the shared assessment.

NECAP is a series of reading, writing, mathematics and science achievement tests, administered annually, which were developed in response to NCLB. More specifically, states created the regional assessment in order to reduce costs and to help simplify a process to jointly adhere to federal requirements.

The assessments include common grade-level expectations (GLEs) for mathematics and reading at grades 3-8 and high school. Writing is also tested in 5th, 8th, and 11th grades. The assessments align to shared standards, developed by member states. In some states, their standards were broad enough to encompass the shared NECAP standards.

Cost savings, the primary benefit of shared assessments, is realized by member states. For example, according to the NECAP Director, the state of New Hampshire saves about \$1 million per year. And, as seen from Table 7.2, Vermont expends \$2.25 million per year for administration of the assessment, much less than the \$10 million per year Wyoming has expended since 2005. Another benefit was a shared experience related to the federal peer review process. All states made individual, but identical submissions, and received timely approval from the U.S. Department of Education.

NECAP also contracted with the National Center for Assessment to compile the alignment data into one study, which allowed member states to escape the cost of conducting individual alignment studies for each state. The technical advisory committee (TAC) was also created for NECAP, which served all member states, thus saving additional funding and other resources.

According to the NECAP Director, for the next two to three years the NECAP assessment will more than likely be used with little change. However, with the creation of the Smarter Balanced Consortium (Consortium) and its successful grant to develop a common assessment for Race to the Top (RTT), it is likely member states who also belong to the Smarter Balanced Consortium will use its assessment. With the exception of Rhode Island, all NECAP states belong to the Consortium. Although, until ESEA is reauthorized, everything is still somewhat uncertain.

Summary Although we were not able to provide detailed information for all states across the country, we heard from state and federal officials, as well as experts in the field, that states are in a holding pattern until after reauthorization of ESEA. Even though states are anticipating change post reauthorization and some are even establishing supplemental growth models, applying for RTT grants, and joining various consortia, it would be shortsighted for states to spend time and money to redo their statewide assessments before reauthorization. For the most part, states appear content to use their existing assessments until reauthorization.



CHAPTER 8

Conclusion

The Management Audit Committee requested an audit of the Proficiency Assessments for Wyoming Students (PAWS). PAWS is used to assess students' proficiency in reading, math, writing, and science (as of Spring 2008), but problems with the vendor have caused legislators and educators to lose confidence in its effectiveness.

More specifically, concerns center around the cost of PAWS, the time in which it takes to administer, and problems with the testing platform. There are also lingering concerns about the subjectivity of scoring certain portions of PAWS, such as the writing component. Legislators and educators also question whether different assessments already being used in Wyoming school districts could replace PAWS.

There is no doubt that Wyoming took advantage of opportunities presented by the No Child Left Behind Act (NCLB) to dismantle and rebuild its statewide assessment. There are also experts in the field who tout that as far as statewide assessments go, *"PAWS is the best in the nation."* This message is often delivered within the context of Wyoming's creation of an integrated statewide assessment, which is instructionally supportive and includes an embedded formative assessment component.

The lingering question however, is whether PAWS, as structured, was the best choice to meet the mandates of NCLB. In other words, it may be the best assessment in the nation because of its integrated capabilities, but was it too complicated and costly, for the modest goal of meeting federal mandates to receive Title I funding?

If the ultimate goal was to meet the minimum requirements of NCLB, as well as passing federal muster through the federal peer review process, a less complicated and costly assessment may have been a better choice. Especially less, given that

districts are already required to develop district assessment systems, which align to state content and performance standards.

Perhaps with the pending ESEA reauthorization, Wyoming can again decide whether it wishes to use an integrated statewide assessment. At a minimum, the experiences of the past will allow it to make a programmatic choice based on new federal guidelines. One thing is for certain however. PAWS in its current form, has experienced problems.

Not only was PAWS costly as a result of its integrated nature, the test administration was a failure for SY 2009-2010. WDE has taken steps to ensure Wyoming's interests are protected by negotiating with NCS Pearson, Inc. and encumbering funds until settlement negotiations are complete, but they are left in an awkward position with respect to their requested AYP waiver request for SY 2010. They are also taking a step backwards in terms of technology by administering the assessment next year via paper and pencil, which is another piece of evidence that speaks to the failure of the current contractor to properly administer the assessment.

Finally, given legislative actions with respect to recalibration, the State is left with using PAWS as an instrument to eventually measure districts' performance. Perhaps this is a productive occurrence, given that PAWS scores are not used for other high-stake purposes in the State. Using the statewide assessment, especially given associated costs, to measure how well districts are doing as part of a district accountability system, appears an appropriate use of PAWS. Given the tremendous expense of PAWS, it makes sense to use PAWS for other purposes.

AGENCY RESPONSE

Proficiency Assessments for Wyoming Students (PAWS)





Wyoming Department of Education

Dr. Jim McBride, Superintendent of Public Instruction
Hathaway Building, 2nd Floor, 2300 Capitol Avenue
Cheyenne, WY 82002-0050
Phone 307-777-7673 Fax 307-777-6234 Website www.k12.wy.us

MEMORANDUM

To: Management Audit Committee
Representative David R. Miller, Chairman
Senator John Hines, Vice Chairman

From: Dr. Jim McBride

Date: November 24, 2010

Re: Agency Response to Proficiency Assessments for Wyoming Students
Program Audit

INTRODUCTION

The Wyoming Department of Education (WDE) appreciates the time, effort and depth of analysis that went into the Proficiency Assessments for Wyoming Students (PAWS) Program Audit. The state assessment system is a multi-faceted, complex program which extends far beyond the WDE's Standards and Assessments Division. The WDE appreciates the courtesy and respect we received from the Legislative Services Office staff as evidence was collected and interviews conducted and the thoroughness of the staff's inquiry into the assessment program.

The WDE acknowledges that the online delivery of the state assessment in 2010 was a frustrating experience for all involved, making it difficult to separate the delivery failure from the soundness of the test design. The department also understands that the online delivery of the assessment created scheduling issues for districts as they moved students in and out of their computer labs to complete the various sections of the assessment, leading to the perception of a five week test for all students. Steps have been taken to alleviate assessment duration and the multifaceted nature of PAWS for the 2011 administration. PAWS will be delivered in a paper/pencil mode over five half-day sessions for students in grades 3, 6, 7, and 10 to complete reading, writing and mathematics and six half-day sessions for students in 4, 8, and 11 that must also complete the science assessment.

The WDE is committed to the importance of assessment and its appropriate place in the standards, instruction and assessment cycle. The department shall continue to provide professional development in the areas of understanding how to use PAWS assessment results in conjunction with other district and classroom assessment results for school program and instructional improvement, as well as an understanding of the related current and future federal and state assessment requirements.

RESPONSES:

Chapter 2

Finding 1: PAWS is not being used as originally envisioned. For example, it is not being fully used as an integrated statewide assessment that is instructionally supportive, includes multiple summative assessments, and includes a formative assessment component. The integrated approach is also extremely costly.

Recommendation 1:

WDE should take action to determine how often districts use the “traffic-signal” reports and its instructional suggestions offered as part of an instructionally supportive statewide assessment. In addition, WDE should consider developing a reporting mechanism for districts to use in order to report this information, as well as identify ways that districts actually use the reports.

Response: Agree

The WDE will continue to support increased teacher and administrator participation in “traffic-signal” related activities. As a part of these activities, the WDE will survey both teachers and administrators on an annual basis to determine the degree to which they have utilized the “traffic-signal” reports in terms of student placement, instructional improvements and building/district initiatives. The WDE is willing to work with a legislative committee to develop a more formal data collection system.

Recommendation 2:

WDE should continue to work with districts to tie in existing use of formative assessments, more specifically NWEA MAP, with PAWS. These assessments throughout the year can provide important information to districts and schools with respect to instructional changes.

Response: Agree

WDE and NWEA are continually enhancing their partnership. MAP testing is now administered in all 48 districts, which makes Wyoming, NWEA’s first statewide adoption for MAP. Numerous discussions and meetings have resulted in an agreement in which, NWEA is working with the WDE’s Mosaic, a consortium of Wyoming educators, to provide additional professional development to assist Wyoming school districts to use their MAP results toward improving student instruction.

WDE also requires MAP results for the Bridges Summer School grant. NWEA is working with WDE at no cost to provide districts MAP results directly to the WDE. WDE is also currently in the process of developing a statewide data analytic tool that will allow districts the ability to compare their students MAP results to their PAWS results along with other critical data elements, for example: attendance rates, disciplinary actions, district and classroom assessments.

Recommendation 3:

At a minimum, WDE should track the cost-effectiveness of PAWS by comparing the dollars expended for professional services against the use of instructional components of the integrated statewide assessment.

Response: Agree

The WDE will work with a group designated legislators to arrive at a meaningful analysis of the cost effectiveness of PAWS by comparing dollars expended for professional services against the use of the instructional components of the integrated statewide assessment.

Finding 2: The Wyoming Comprehensive Assessment System (WyCAS) could have been modified to meet the requirements of NCLB, according to past and current WDE officials.

Following the direction provided by the 2003 Wyoming Statewide Task Force on Student Assessment and Educational Accountability (Task Force), the WDE was directed to put into place an assessment system which surpassed the reporting capabilities of WyCAS. Thus, at the time of the transition from WyCAS to PAWS, work required to have modified the WyCAS was comparable to the work necessary to develop a new assessment (PAWS) which, in addition, was aligned to the newly implemented Wyoming Content and Performance Standards.

Finding 3: Wyoming does not require that PAWS scores be used as high-stakes criteria related to graduation requirements, college entrance requirements, college scholarships, or pay for performance of teachers or school administrators.**Recommendation 1:**

Pending reauthorization of ESEA, WDE should work with the Wyoming Board of Education to determine if the statewide assessment and related reports should be used for other high-stakes purposes such as graduation requirements, college scholarship criteria, or pay-for-performance for teachers and administrators. Based on the conclusions, should then be factored into the next RFP.

Response: Agree

The WDE will continue to be an active participant in all discussions regarding whether or not the statewide assessment and related reports should be used for other high stakes purposes. These decisions will be far reaching and the WDE is supportive of these decisions being made with representation from all stakeholders, along with guidance from experts in the field. Final decisions will be factored into the requirements for the next statewide assessment RFP.

Chapter 3

Finding 1: Local districts are critical of PAWS, but may be lashing out because of test administration problems.

The WDE is aware of the negative impact on public confidence and perception of PAWS as a result of the online platform failure in 2010. The WDE has made and will continue to make every effort to assure the public of the soundness of the test design of PAWS versus issues with the online delivery software.

Finding 2: There is a desire among some districts to use formative assessments in lieu of PAWS. However, such use would be contrary to NCLB requirements.

Recommendation 1:

The WDE should develop and administer training on NCLB requirements, as well as pending ESEA reauthorization with respect to differences between various assessments as they apply to NCLB requirements. This should be applied to additional training on the PAWS program from RFP to contract.

Response: Agree

WDE agrees that training on the current NCLB requirements and the pending ESEA reauthorization will assist districts in understanding the federal assessment requirements. WDE will continue to provide trainings specific to the PAWS program to support districts in the use of the assessment results to meet the primary purpose of the state assessment, “improvement of teaching and learning in schools and fostering school program improvement”(§ 21-2-304 (a)(v)).

In addition, early each fall, the WDE hosts an All Superintendent’s Day. During this meeting both new and veteran superintendants are informed of WDE policies and procedures regarding all aspects of the agency. Plans will be put into place to better address the relationship between ESEA and various aspects of the PAWS program during this meeting.

Chapter 4

Finding 1: NCS Pearson, Inc., encountered problems and performed poorly during the 2009-2010 school year as it carried out contractual obligations.

Recommendation 1:

WDE should continue seeking the maximum amount of liquidated and default damages based on problems encountered by NCS Pearson, Inc. as it carried out contractual obligations.

Response: Agree

The WDE agrees that the efforts to seek the maximum amount of liquidated and default damages from Pearson for the problems encountered as it carried out contractual obligations, should be continued until the state is compensated for the 2010 PAWS administration failures.

Recommendation 2:

WDE should continue working with the Department with respect to its AYP waiver. It should also ensure that the recommendations from NCIEA's report are fully implemented.

Response: Already in place

The WDE received formal notification from the United States Department of Education on November 16, 2010 approving the AYP waiver request.

The NCIEA report (7/27/10) contained the following recommendations:

1. All reports that were originally planned should be produced and distributed. Without evidence to the contrary, it should be assumed that the reports provide a valid estimate of student achievement.
2. If it is known that a student was affected by administration problems, and the achievement of the student on PAWS was inconsistent with other information about the student, the PAWS result likely should be discarded. Note, however, that this recommendation is consistent with all good testing practice; any time an individual test result is not consistent with other known information about a student's achievement level, the other information should take higher priority in judging the student.
3. WDE and Pearson should make an offer to any district that feels it can identify subgroups that were clearly affected and clearly not affected to conduct the kind of impact study we were unable to do under the time constraints provided by this contract.

With respect to NCIEA Recommendation One, the WDE publicly announced on 7/30/10 the decision to publish all 2010 PAWS results. Subsequently, the WDE and Pearson publicly released all 2010 PAWS results online on both PearsonAccess (10/8/10) and the WDE Fusion site (11/12/10).

Regarding NCIEA Recommendation Two, an interpretive guide written specifically to address potential testing anomalies and how to judge a student's performance on PAWS was posted to PearsonAccess (10/8/10). Superintendent Memo #2010-151 went out informing districts of this recommendation that should be used when interpreting a student's PAWS 2010 test result.

Lastly, regarding NCIEA Recommendation Three, on August 27, 2010, the WDE invited all districts to submit their student level information regarding any subgroups of students that were clearly affected and clearly not affected. The WDE and Pearson analyzed all the districts' submissions and were unable to detect any score anomalies at a district level. Results were provided back to those districts which participated.

Finding 2: The PAWS writing component is not required by NCLB.

Recommendation 1:

WDE should continue to address concerns with respect to the writing portion of PAWS. It should also determine whether eliminating the writing portion in the future, would diminish the importance of writing within the districts. It should work with the local school districts to obtain their feedback as well on this option, as well as modifications to remove the "drafting" portion of the component.

Response: Agree

The PAWS includes a writing assessment as required by §21-2-04 (a)(v) (B). Additionally, the Statewide Assessment Taskforce recommended writing be included as part of PAWS to ensure that writing was a skill that would be taught on a regular basis and that the writing assessment mirror the Wyoming Language Arts Content and Performance Standards which states: "Students use the writing process and use appropriate strategies to write a variety of expressive and expository pieces." Therefore, the assessment includes an expository prompt and an expressive prompt and students use the writing process (a drafting session and a final copy session) for each prompt.

WDE would support the opportunity to discuss various options/modifications regarding the writing portion of PAWS with districts.

Chapter 5

Finding 1: The federal peer review process has frustrated WDE, and appears to be inconsistently applied across all states.

Recommendation 1:

The Management Audit Committee may wish to consider requesting that the U.S. Senate Committee on Health, Education, Labor, and Pensions request that the Government Accountability Office (GAO) conduct an audit of the federal peer review process pending reauthorization of ESEA.

Response: N/A

Chapter 6

Finding 1: Elementary and Secondary Education Act (ESEA) will surely impact PAWS.

Recommendation 1:

WDE should work with the current vendor to correct discrepancies with respect to PAWS administration during Spring 2010. In addition, it should not make drastic changes to the current assessment, until ESEA has been reauthorized.

Response: Agree

The PAWS 2010 administration had many issues in respect to the online delivery. The WDE conducted a district survey to determine if the districts would like to continue with an online assessment or return to a paper/pencil mode for the assessment. As a result, PAWS 2011 will be administered paper/pencil. The WDE is working closely with the assessment vendor to track each deliverable from development through delivery for the purpose of ensuring on time delivery of all assessment materials to the districts.

In order to minimize impact on districts, the WDE agrees that the future of PAWS should be considered after reauthorization.

Recommendation 2:

WDE should study the activities at the federal, state, and local levels with respect to the direction of ESEA reauthorization. Information related to the blueprint, RTT funding, as well as the development and funding of various consortia can be used to position itself for eventual ESEA reauthorization.

Response: Agree

WDE will continue to monitor the United States Department of Education (USED) website and attend technical assistance meetings provided by the USED, as well as Council of Chief State School Officer meetings for the purpose of staying current with all information regarding ESEA reauthorization, the blueprint, RTT funding and the activities of the various assessment consortiums.

Chapter 7

Finding 1: Wyoming and other states are not making changes to statewide assessments until after reauthorization of ESEA.

The WDE agrees that many states are waiting for reauthorization prior to making significant changes to their state assessment systems.

APPENDICES

Proficiency Assessments for Wyoming Students (PAWS)



APPENDIX A:

District Superintendent Survey Comments

Comments Section Related to Individual Questions

Question 1	Overall, what is your opinion of PAWS?
Comment Section	
We are not opposed to reasonable accountability measures. As a state test PAWS is a good test but the implementation of it has been problematic. Every year there have been different administrative issues.	
The length of the test is too long. The level of difficulty is inconsistent with grade levels. The writing prompts are poor. The results are not timely. Each year the test changes significantly thus inhibiting year to year comparisons. There is no "teeth" to PAWS thus as students get older their seriousness of the test declines. The writing portion of the test is too subjective. Results don't portray student's ability.	
PAWS is too long, should just focus on reading and math and in its present format does not provide meaningful student assessment data. It needs to be replaced or drastically revised.	
The original intent of PAWS was to provide an assessment tool and platform that would provide educators with formative data that could, in turn be used to inform instruction. This was a good idea; however, the concept never came to fruition. Consequently, PAWS became a disappointment that was exacerbated by a poor delivery system.	
With how much PAWS changes from year to year I don't know how much reliability the test holds. 08-09 had writing questions that were extremely hard for the kids to relate to. Last year was much better, however, there's that reliability issue again. I think that we need to get rid of PAWS and go to MAP's.	
PAWS provides a one-shot measure at the proficiency of students in reading, writing, math, and science. Wyoming has set high cut scores that need more transparency for both the education community and the public at large. Wyoming must know how difficult their state assessment is in relation to other states, define what "proficiency" means in terms of the cut score process, and clearly communicate that process to the public before making judgments on the achievement level of Wyoming students. The assessment is too long and we have experienced too many changes. After last year's PAWS administration, the assessment has serious credibility issues that will be difficult to overcome. Pearson has been unable to deliver on its promise to properly assess Wyoming students. Additionally, we would suggest strong consideration for a growth model.	
Way to long, very poorly delivered (tech platform untested) results are not timely, the ability to disaggregate data to inform students instruction is very poor, the district spent 50 k to contract with a company to help us use the data appropriately, now the data we have is invalid.	
Make shorter	
Last year was disappointing	
I'm not impressed with PAWS. It can't be used as an assessment tool, because we get the result back during the summer and this year in October. How can a teacher use this information as instructional tools as the students have moved to the next grade? It is a good snap shot in time, but nothing else.	
I think having summative test such as PAWS is helpful for ... However the technology issues with it the inconsistency of it questions and the length of the assessment seems to interfere with the validity of the results	
-Results don't come into districts at times when teachers can check for score validity. -Test has changed over the years. (Example: We used to teach and test the 6 traits of writing, now (as of last year) it is the 4 traits. -Computer issues should have been taken care of before the testing began. -Community has very little faith in test as it	

currently is.
Paws is not a tool that is significantly helpful to teachers or schools. Testing parameters vary significantly, results are not timely, questions of validity and reliability abound. The ideas behind the traffic signal reporting are good but are not timely for teachers
I think that it keeps changing and being adjusted to such an extent that very little value can be derived from the results. The return data can be used for a very short period of time because it comes so long after the testing date. It is outdated, because of its changing nature, it is hard to credibly use it as a longitudinal expression so a students learning.
We have been told that PAWS meets the requirements of NCLB and that with the mandates set by the feds along with the states "accountability plan" we are stuck with it. I wish we could get reliable results sooner. Wyoming has many educational experts and I believe if given the chance we could come up with something better that actually assesses both proficiency and growth and provides good information to enhance a mode of continuous improvement.
The test is prompted by federal mandate. It has been a work in progress where many of the problems have been fixed. However, it is still to long and defeating to students. There are still many problems with the test. The technical problems are many and have many instances and horror stories related to the problems. These problems reduce the credibility of the test and Pearson. The test is plagued with problems of which credibility with students, teachers and parents is very problematic. There is a need for a reliable and valid assessment so programming and triangulation of problems is possible to accomplish both. We threw out Wycas when it was refined to the point of making a difference and now we risk the same by throwing PAWS out as well. We need to form best practices and draw on expertise to provide a system for Wyoming that would work for multiple purposes and audiences it must serve.
PAWS serves its intended use in that it meets NCLB requirements, is designed to attempt to provide data that measures a deeper cognitive depth than other large scale assessments do, and provides data to inform classroom instruction in a broad sense. However, there are several areas of concern with the PAWS. These issues stem largely from decisions made by the Wyoming Department of Education on the administration of the assessment and have plagued this assessment from the onset. The original task force recommended no computerized administration for two years until the assessment was established and a thorough technology capacity study was complete. If followed, this recommendation would have allowed the state to have two years worth of testing data before introducing a potential for problems with reliability of the test based on technology issues in administration of the assessment. The recommendation also provided for time to assure adequate capacity of the state's and districts' computer systems and training of district staff. This recommendation was not followed and the first administration included a computer based component. Some will say it was fine and there was no impact on student results while others would disagree. Regardless, we have no baseline data to make that determination. Five years later, we come full circle to the 2010 administration of PAWS which was administered using a new software platform that was not properly field tested and caused multiple problems. Within a month of the close of the testing window and, in my opinion, before a thorough investigation of the issues – the WDE decided the 2011 test will be completely paper/pencil. Reverting back to teaching elementary students to take tests by filling in bubble sheets with #2 lead pencil (fill in completely but don't make any marks outside the bubble and don't let your wrist rest on the answer sheet and smear your answers) is an overreaction and not good role modeling of 21 st Century problem solving skills.
"I think it takes far too much time. MAP testing is far more usable and takes far less time to administer. The on-line portion is well-intended but technically not acceptable. We need to get rid of PAWS. We do not get kid's best effort because it is way too long. I'm not sure getting rid of PAWS is wise until we know how this would be accepted at the Federal level. What would replace PAWS?
PAWS is a test that tries to cover too many aspects of assessment. While the original intent was for the test to be a criterion referenced test it is now norm referenced for the sole purpose of comparing districts. For example it too many students get questions wrong it is thrown out. Also cut scores are established after the students take the test rather than setting the standard prior to the test.
We have experienced issues each and every year it has been in place. Issues of incorrect scoring , mix ups on testing materials, students scored or coded that should have not been counted or vice versa, special ed issues,

and of course the major issues this past spring with technology platform were just plain ridiculous. The time to take the test is too long for younger students. We should not test science until 8th grade and even then it is not really a science measure only a vocabulary or comprehension assessment. Stick to reading, writing and math only. WE can better manage science and other core subjects. The time to get the results is crazy. Waiting for 4-6 months and then have to spend two more months getting the scores accurate does not help in using the test as a feedback tool for staff. Probably the most important issue with PAWS is it has no incentive for a student to perform well on the test. It does not count for anything but AYP. NO incentives like Hathaway Success Scholarship ranking, grade advancement, college acceptance, etc. Just does not have value for students or staff and in many cases school personal and patrons in general.

Currently my opinion is very low. There was a time that the aspect of what PAWS brought for student assessment and results was very exciting. Still the issues that were always associated with the Pearson product remained. Extremely slow turnaround time and there always seemed to be a question, was the assessment articulated or presented correctly for students. In short most members of the education community just didn't trust the product or the state assessment. The latest issues weren't any different than previous issues it just that it was a statewide problem instead of a few district issues. It also seems that the writing component of PAWS and the assessment are problems. Check the mandates of NCLB and ESEA and writing is a component not a standalone facet. The time and cost to assess writing is huge. I believe that this can be completed under different approaches.

Not reliable, problem plagued administration and or reporting problems every year. Results too late to be used for anything in the classroom. Test has definitive validity and reliability issues due to factors such as test design and administration and lack of accountability as student level. We need accuracy and we need results in a timely fashion.

The idea behind a state-wide test that allows us to measure students' achievement is a good thing. The PAWS assessment has some serious flaws. First is the amount of time it takes to administer. We are spending far too much time assessing as opposed to teaching. The items that we're testing are appropriate. The format of the test leaves much to be desired. The biggest problem is that the information on how students have done has been so delayed in getting back to the district that it offers little to no value. By contrast the MAP test, which many districts are using, provides immediate feedback which allows teachers to adjust to the needs to the students. The move back to paper and pencil was not well thought out. In today's world students are more comfortable with the technology platform. We need to find a way to make that platform work.

It is a well designed test for its intended purpose. It measures a few big, important skills in each content area, which helps us implement our Body of Evidence System. This is our opinion of the test itself. We need a contractor who can deliver the test without glitches.

PAWS is better than WYCAS which it replaced. It is not the best test to measure student growth but does give the district one measure, one snapshot in time as to how a student or group of students is progressing. I also believe the technical problems encountered this year should not be a reason for PAWS to be scrubbed or for us to return to paper and pencil testing.

PAWS has many merits as a state test. It represents the standard input of the state and its' instructors. The test does provide great feedback for schools and individuals related to areas of concern and areas where we are strong. The online format was a positive move for the state even though some problems developed. The problems were that the test is too long and the results were provided late after the test. The results for the test format can't be immediate, but there is a need for quicker results. The level of rigor for this test is lower than we felt it should be. The frustrations of continued changes did not provide a quality evaluation of the data. In 5 years, the targets and the procedures continued to change. We need consistency.

First of all let me say, in relationship to this question and all additional applicable questions, that our district is not at all opposed to the establishment of accountability for the determination of student performance. I believe that the utilization of state and local resources requires reasonable expectations for student performance. Having said that, I believe that the WyCAS/PAWS creates a definition and expectation for student performance. Performance, PAWS, defines static performance at the time of its administration. It is a high stakes test that essentially requires Wyoming districts to spend a great deal of time and effort to teach and otherwise prepare Wyoming students to do

well on PAWS. This effort also essentially strips away time and focus from other curricular programs that are equally important but not tested on PAWS; we can all be thankful for that, as I am not sure what the result would be if all curricular areas were high stakes tested and reported. So, PAWS does require a maximum amount of effort and resources to prepare our students to perform well on the state assessment; The results obtained from the PAAWS are actually of little use to districts in determining curriculum and instruction adjustments that would be helpful to the overall benefit of our students and their continued achievement. Finally, as it concerns this particular questions, prior to NCLB/State Assessments our students attended school for 175 and during that 175 day period, schools applied their resources and skills to prepare students to be as successful as possible in all of the areas of the curriculum. It is my estimation that we have lost many days of that instructional opportunity because schools are no focused on attempting to do everything possible to meet the proficiency areas tested. To, that narrow focus has actually diminished the total amount of days available for all curriculum areas; I suspect that of our 175-day calendar that we have focused 20 school days in preparation for and taking the PAWS test. If twenty such days of preparation for PAWS is the norm for most Wyoming district, it would be my position that we have lost a lot more educational opportunity because of PAWS than we have gained because of PAWS.

PAWS has evolved each and every year. The data we get is untimely and many times useless. I understand why we do PAWS but we have much better assessments in place (MAP) that actually affect instruction.

Feel we need something different.

Paws is a potentially powerful way to provide one piece of information about student progress on some of the Wyoming Standards in Language Arts, Math, and Science. Unfortunately the results are not returned in a timely manner and educators have little confidence in the writing portion of the assessment. The writing results have not been stable enough to provide valid information about student progress from year to year. So there is almost no way to judge whether curriculum or instructional changes that are made one year will have an impact the next year. This year there was an attempt to correlate writing with other aspects of language arts. It is difficult to know if those changes will make it possible to believe that the writing scores are accurate, but they have the potential to create a more consistent, accurate score. The length of testing time is not an issue for me. I am in favor of keeping PAWS, while taking necessary steps to insure that the assessments are valid and reliable, that educators have confidence in the results, and that the results are returned as quickly as possible.

It's time consuming, probably little or no correlation with other states. Provides a "snap shot" in time of a classes progress, nut not necessarily of individual student progress.

I don't believe the PAWS scores will reflect an accurate measurement of our students' abilities due to the serious issues with Pearson test software during the complete cycle of testing. The issues may have caused variances in the final test scores that would not have happened if the software had functioned correctly. Our students were kicked out several times and needed to be logged back in after the test administrators reset them. Some kids needed to log back in up to 15 or 20 times for one test. Even those who did not have challenges with the software were disrupted by the students sitting next to them who were raising their hands to alert us of their challenges. I believe our students were frustrated and their attitudes may have carried over to the pencil and paper sessions skewing those results, too. The idea behind PAWS is excellent. I would love to see the NCLB assessment tied to our standards and benchmarks, and to be efficient enough to be used both formatively and summatively. In its present form PAWS has not performed as intended. The results are not timely and it takes too much time for elementary students to complete the test. The upfront management of the test is also cumbersome. The State of Wyoming can do much better than PAWS.

PAWS serves purposes beyond the child and the classroom which don't always connect to educations hierarchy purposes. It is basically a pass/fail test which does not consider growth or inform instruction related to growth. By using cut scores the assessment masks the reality of student achievement. A large group of students can be 1 point below a cut score and then be compared to a large group of students 1 point above that cut score which then inaccurately represents both groups as failing or passing. The reality is on the continuum of learning they are in similar places.

The original design was sound providing for both formative and summative assessment opportunities. When the original design was abandon best practices in assessment were abandoned. Over the life of PAWS there have been no two years where a consistent administration has occurred, some variable has changed such as one opportunity, then two opportunities removing extended response items from the math and reading sections and

adding a language arts section. There are also issues using multiple choice items to measure writing, as it is not authentic and requires only passive not active response from the student.
Overall, I believe the assessment is a good one. I'm concerned about moving backwards to all paper and pencil.
Overall, I think PAWS is a well designed test that gives a snapshot of student performance. However, I do not believe we should be testing writing. The length of the test for the younger students is problematic.
My overall opinion is that PAWS is a meaningless test with few instructional implications. It is too low level and because of the numerous changes, has created little buy in from our district and staff. This is especially pronounced at the high school level.
PAWS is not what it is intended to be. It is not a valid and reliable measure of individual student achievement for our schools or our district. Further, it is not a timely fashion to gauge student learning and adjust classroom instruction.
PAWS too long; way too long. PAWS results are so slow coming in that their value is practically nil, particularly for planning purposes. PAWS seems to have a changing target for instance when writers became proficient, proficiency levels were raised.
The intent of PAWS is exactly what WDE intends for it to do – measure student proficiency to monitor student performance and to improve instructional delivery. Unfortunately, there are a number of obstacles that prevent appropriate (1) administration and delivery and (2) timely access to data. Our district recognized that the increase need for technology has driven us to examine our technology needs. There has been much concern regarding writing prompts. Lesley Wangberg is cognizant of the concern, and it seems as if that is being addressed. Affording high school students to use the computer with writing prompts is an excellent opportunity for students to effectively express themselves. The pressure of schools' performing to make AYP, with published comparative data, poses a huge amount of stress on the staff. For example, many teachers seem to focus on explicit genre of writing and the opportunity for students to initiate creative writing seems stifled.
Because of the problems with the testing and reporting, I believe many students, parents, and staff view the test as a disaster. In my opinion, the test does not do what it was intended to do.
Too long of an assessment. Problems with technology piece. Question validity.
In our district the greatest concern has been with the technology and the writing assessment. The online version had many bugs that prevented students from completing the assessment in a timely manner. It was also a cause of frustration for students and Staff. The second concern is in the area of writing. The scores in this section of the assessment have been up and down with no real explanation other than the makeup of the assessment and the scoring. I do not think the other areas of the assessment are a problem. The assessment is fine with the exception of the two areas above.
I believe the PAWS to be a fundamentally sound assessment, but despise the platform. It is too time consuming and students become test fatigued.
It is a poor statewide assessment of student achievement and has been since its inception. It has some striking deficiencies beginning with a lack of communication with practitioners, parents, and community leaders about the purposes and uses of statewide assessment information. It has been plagued by a series of administration and reporting difficulties. The administration directions change annually, leading to differing results that are difficult to compare from year to year. The reporting has been especially poor. In two to four, the results has been significantly delayed making them essentially useless. Each year, parents complain that the items are poorly crafted, especially the writing prompts. The scoring rubrics have also been problematic and changed from year to year, making comparisons meaningless. This is a very poor assessment.
The concept of the test is a great idea and the accountability school districts are essential, but this testing system has flaws that need to be addressed. Because students are not held accountable there are students that just mark answers or just refuse to take the test at all. This is one of the major complaints from teachers and administrators. Also the amount of teaching time lost is another complaint. To allow students the opportunity for quality testing time, it takes about 2 weeks of instruction time. Their set testing procedures do not follow the research based testing norms. Tests are too long and redundant.
The validity of the test is in question. The changes that occur year after year have made it impossible to take the test as well as the dept. in charge of the assessment.

Question 2	What is your opinion of PAWS as a valid and reliable instrument to meet the mandates of No Child Left Behind (NCLB)?
Comment Section	
To many changes over time, too many variables. Only valuable as broad indicator, not reliable for high stakes individual educational decisions	
N/A	
PAWS is neither reliable or valid.	
N/A	
Only comments I have are the as on question one.	
We believe the administration difficulties in 2010 will produce results that are not valid or reliable. We also have doubts about the quality of scoring of the open-ended items.	
It has potential	
When tech pieces work, would be more of each.	
N/A	
Good definition, but PAWS had so many problems with the administration of the test it leaves doubts in everyone minds that the scores are not accurate.	
It seems as though this has not consistent over the years with the test One would expect that it would.	
As stated in #1, test has changed over the years – length in time when getting results back to district and to staff. Very little time for districts to be able to go over the results to determine if scores are valid and reliable.	
When state wide results vary as widely as they do with PAWS then we do not have a valid assessment tool. The number of scoring errors have also made it not reliable.	
See Above	
Reliability is an issue because the test changes yearly. One year a pilot, this last year a fiasco. We can't determine trends because we haven't taken it for long enough. If it is completely changed again, we will lose the potential for reliability again.	
With changes in administration procedures from year to year and the changes in the format and content, the test has limited reliability. There has never been conclusive correlations completed to determine the validity in terms of correspondence of the state.	
I answered this question based on the fact PAWS is approved by the federal government to meet the mandates of NCLB, including having processes in place to assure adequate levels of reliability and validity. If I were asked whether I perceive the PAWS as a	
Valid and reliable with the exception of writing. The test items may not align with 21st Century competencies—real world essential learning's.	
While the questions do relate to standards in a critical way the format and the content of the test change frequently therefore making it not reliable.	
N/A	
N/A	
Test scores have shown wide swings in the same group of student from one year to the next- especially in writing; also, English seems to be inconsistent and non-reliable.	
The test is valid but excessive. As I stated above it is consuming far too much of our instructional time.	
The PAWS test meets standards for reliability and validity that state tests are expected to meet. Of course, poor delivery of the test can negatively impact the validity of test scores.	
N/A	
The lack of consistency through the years provide our decision on reliability. It was not the content. If consistency was there for the test, the results would have been valid.	
I was not and have not been a participant in the development, analysis, or as an over-viewer of the construction of the PAWS assessment. It would be my lay opinion that the PAWS instrument is reasonably reliable and valid.	

N/A
Validity is questionable as reports of PAWS assessing standards that are not grade level have been heard.
Since science is only administered at grades 4, 8, 11 and we haven't had 4 years of administration I wonder if we can know about reliability and validity. 3rd and 5th grade writing have been questioned each year the PAWS has been administered.
(blank)
N/A
The reliability is probably good within a particular grade level, but there is very little consistency between different grade levels. The proficiency levels vary between grades. The proficiency levels vary greatly from one grade to the next when follow
Instead of focusing on meeting mandates of NCLB we should be focusing on conducting how well our students are learning what schools, parents and communities expect. Of course this would mean serving our children in lieu of serving the federal government.
As stated earlier, Over the life of PAWS there have been no two years where a consistent administration has occurred, some variable has changed such as one opportunity, then two opportunities removing extended response items from the math and reading sect
N/A
N/A
The test does not truly measure Wyoming standards, nor does it tell us if students are career or college ready.
PAWS have never been a consistent measuring instrument. The format has changed almost yearly.
N/A
Concerns regarding technical issues were cause for question during Spring 2010. Other than technology concerns, PAWS is aligned to the standards of measure in all content areas. Some of the writing prompts do not always seem appropriate at various grade l
When the state of Wyoming has to hire a third party (Center for Assessment) to assure, validate, and reliability, we are not applying tax dollars wisely. Validity has to have several years of good documentation. Example: ACT
For 09-10 length and problems with technology. I don't believe results will be valid or reliable.
I would have marked valid and reliable except for the area of writing and for when students experienced difficulty with the technology.
It seems as if cut scores and test items are changed too much to be fully valid or reliable.
Of course, these are technical indices are the PAWS TAC never sends the results of these analyses to superintendents or other interested parties. Our principles and faculty believe this assessment to be deeply flawed both technically and substantively.
The reliability and validity are questionable because of the time factor, testing issues, and the continuous changing of the tests and testing delivery.

Question 3	How satisfied are you with the process used to make recommendations for developing a new statewide assessment (PAWS) after the passage of No Child Left Behind (NCLB)? More specifically, the process used by the Wyoming Statewide Task Force on Student Assessment and Educational Accountability and the Commission on Instructionally Supportive Assessment, which ultimately served as the basis for an initial RFP?
Comment Section	
There was an attempt to involve many constituents.	
Not sure we had any input.	
N/A	
We can't forget that Wycas was the initial assessment instrument that was created in response to NCLB -not PAWS. Wycas was far superior in validity and reliability. PAWS was created as a political move from a new state superintendent, Trent Blankenship.	
N/A	
The task force recommendations on cut scores were not always followed. When this happens, a transparent	

explanation should follow.
Not present for the process.
N/A
N/A
The process was good but administration of the test is greatly lacking.
I don't see that PAWS is any different than WYCAS generally and the time, effort, and state funds do not seem to be worth the effort.
It is the changes to PAWS that have occurred in the years after NCLB that I don't believe we have had much input on. In the beginning there were committees which included educators, but each year changes have been made without our input.
I was not involved in the process personally. We did have staff members who participated
was not here during the process- so no comment
The 2002 committee did some really great work. Those recommendations were only partially fulfilled. The test we take now has no resemblance to what that committee envisioned. The state needs a growth model assessment. One that is meaningful to all students.
The process occurred with surprisingly limited input from state educators. In some cases teachers were asked to work outside of their content area or grade span expertise to develop and or/ modify stems and distracters. This was develop, in the perception
I served on the original task force with then Senator Job and Representative Wasserburger, so I declare some bias in my response. I know firsthand the time and training provided to task force members during this initial work. The task force had legal co
Due to the limited input from districts, we are not all that satisfied with the process that was used. How satisfied were you with PAWS last spring?
I was not here when this was developed but those who were state that there was not enough involvement from teachers and staff.
Don't have an opinion. I was not involved with the process. Seemed pretty well decided long before input was taken so our input seemed a bit like window dressing. The WDE likes to say they ask for input and in many cases they do a good job of taking in
I know that the attempt was to develop an assessment that also aligned with the requirements of BOE legislation. I don't think that at the time of the decision for PAWS the aspects of BOE and the Hathaway Curriculum requirements had been developed. There
The development of PAWS was an improvement over the old WYCAS but it didn't go far enough. NWEA was in the running for the contract but it seemed WDE didn't want to deal with USED and find a way to make it work. NWEA has been reliable, and gets results.
Questions are ask but it feels like lip service
The process was objective, with representation of stakeholders, technical advisors. Annette Bohling's leadership was well intentioned and well executed.
The process was fine. Many stakeholders were involved and the test was designed. However, it was to be used as a formative assessment rather than summative and the way it has turned out doesn't seem to work exceptionally well for either, but the process a
We think there needed to be more input by groups. Things were controlled pretty tightly.
N/A
Process was too driven by WDE philosophers.
N/A
There was an initial effort to make the new assessment sensitive enough to provide teachers with sufficient data to improve instruction. There were attempts to limit the amount of time students would need to complete the assessments. Concepts or skills
(blank)
N/A
WyCas was much better.
Then the intent was on serving Wyoming's students. Nonetheless a decade later, this intent actualized into

something different.
Communication to districts and transparency has been poor causing the actions of the task force to appear suspect.
N/A
N/A
This was a very secretive group with little input from districts.
N/A
You appear willing to listen, but there is so little evidence that you've paid attention to what we've said. An example Reading test too long -you made it longer.
WDE has a strong task force with expert individuals as James Popham, Mike Flicek, etc. Districts must recognize, respect, and support their expertise to make appropriate decisions and rely on the process to meet the needs to provide student information to
If they keep the process simple; back in the hands of students and teachers so instructional modification can be made, it can work.
But not pleased with the end result.
I believe the task force is very knowledgeable and focused on providing instructionally based assessments for students.
I was not in Wyoming at the time.
Our representative to the task force felt that her opinion of many other educators from the state were not listened to during the entire development process. They have met infrequently and have not been consulted during this recent debacle of the cache serve issue.
We don't feel that they listen to districts' recommendations. Their testing procedures do not follow education based research on best testing practices.

Question 4	How familiar/satisfied are you with the PAWS administration process from 2005 to present?
Comment Section	
There have been concerns every year. The double testing was problematic.	
N/A	
Again is long, and test administration changes each year. The tech component this year was disaster.	
N/A	
The trainings we have with Pearson there is great. However, sometimes we are given conflicting information from training -testing time.	
School districts have been faced with yearly changes leading to yearly frustrations. The 2010 administration was particularly challenging and was unacceptable.	
Very poor in 2010	
N/A	
N/A	
To not have done a beta test on the administration of the test was a great mistake. With the amount of money we are paying Person they need to get it right.	
Although we had limited issues this past year with the online portion of the test. That has not always been the case. Many districts around the state experienced much more difficulties.	
I had staff attend the PAWS administration training last year. One group went to the Casper site and another group went to the Thermopolis site. After the testing window started I had several staff confused over an administration aspect of the PAWS.	
Paper Pencil was fine, computer delivery had issues. The whole process takes too long	
N/A	
Perhaps the administration between Pearson, the state and the schools is something I should know but it hasn't become a priority for me. On staff we have a trained PAWS coordinator. The test is administered well with the district.	

N/A
Please see comments to question #1
2010 was obviously a disaster, one more year like that would be unfortunate for Wyoming student assessment
The obvious technical problems of last year are the basis for my answer.
Takes away from teaching core teaching.
The actual application of the assessment seems to be a work in progress with new software being developed in an effort to speed the student results return expectation. But the process of the paperwork to return for scoring and the precise packing methods
Test results are too late to be useful for corrective action. The test platform have consistently had problems, 2010 was the worst.
Members of my staff are very familiar with the process. There were problems last year in the fall but the concerns were addressed during the year. Those concerns should have been addressed in the very beginning of the tests in 2005.
Harcourt/Pearson's project management approach was placed at the school level. This caused district personnel to fix problems and create our own solutions because Harcourt/Pearson underestimated the scope of a project in a district our size. This was con
Pearson has worked with our district to be sure the implementation goes well. This year we received the tutorials about 3 weeks in advance of the test window. We loaded the tutorials and found they did not work. Pearson worked with us for two days and the
If yes, how satisfied are you with the PAWS administration process from 2005 to present?
I am somewhat familiar; I rely on our principals and technology folks to properly prepare our staff and students for the administration of PAWS It takes our junior students 3 hours to take the ACT test; it takes our 3rd grade students 14 hours to complete
It has improved
N/A
This year was filled with glitches especially as related to on-line components. But signing students up, preparing the rosters, and getting the entire tests ready to administer has been a nightmare in previous years as well. This year was the worst.
(blank)
N/A
The process to set up the online assessment was time consuming.
N/A
As stated earlier, Over the life of PAWS there have been no two years where a consistent administration has occurred, some variable has changed such as one opportunity, then two opportunities removing extended response items from the math and reading sect
N/A
I was pleased with the transition to online until the platform problems last year. Directions to schools are thorough. Customer services and clarity of documents leave much to be desired.
As a Wyoming high school principle from 2001 - 2009, I have been very frustrated with the administration of PAWS. It placed to much burden on the students (length, frustration with technical delivery.) As a superintendent in 2009-10, the technical problem
N/A
During this year, it would have been appropriate for Pearson to have notified WDE regarding the technology platform systems change from what was tested in the fall (2009) and used in the spring (2010). Pearson should have addressed the problem at the beginning
Too many flaws in the system; Our students are more accountable to them then they are have been to students in Wyoming.
I believe it was somewhat successful and that they have worked with the state to make the assessment better and more error free.
The administration of the assessment was just starting to smooth out and then Pearson changed a large amount of safety protocols and added language arts qualifier that lengthened the assessment.

The entire administration process has been bungled. In 2006, there was a single test administration. In 2007, there were two administrations. In 2008, 2009, and 2010, there were single administrations. Each year, there were different directions about what information could be posted around the classroom walls. Each year, educators from around the state complain that the test is too long and each year the test administration process for students is made longer. A simple examination of the statewide results shows wide swings in student performance changes from year-to-year. This is one simple indicator of administrative imbalance.

There was one year, in 2008-2009 school year, that the process was a positive experience.

Question 5	Are you familiar/satisfied with the process used by NCS Pearson, Inc. to score PAWS each year?
-------------------	--

Comment Section

It is quite expensive and screening and selection of scorers is an unknown variable

Length of time to complete test is an issue. Results are not returned timely.

We do not get test results in a formative way to help guide curriculum and instruction for students.
--

N/A

N/A

The scoring of the writing responses is in question. We continually find errors/mistakes/problems with the PAWS assessment in the portion of the system that is visible to us, so we must question that part of the system that we cannot view.

N/A

It seemed that the test results are useful but the timing is lousy.

N/A

Errors in reporting created issues.

N/A

We have issues with our writing scores every year. As part of our district accountability plan, we have implemented fully, The Six Traits of Writing. With the Six Traits, students do a district writing assessment and we have them scored by an outside evaluator– the scores don't align with the PAWS writing scores.
--

N/A

The scoring process is lengthy and involves many Wyoming educators. This system may have been an overreach with the original task force and subsequent design team. The resources expended may exceed the benefit of such an involved process. I would recommend a review to make the process more effective and efficient while maintaining some levels of training and involvement of Wyoming educators.
--

The scoring of open-ended questions has been a problem. NCS seems to have struggled with scoring work from Wyoming students.
--

It takes entirely too long to get results we can do nothing with them from an instructional standpoint.

N/A

We don't know how to help a student because we don't actually understand the weights and percentages of specific groups of questions. One year there may be 15 questions on number competency and the next year 5 questions. How do you plan to work students through a procedure like this? Learning the standards is difficult enough without teachers trying to guess the progress of their students that are actually other teacher's students, when the results are provided.
--

Writing scoring is not done by Wyoming teachers (as was promised) Writing scores show wide variations in the same group of students from year to year.
--

I'm not familiar with how the test is scored but I am familiar with how long it takes for us to get the results. This is a major issue that needs to be addressed.
--

We are satisfied with the technical scoring process (i.e. psychometrics). We are not satisfied with the timeline of reported scores. We need scores by the end of June, so that we can prepare for the new school year. Human

scored items should be scored by more than one person. For most years, there have been errors in reporting, which the WDE or local assessment personnel have had to address after initial reports were delivered.
I just can't believe that scoring the number of tests we have in Wyoming and getting the results back in a couple of weeks is not possible. If we go to paper and pencil it may be months before we see results. We need test results almost instantly for good planning a curriculum delivery.
This was one of the strongest partnerships between the state, Pearson, and instructors.
I am not personally responsible; those responsibilities are delegated to other staff members, generally speaking, building principals and district technology personnel. We are reasonably satisfied that the results of the PAWS testing provides a fair picture of our student's knowledge and performance capabilities.
Fluctuates year by year
N/A
Writing is an issue and has been a problem since the first administration. Who scores the writing and how they are trained is often questioned. How do we build up teacher, student, and parent confidence in the accuracy of the writing results?
N/A
N/A
N/A
The writing scoring process has resulted in inconsistent proficiency determinations across year and across grades in a single year.
N/A
N/A
They follow established protocols and norms for large scale assessments.
We do not get final results until 3-4 months after administration. For 2009-2010 it will be 5-6 months
N/A
Scoring seems appropriate, but the concept of online testing loses its purpose if data is not available for 5-6 months after administration. It is very difficult for teachers to use student data to make instructional decisions and to develop individual learning plans. Representatives from our district have participated in the summer PAWS item review in Laramie and the fall anchor pulling in San Antonio. It seemed that all Wyoming representatives took these tasks very seriously. The writing scoring seemed somewhat inconsistent. It is unfortunate that WDE has not considered artificial intelligence for consistency in scoring, which could be easily done through an online program called My Access.
We have several staff members that have been trained, however; in the case of writing the training does NOT match the scoring.
N/A
My concerns in this area are twofold. The first is the lack of consistency with the scoring of the writing and the second is in the length of time to receive meaningful results.
N/A
Part of the problem of trusting NCS Pearson is the secrecy surrounding the scoring process. Some educators are invited to Texas to participate in aspects of the scoring process, but the scoring itself and rubrics change without rationale a and are not subject to any transparency.
The scoring process also changes every year with rubrics and cut scores.

Question 6a	In your opinion, how do students in your district perceive taking the PAWS assessment? 4th
Comment Section	
We have tried to make it a high stakes test but some students are nervous	
N/A	
N/A	
N/A	

Young enough to still take it seriously.
We responded "somewhat seriously" because of the importance placed on the assessment by teachers, parents, and administrators. This is a generalization as we have students that would fit each of the three categories. Our 4 th graders are the most likely group to take the test seriously of those grade levels presented in this survey. This may drop to "not seriously" in 2011 based on the negative experiences in 2010.
N/A
Depends on the school
N/A
As a fourth grade student's teachers have a great influence on getting students ready for the test and to do their best. I believe our teachers did a great job of trying to motivate our students to do their best.
The students seem still willing to try even though there is no reward for doing well.
See Comments under 11th grade
N/A
Our teachers do a good job prepping these students
We work very hard at educating the kids to take the test seriously. It is difficult because the teachers don't really find the results meaningful and the information provided is broad which makes it difficult to utilize it to inform instruction. Also, by the time the results are back, teachers have an entirely new cohort of students.
We have expended a great deal of energy to make the tests meaningful to students. We have tried to support the test and the importance of doing one's best. This has been difficult with a beleaguered test has an abundance of flaws that students have pointed out and called attention to.
Generally speaking, students take the PAWS somewhat seriously but more out of compliance than authentic participation. The same might be said for some Wyoming educators administering the assessment.
Again, comments vary by building. The most common complaint pertains to the length of the test. Nothing was mentioned about validity and reliability—which obviously needs to be considered.
N/A
Good kids will try but the time spent on the test 10-12 hours is just too much time on one test. Maybe 2 hours for math and 2 hours for reading would be enough for a quick summative assessment. WE can manage the writing in house with state guidance. No science please!
Students start the assessment with serious work habits. As the test continues along this starts to wane and diminish. Attention spans are not considered in the assessment and least of all the actual work of standards that has probably been completed for the school year. Students may see questions that are out of alignment with the standards or the interpretation of a standard. Shock and fear comes in on the student least prepared to handle failure and they stop trying. In general most 4 th graders still want to please their teachers and need reassurance so the work quietly through the assessment, knowing full well that they are doing their best. The bottom line is, perception by most 4 th grade students is that that they are required to take this assessment and teachers request that they be in attendance and take the test seriously.
N/A
In prior years our secondary students have not taken this as seriously as we would have liked them to. This past year we made the effort to educate them as to why this assessment was important. I feel that the majority of the students made their best effort. The fact that the technology not working led to frustration may have caused some to be less than motivated to do their best. That's why the technology piece needs to be worked out.
We have worked hard on this area. It is a district priority and is better at the earlier grades than at the higher grades.
At this age students still want to please their teacher and principal and will try their best.
Our 4 th grade students work very hard on this test.
Generally speaking, it is my belief, based upon my conversations with principals and teachers, that our younger students do take the test seriously and give their best efforts to perform well.
Younger students are more serious
N/A
They usually do the best they can do, but are not always supported by their parents who make choices about attendance.
N/A

We have expended a great deal of energy to make the tests meaningful to students. We have tried to support the test and the importance of doing one's best. This has been difficult with a beleaguered test has an abundance of flaws that students have pointed out and called attention to.
I don't perceive the 8 th graders as taking the test seriously for intrinsic reasons but rather due to teacher pressure to do well or group incentives for performing at a proficient level.
Quote from staff; "Most of the students take it seriously; quite a few do not."
N/A
Again the time commitment defeats the outcome gains. Just takes too long for each section. Maybe 3 hours for reading total and 2 hours for math. Again, Writing could be handled by teachers locally with state guidance.
I don't believe that all 8 th grade students devalue nor do they value the PAWS assessment. And the truth is why they should. The assessment and the results either confirm that they will do well in the next school or that they will be lucky to make it through, so start thinking about dropping out. The assessment is seen as mostly an interruption in the normal school year, perhaps a break from homework, but not much more. Very little of the PAWS information, in their perception makes back to directly impact their learning at the high school level.
N/A
In prior years our secondary students have not taken this as seriously as we would have liked them to. This past year we made the effort to educate them as to why this assessment was important. I feel that the majority of the students made their best effort. The fact that the technology not working led to frustration may have caused some to be less than motivated to do their best. That's why the technology piece needs to be worked out.
We have worked hard on this area. It is a district priority and is better at the earlier grades than at the higher grades.
There is no reason for students to take the test. The older they get the more they realize that there is no reason for them to take it. Somehow it has to be meaningful to them and not just to districts. It really depends on the class in a small district as some classes think it is very serious and some laugh it off. Staff can predict how classes will score because of the pre test chatter. Make the test meaningful and students will take it seriously.
As with the 4 th grade, our students took this test seriously.
With some individual student exceptions, it is my belief that most of our 8 th grade students do take the test seriously and they do give a good effort to perform up to their individual capabilities.
Time commitment is difficult for 8th graders
N/A
They usually do the best they can do, but are not always supported by their parents who make choices about attendance.
N/A
Tired of testing
N/A
Our staff, parents, and students take PAWS testing very seriously.
We have used the Olympics as our middle school theme to get students ready. This is our only school in a plan of assistance. The students, staff, and parents are very upset at the recent events.
N/A
N/A
N/A
Our students take all assessments very seriously. They know that the assessment results matter a great deal to their teachers and principals. Consequently, they persevere in spite of poor item quality and lengthy administrative session. The time estimates offered by NCS Pearson are approximately half the time our students spend on this

assessment. Our eighth-grade students are forced to take the Explore test and these results are not used in any material way. It is difficult to motivate students to try hard on the PAWS test when they are forced to take other meaningless tests the same year.

N/A

Question 6c In your opinion, how do students in your district perceive taking the PAWS assessment? 11th

Comment Section

Allowing 10th graders to bank scores has been a motivator for students to do well so they can participate in all other scheduled activities we offer during the testing window.

N/A

N/A

N/A

They are worried about ACT at this point, not PAWS. ACT's can get them into college, PAWS can't.

We responded "somewhat seriously" because of the importance placed on the assessment by teachers, parents, and administrators. This is a generalization as we have students that would fit each of the three categories. This may drop to "not seriously" in 2011 based on the negative experiences in 2010.

N/A

N/A

N/A

Your highly motivated students work hard on the test, but the vast majority of them see no meaning in the test

Some are serious because Hathaway is connected and families care. However at risk students and do not see how this matters.

We believe a majority of our students take the PAWS assessment very seriously. Before the students take the test we have elders from the community come into an assembly and speak to the students as to how their effort on the assessment reflects back onto the school and community. Students are stressed out when taking the assessment and ask questions as to how they did when they are finished. In our opinion that indicates our students do take the test seriously

N/A

These students are mostly ones who have not succeeded in school and have come to our charter high school.

There is no incentive for a junior to get good PAWS scores. By this time the kids care about ACT scores and college entrance exams. PAWS has no meaning for most HS students. Banking of the scores was a good idea and was part of the initial planning committees work. It allowed the students the gift of not taking the test again.... If you were proficient, and that was meaningful.

We have expended a great deal of energy to make the tests meaningful to students. We have tried to support the test and the importance of doing one's best. This has been difficult with a beleaguered test has an abundance of flaws that students have pointed out and called attention to.

At this grade level, the ACT is of more concern to students. Generally, they want the senior privileges we've attached to demonstrating proficiency on the PAWS and that has increased their efforts to do well on the test.

Quotes from staff; "Some students are very serious. Others get tired of the length of the test with both an on-line portion and a paper/pencil portion in each subject area." Several of our students respectfully questioned the intent and purpose of the PAWS assessment. Their main objectives to the assessment were the large amount of class time devoted to the assessment, their performance being used to determine school effectiveness, and the lack of perceived value to the individual student. As a result of the individual views of the PAWS assessment, some students indicated that they would comply with completion of the assessment, but with little value being placed upon the individual results.

N/A

High school students think it is a joke. Again, there is no incentive for students to do well on the test. They try to meet BOE and ACT scores are important to the ones who want to go to college and apply for scholarships. The Hathaway may be an avenue to create an incentive for kids to perform at their best. I do think a summative test is

needed but for whatever reason (problems) and no incentive for students we will continue to get less than best performance out of our students. I can personally attest as students score in the upper 20s on ACT but don't reach proficiency on PAWS. That is just wrong.
Again, students that are not making their directions and plans for college or university learning want to see once again how they are doing and the PAWS assessment gives them another review. The student that is going through the motions of keeping up in school requirements and trying to graduate is seeing most of this test as trying to answer the questions, And, So? Example, You are required to take this assessment and we want you to do your best! And if we don't, so what happens?
There are no rewards/sanctions for these students
In prior years our secondary students have not taken this as seriously as we would have liked them to. This past year we made the effort to educate them as to why this assessment was important. I feel that the majority of the students made their best effort. The fact that the technology not working led to frustration may have caused some to be less than motivated to do their best. That's why the technology piece needs to be worked out.
We have worked hard on this area. It is a district priority and is better at the earlier grades than at the higher grades. 11 th grade is still a challenge because of the perception that the test score does not affect the students' future.
There is no reason for students to take the test. The older they get the more they realize that there is no reason for them to take it. Somehow it has to be meaningful to them and not just to districts. It really depends on the class in a small district as some classes think it is very serious and some laugh it off. Staff can predict how classes will score because of the pre test chatter. Make the test meaningful and students will take it seriously.
We have had great success with our high school results.
It is my belief that most of our high school students give a serious effort to do their best, however, we know of many students who annually do not care about giving their best or any effort to perform well on this assessment. For these students, there is no perception that doing well on the assessment will be of any assistance to them individually. Therefore, many students provide responses in the shortest possible timeframe with even reading the question. Some students are serious; some students are not serious in any manner.
Has little value to 11th graders
N/A
On the 2009 PAWS tests 100% of the 11 th graders in one of our high schools tested proficient in math. That means they all took it seriously. Most high school students do try, but there is little incentive for 11 th graders to do their best and relevance is often not apparent to them.
Perhaps this grade level is more attuned to the need to do well on the test.
N/A
N/A
N/A
N/A
Not meaningful to students
At this level especially, students see this as a real burden with little benefit or stakes for them.
ACT is more relevant to them
N/A
Our staff, parents, and students take PAWS testing very seriously. When Pearson makes huge errors in administration for the testing online, our older students question how seriously Pearson views the importance of our state test, which affects student attitude – it is easy to give up and blame others.
The question you hear way too often at this level is why. I need a good ACT score. The test is viewed way too negative by the group.
N/A
Students at the secondary level quickly understand the score on the assessment has no implications for them and so some do not put out a great effort.
The changing rules, cut scores and banking have soured the students and staff.
Our students take all assessments very seriously. They know that the assessment results matter a great deal to their teachers and principals. Consequently, they persevere in spite of poor item quality and lengthy administrative

session. The time estimates offered by NCS Pearson are approximately half the time our students spend on this assessment. Our high school students are forced to take the ACT test (or Work keys) though for many students these are actually useful. It is difficult to motivate students to try hard on the PAWS test when they are forced to take other meaningless tests the same year.

The fact that we are a small school, lends to our students caring about the results, But we do have a few that know it means nothing to them and just write in answers. This of course with the small class sizes can give us a failing AYP.

Question 7	In what ways should PAWS be used – Exit Exam, Criteria for Hathaway, Salary Decisions, Funding Decisions, Body of Evidence.
-------------------	---

Comment Section

We use PAW as an additional indicator for proficiency if necessary.

Could be used as one of the criteria, but not the sole criteria.

PAWS is not reliable or valid therefore its results should not be used for any of the above.

N/A

Some of these we can't use PAWS for until it becomes a more reliable test. (the ones marked "no")

It is a requirement that PAWS be incorporated into the district assessment system. We have been advised by WDE to include PAWS in our BOE systems.

BOE process is unwieldy and should be replaced.

N/A

N/A

N/A

The current BOE we use is much more informative because the work is immediate available to inform parents

PAWS was never designed as a "high stakes" test and by using it for any of the above stated reasons changes the intent of the test. It was originally designed for school improvement.

Not until it's proven to be a good tool. Then some of these areas could utilize PAWS results.

This test needs to become valid and reliable before it is used for anything other than an annual test of basic achievement.

No

N/A

The PAWS is not an appropriate graduation exam because it is intended to measure a greater depth of cognitive complexity than most large scale exams. Graduation exams, by nature, need to measure minimum competencies for graduation which, in my opinion, would destroy the positive aspects of the PAWS design. The PAWS could be used for Hathaway Scholarship eligibility. However, there would be students who need the scholarship to attend a post-secondary school that would not meet this qualification. With the opportunity currently afforded by the Hathaway, some of these students will succeed in furthering their education and breaking out of generational patterns of low educational aspirations. I would support PAWS as one of several criteria for consideration of salary decisions if that authority remains with local school boards and not the Wyoming Department of Education or other state entity. The danger presented with attaching salary to PAWS scores is further promoting the pull of your most qualified educators away from the most high risk schools. Furthermore, this would lead to further pressures to apply a vertical scale to the PAWS and then call it a growth assessment. This bears all types of problems. It would be very complex and carry many unintended consequences. Please be well advised from experts in the field of psychometrics before considering such a move. PAWS performance is already a criterion for making funding and other decisions at the district and school level by local school boards and should remain with the locally elected body. There are a number of methods in which the state may support and intervene with low performing schools other than negating a local board's decision making authority. The PAWS is an assessment of 23 essentials skills found in current Wyoming Standards. It is appropriate to use the PAWS in a district's Body of Evidence system as one indicator within the multiple opportunities to prove proficiency.

PAWS will not meet the requirements of Body of Evidence, in our opinion.

I believe the body of evidence should be eliminated and replaced by an assessment based upon the components

at the local level. These would be turned in when districts were underperforming. If districts are doing well allow them to keep moving forward.
Maybe this could be used for determining levels of diplomas if we keep the tiered diplomas.
If the results of PAWS belong anywhere I guess the BOE is the place. Of course with that comes the question of why do we need PAWS or the BOE when review the guidelines of the NCLB, neither is required, under the scopes that they are used.
Why complicate it more? So we can hire more people at the state level? To make PAWS a reliable instrument for any use, students need to take it seriously.
I do not on one high-stakes test. I think a more thorough evaluation of student achievement and review of the transcript ought to be the basis of determining whether a student graduates or not. I feel the same way about using one high-stakes test for determining Hathaway scholarships as I do an end-of-course assessment. We already have research-based assessment in the ACT that weighs in on students' awards for Hathaway scholarships. I do support a growth model as a means of evaluating teachers. The PAWS does not currently support a growth model on student achievement. We do, and will continue to, use the resources provided through the block grant to make sure that schools that need additional support are getting it. I do not support the current Body of Evidence program. When the statewide assessment is working well and has been determined to be a valid instrument I would support it as the Body of Evidence process.
PAWS is used as an alternate indicator for 2 of 9 content areas in the Boer.
Not in the current form, but with some modifications could be used. The problem is that it is not a growth measure for individual students.
This could be one and only one small part of the BOE for the state.
If a student is not proficient, would that mean that the Body of Evidence could not be satisfactorily achieved, despite all other elements of the Body of Evidence being achieved? Is this to be another single snapshot of evidence to could undermine the Body of Evidence process and possibly deny a student a diploma?
I do not feel the required body do evidence meets our districts needs. There are better solutions.
N/A
One problem is that PAWS only assesses some of the standards for graduation. Another problem is that if we wait until the beginning of 12 th grade (when students get their results) it might be too late to rectify. A third problem is that requiring PAWS as criteria for BOE would mean that virtually no one from out of state could enroll in our high schools as seniors and hope to graduate. Currently we do use PAWS scores for BOE only as an alternative piece of evidence to demonstrate proficiency if students have not demonstrated proficiency on common district assessments in prior years. My question would be why do we even need a separate BOE system for graduation?
N/A
As long as there are challenges with test administration, the results should not be used to make important decisions. I truly believe that ACT scores could be used in place of PAWS scores for any of the listed items in the high school. Parents and students see value in the ACT results and work hard to achieve. They are offered multiple opportunities to take them and therefore they results are more accurate than a one shot test. If we could use ACT results than we would align our curriculum to the ACT test.
An appropriate statewide assessment would serve many functions and purposes like those mentioned previously. However, PAWS in its current state fails to serve anything very well- least of all instruction.
The use of a single indicator as the determination of a student's master of content is contrary to the research and best practice in the field of educational assessment. The use of course embedded performance assessments provides an authentic opportunity over multiple measures for a student to demonstrate mastery and allows for greater classroom credibility.
N/A
PAWS is a snapshot test that does not necessarily capture the learning progress of students.
The test is fatally flawed and uses a standard percentage of proficiency models instead of a growth model. This measure is especially sensitive to levels of poverty and language issues. This test is also not instructionally sensitive or timely enough to be of any benefit.
It is not a valid and reliable assessment.
Already use it as an option.

There are a number of variables that affect student performance in a state test, which impacts the credibility of a true measure of individual academic performance. The tested areas in reading, writing, math, and science does not always allow for students to show understanding and critical thinking skills that are extremely important in for an adequately educated life-long learner. PAWS does not measure all standards and is only a snapshot of student proficiency. Testing is a mere snapshot of student work – there is no measure of authentic, performance-based learning. The test is biased; it does not measure the skills of all students due to different learning modalities, especially the auditory and kinesthetic learner. Effective assessments should be congruent with student learning modalities – the state test does not measure this.
Evidence is fact based; not a once in a school year shot.
N/A
Good measurement of student learning should always involve multiple assessments. Putting high stakes on one test should never be done. In addition, if we believe learning should be the constant and time the variable then one assessment without a chance to improve scores to demonstrate improved learning should not be considered.
N/A
PAWS is such a poor test it should not be used for any other purposes. But if Wyoming had a sound statewide assessment system it could be used for any of those purposes (except the high-school stakes for graduation, which is a poor idea). Wyoming educators are not fearful of accountability. Rather, we are anxious about being held to account to poor instrumentation.
The way the current test is designed, holding graduation, funding teaching jobs accountable is not fair to anyone.

Question 8	What innovations do schools in your district employ to encourage students to take PAWS seriously?
Comment Section	
N/A	
Parent communications, class challenges, practice, reward for proficiency.	
We work with students throughout the year to take all classroom work and assessments seriously.	
At the elementary level student who take PAWS seriously earn coupons than can be traded for roller-skating and ice cream cones. Students at the middle school (8th) grade are given awards for showing growth on PAWS scores overtime.	
Pep assemblies, snacks, bowling afterwards, novelties, and no school for those who bank a previous year, pizza parties.	
Parent education and support, Rewards such as class parties, field trips, etc..., Nutritional food Teacher supervision of the assessment, Teacher discussions with students about the importance and impact of the assessment, Principal pep talks	
Various incentives for younger students, Middle school students lose an elective if they are not proficient.	
N/A	
Incentives for behavior during assessment	
Our teachers talk about the test and to put forth your best effort. We have breakfast for the students during the testing to make sure their minds are focused on the test and not their stomachs.	
Extra length periods, excused absences accumulated, food, drink, parties.	
As stated in #6, students are encouraged by staff and community members to take the test seriously. We remind parents to make sure their children get plenty of rest and encourage the students throughout the testing period.	
No prizes or rewards. We talk the importance of testing in learning and about the significance of doing your best.	
This past year- teachers went to testing with the students an activity dog or movie, in movie theater was promised along with pizza and ice cream bars.	
We make the allowable accommodations for every child by utilizing many staff members as listeners etc...We provide snacks and high energy foods during testing. Based on Quantum Learning and brain research, we get kids up and moving at specific intervals during the test to increase brain stimulation and activity. We communicate with and educate parents about the need for a good nights rest and the detrimental effect video gaming has on brain functioning – especially prior to test taking.	

We emphasize the importance of the test. We feed students breakfast during PAWS window. We use the test results for programming in to extended learning opportunities. We use the tests as an indicator of transcripts status, and we make students aware of the uses of these data.
We have implemented a Senior Incentive Class which is offered first hour to 12 th grade students who meet the criteria, one of which is proficiency on the PAWS. The class meets in an informal setting to study, work on scholarships, online courses, etc. Credit is not given for the class so students must be on target with credits to graduate. Students want to be eligible for this class as our district does not include any study halls in our school schedules.
Encouragement from teachers to do well. At our school, teachers really rally our students by sharing the publicity that the test results get. Other than that, we do nothing. Admin/Counselors met with students prior to the assessment to discuss testing procedures and value of assessment results. School programs and interventions designed to assist students with improved PAWS performance were discussed. Breakfast was served to students by building staff on the first day of testing. Snacks were provided throughout the testing schedule. The testing schedule was designed to minimize missed classroom instruction and to provide an appropriate testing window. A letter is sent to all parents outlining the PAWS assessment schedule and the role of parents playing in achieving student success.
Part of a comprehensive awards system for funds exemption at the high school level.
We have tried them all from rally nights, pre test celebrations, post celebrations, parental support ideas, etc. Our problem is the test is not available for immediate feedback so you lose some of the excitement if you get them ready in a good way prior to the test. Kind of a letdown waiting for results.
Fortunately the students at a small district take assessments more seriously than we might suspect. Not every student makes funny pictures or writes inappropriate materials on tests. The over glamorized opinion that all students don't take assessments seriously is a mistake. So we may offer a food break at a different time. The principals let students sit outside for a few minutes if a nice spring day comes along. Students may be release from core classes to elective classes for times. Students are motivated by praise and respect for themselves to see what they can do.
Pride and self satisfaction, relational approach, proof of learning exclusion from remediation.
Educating them on the importance of the exam and how it impacts AYP for the district. Many students simply did not understand how important it was.
PAWS Academic Assemblies, PAWS Rap Song, Turn the principal into a Banana Split, HS Seniors earn their own parking lot if they met PAWS targets, and Students could wear hats in the building if they met PAWS targets. Schools have implemented school, classroom and individual student goal setting conferences with students working to meet expected target scores. Our district has developed PAWS Skill specific rubrics for routine use in classrooms prior to PAWS administration.
We really don't offer any innovations. We used to have a barbeque and breakfast on PAWS mornings but they seemed to make little difference and used up valuable instruction time.
We do not do anything. We ask them to work hard and thank them for the results.
Like other school districts within the state, each organizational level, indeed, each building, has their own strategies to engage students in the process of trying to do their best on PAWS; there are rallies, incentives, encouragement and any other number of strategies and tactics employed to encourage students to do their best.
Rewards, Parent info, media Promotions. Snacks, pep Rally
We provide snacks and water prior assessment. Student are reminded to do their best.
We help students set goals for their achievement in math, writing, and reading. We teach them to be good readers, writers, mathematicians, and thinkers. Then they will do fine on PAWS. We don't pay them or give them time off. We do celebrate successes the next year with an ice cream treat.
Banking test score, celebration of school and community by a parade, non proficient students are put in classes the next year to aid in instruction to become proficient.
Incentives program for different grade levels.
Teachers encourage students to put forth their best effort (Elementary). We promote it prior to testing. We discuss its value during advisor/advisee time. We send letters home to parents and students. We brought in a speaker to talk to the kids about life goals and making choices to support your goals. We serve the students breakfast prior to testing and serve snacks between tests. We also gave students a longer lunch break on test days. We use PAWS

<p>scores as tie breakers for valedictorian and salutatorian (High School).</p>
<p>There is a fine line between berating students and seeking authentic results. Nonetheless there is a line and it should not be crossed. Our schools encourage growth from previous years and emphasize doing your best. At the high school those students who are proficient in all areas have more options as seniors in terms of coursework and class schedules.</p>
<p>Parent informational sessions and newsletters, assemblies, breakfast at the school, rewards and early release.</p>
<p>N/A</p>
<p>I would be happy to share video clips of our student's news programs @ 4th grade. No homework @ 8th grade also works.</p>
<p>Effort awards are given at all elementary and middle schools.</p>
<p>Motivational assemblies, recognize previous success, food/beverage rewards, and parent's awareness.</p>
<p>All year we work towards PAWS in classrooms and remediation. We talk about the use of test results to determine placement and remediation. Teachers encourage student's everyday.</p>
<p>Elementary Level: Internal Motivation: Teachers tell learners here the importance of the tests (to give us information about what they know so that we can teach them better) and that it is very important that they do the very best they can. The teachers talk a lot about self-respect and being able to be proud of what you as a learner do. External Motivation: Some teachers do give treats after a very successful testing session but more importantly the teachers praise individual students. Teachers refer learners to the principal's office and other people in the building to be praised for their hard work and success. Middle Level: Reward Days, Dances, Gift certificates, McDonalds, A and W, Snacks and Food for Dances, Air slides - activity day High School: Put names into a "hat" for each time that a student were proficient, two slips if advanced. Prizes were given if one's name was drawn out of the hat in front of the student body. Note: I had read that Natrona High School worked "with" students to determine what would motivate them. Students wanted better parking at the school. The school administration worked with the students to provide this request, established a target for PAWS achievement for those students, and agreed on students achieving a certain level of proficiency to get a student parking lot. I don't know the outcome, but this seemed to be a win-win situation for the school – all stakeholders had input and commitment for students to achieve.</p>
<p>We supply healthy snacks, pep assemblies, incentive days for both growth and proficiency.</p>
<p>Reading intervention class for students not scoring proficient or advanced</p>
<p>For the most part, students are talked to and encouraged to do their very best on the assessment. It is pointed out that it may not impact a student in any way it does have significant implications for the district.</p>
<p>We have created a culture of competitive spirit. We want to be the best in Wyoming, administration, teachers, and staff.</p>
<p>We hold assemblies at each of our school to help engage students' motivation to do well on these statewide assessments. But we do not offer (financial or other) inducements to do well on tests. Learning should be its own reward</p>
<p>Teachers display a good attitude about the testing and discuss with students the use of the testing results and the impact it has on improving instruction.</p>

Question 9	Does your school district use PAWS scores as part of its "Body of Evidence" with respect to graduation and mastery of the "Wyoming Content Standards"?
Comment Section	
The BOE is not needed	
N/A	
N/A	
N/A	
N/A	
It is a requirement that PAWS be incorporated into the district assessment system. We have been advised by WDE to include PAWS in our BOE systems. Our system in not dependent on PAWS and any changes with the assessment would not impact our system.	

N/A
N/A
N/A
We have it in our plan as one way to meet the criteria, but have not had to use it.
We do our own benchmarks to the PAWS indicators so in a round about way there is a connection.
Our district uses the results of the Paws at the high school level in mathematics for determination of the proficiency levels. If a student desires an advanced endorsement on their diploma a combination score of PAWS, along with MAP assessments are combined to give the student a proficiency in mathematics.
Part of the standard setting process to determine cut scores. But not used annually for analysis of student mastery.
N/A
N/A
PAWS serves as one of three indicators in our BOE system. The other two indicators are the districts standards based test and the Northwest Evaluation Association Measured of Academic Progress. Transcript endorsement uses the PAWS scores for determining proficiency levels.
It is one indicator for Language Arts, Mathematics and Science.
N/A
N/A
N/A
The answer is yes but we also are required to use additional evaluation sources. Some of these sources provide information that can be used in a timely manner to plan for the next year.
It may be used in the appeal process.
N/A
PAWS could be used as an alternate indicator for 2 of 9 content areas in the Boer, but is rarely, if ever, used.
It is one part of the entire picture of a student's academic success or lack of. WE don't use it as a sole factor and would hope no one does.
It can be used in Math, Language Arts and Science as a support when there is a challenge on the student's performance decision.
Students who are notified that their Body of Evidence materials are not sufficiently acceptable for graduation may appeal and utilize their PAWS test results to further their case for a satisfactory Body of Evidence program. This appeal route has never been requested by any student in this district.
We utilize other assessments
N/A
Currently we do use PAWS scores for BOE only as an alternative piece of evidence to demonstrate proficiency if students have not demonstrated proficiency on common district assessments in prior years. My question would be why do we even need a separate BOE system for graduation?
N/A
Since it does not really measure state standards, this has not been an area we have used in our body of evidence.
It is one of four data strands in language arts and math.
N/A
If a student has not shown proficiency on BOE district assessments regarding standards and/or benchmarks that are assessed by PAWS, those PAWS scores are used as an alternate for a student to show proficiency in reading, writing, math, and/or science for graduation.
N/A
N/A

We do not use PAWS to make the determination if a student is proficient for not. We use it as an additional opportunity to demonstrate proficiency. If a student is not proficient on PAWS students can still continue to take district assessments until they demonstrate learning at the proficient level.
N/A
While we allow PAWS proficiency status to count towards the entire mathematics content standard, the reading and writing assessments are only part of the overall language arts content standard. Thee body of evidence has been poorly managed by the Wyoming Department of Education, but that is the subject of another survey.
N/A

Question 10a	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your “yes” and “no” answers.: • Northwest Evaluation Association MAP (Measures of Academic Progress)
---------------------	--

Comment Section

N/A
It is more reliable than PAWS.
N/A
This is a formative assessment that could be made more robust to possibly meet NCLB standards.
MAP's would be perfect in my opinion. It's already catered to Wyoming standards. And best of all: results are instant and reports are ready in hours/days, NOT months. Love this deal!!! Plus a lot of districts are using it anyway. We have instant instructional data.
This assessment would not meet the necessary requirements
A bench mark assessment that yielded usable results quickly.
Easy to give and reliable
Have not used this in our district
We could use it, but what I understand that the feds will not let us use MAP. We do need to develop an assessment that is move of a growth model that would shows grow over time. This I believe would give us better information and would show students how they are improving over time. This would make them move motivated. They could see their own growth.
It is one thing with immediate results and with a bit of tweaking could match Wyoming standards exactly
NWEA's current MAP assessment is an interim or periodic assessment used to measure student's growth. It is an adaptive assessment meaning the assessment adapts as to how the student answers an item. Items get harder when students answer correctly, gets easier when students answer incorrectly. The results help teachers differentiate their instruction based on where students are. In order for the MAP assessment to be used to meet the current NCLB laws NWEA would have to design a test that would pass the current US Department of Education requirements. Until reauthorization allows states to include growth models MAP could not be used as presently formulated.
MAP is a wonderful assessment tool which we utilize in our district. I understand there are issues with NCLB requirements because of it's adaptive rather than mastery approach.
Not without adapting the test from its current format. I would prefer to use it and I think it can be readily adapted, but in its present form and usage I do not think it currently is acceptable to the feds.
MAP is a growth model assessment. Jim McBride has said that we can't use it but he offers us no policy or evidence regarding that decision. MAP is a valid and reliable assessment and it provides us with usable data that can inform instruction. It allows us to look at the growth levels of individual students to see if they are making adequate or expected growth, no matter where they start. This is so important because it hold us accountable to those students that are highly proficient but may not be making a years' growth in a year. In our district teachers create individual learning plans for every student based on MAP data. The plans include specific learning goals that the teacher creates and then the students meet with their teacher to discuss their MAP assessment results. At this point the students set their own learning goals that are specific, measurable, achievable, and relevant and then set a timeline set to achieve the goals. This year students will lead their own parent/teacher/student

conference and share their assessment data and progress on personal goals. MAP is a meaningful assessment that we (the admin, the teachers, the parents, and the students) tied directly to student learning and performance. MAP assessment data is our foundation for continuous improvement and we believe that our district will achieve success – one learner at a time. We may be punished for our PAWS scores, but we have so much to celebrate when we look at the growth of every (almost) student. This is our focus – we live with PAWS but we believe in MAP>
As part of a system. The MAP test is a very powerful test that is computer adaptive and yields growth scores. It includes diagnostic reports and prescriptions for individual students and can also be aggregated for class level and building level data patterns and profiles. The MAP test would be insufficient in isolation, but as part of a system it would be a powerful component.
The MAP assessment would not meet NCLB requirements. It is a vertically aligned skill based assessment which provides a helpful rather quick snapshot of a student's skill level but does not measure the cognitive depth PAWS is designed to measure. It makes sense that individuals are more comfortable with teaching and assessing skills but that does not reach the depth of learning reflected in the state standards.
NCLB requires the use of standardized achievement tests and MAP is norm-referenced. There are also significant alignment issues. How would science be addressed/assessed?
This is a growth measure and accurately reflects student learning.
We would need to create a summative level. Still use the formative during the year. I realize this has not happened yet but it sure would have staff and parental buy in if we could pull it off. We would have specified and immediate feedback that would assist in teaching and parents understand and trust the results. WE need something that works!
This is a tool that actually offers ideas for student improvement and contains growth allowances. This can be used but not under the expectations for graduation assessments as they currently exist.
Teachers have faith in the MAP. MAP results are available immediately and teachers use them in lesson planning and program improvement. The test is very accurate and nationally norm referenced. It would give an accurate picture of Wyoming Students compared to others.
N/A
NWEA could develop a state test, but it would have different characteristics than the MAP that is currently used. This would require an RFP process.
All districts in the state currently use these tests and with very little alteration they could be used. The really good thing is that the scores are returned instantly and can be used for instructional purposes.
MAP is an adaptive assessment and if we were to use it we must consider different evaluations focuses. It would work provided the rules were changed to allow growth model format or MAP would have to be adapted in it's current format.
I am not familiar with this assessment and therefore could not give a yes or a no.
It would have to get NCLB approval. MAP has been working well for us.
N/A
I don't think it meets the rules and regulations of NCLB, nor is the assessment aligned sufficiently to state standards.
If aligned with new national common core standards
N/A
National comparisons and state comparisons are available. The results are immediate and personalized. It would serve students, parents, and teachers by informing instruction and targeting improvement immediately.
Only if it is used as a component of the system for growth purposes, as MAP is not standards based assessment.
Map is a growth test. It's different type of instrument.
I am not sure that this should replace PAWS, it is a complementary indicator on growth do students learning v proficiency level.
However,, this only works at a low level, this would have to supplemental for other levels of cognitive depth.
Student friendly, easy to administer, immediate feedback, normed, All districts use it.
Quick results, short test, teacher and student friendly
MAP is used by all districts. It is an excellent growth model for teachers to examine individual student performance and to target individual learning needs. A growth model makes more sense. MAP is very helpful in

that it provides “quick” results – the data does not linger for a long period of time as with PAWS. Originally, this was the intent of PAWS, but Harcourt (now Pearson) changed the timeframe for delivery as WDE staff changed; WDE staff members, who were new to WDE, accepted the change in the timeframe without knowing that the original agreement for delivery was by the end of the school year – now it is within 6 months after the following school year has started. As long as MAP is aligned to Wyoming standards, this would work. The cost for MAP testing isn’t nearly as costly as PAWS.
We have the results quickly and students are able to use the data to self regulate. Students and teachers can make data related decisions.
Immediate feedback and shows growth of students
I do not believe MAP could be used at this time. The primary reason is we would have to first do alignment studies to ensure it is measuring what the Wyoming State Standards call for. In addition, a study on cut scores would have to be initiated to determine what proficient looks like. If that work was done, then I think it would be possible to be used as a state assessment.
This is not a st
With some minor modifications to meet current requirements for no child left behind, the MAP tests would be a far better measure for statewide assessment purposes than PAWS. It has greater validity and reliability, comports well with the state standards, ahs greater efficiency in terms of time and dollars spent, and can be used for other purposes including classroom diagnostics and program evaluation. Utah is currently working with the US Department of Education to get MAP approved for NCLB purposes. We should cancel the contract with NCS Pearson immediately.
They are two different tests that are meant to measure different aspects of learning. PAWS is summative and MAP is a formative test.

Question 10b	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your “yes” and “no” answers.: • AIMSweb
Comment Section	
N/A	
No multiply choice questions the scope is too broad.	
N/A	
Same as above (MAP)	
No idea what this is	
This assessment would not meet the necessary requirements	
N/A	
N/A	
Have not used this in our district	
This is a growth model and we can show how students are growing over the course of time. I’m not sure if this would meet the feds requirements.	
More appropriate to track standards with disabilities	
Does not currently assess Wyoming Standards, so could not be used to measure NCLB requirements. Diagnostic Assessment (classroom)	
N/A	
This is a great assessment as well but I don’t think it meets the requirements of what PAWS is supposed to be.	
Don’t know about it.	
Aimesweb is a comprehensive monitoring system comprised of benchmarking, Strategic monitoring, and progress monitoring.	
N/A	
Not sure about this assessment.	
N/A	
N/A	

Difficult to understand the use of this tool in any view point. Seems like the first of many that jumped on the wagon after NCLB was adopted. Promised a data strand for student progress and monitoring but not as a final evaluation of a student learning level.
N/A
N/A
This test is designed as a progress monitoring tool.
We use AIMS WEB but is somewhat limited and would probably not be a good measure for NCLB.. It might be a part of two or three tests and then could be used.
This is more of a diagnostic assessment and not a formative or summative evaluation.
I am not familiar with this assessment and therefore could not give a yes or a no.
Could be one piece for triangulation.
N/A
I don't know if the Aimesweb assessments are aligned to Wyoming state standards nor is it a summative assessment. I don't think it meets the rules and regulations of NCLB.
N/A
N/A
Primarily diagnostic testing to determine placement for instruction.
AIMSweb probes are screeners only; they are not in any way a measure of a student's mastery of standards.
N/A
This is screening test for RTI programs
N/A
N/A
N/A
Districts were asked by WDE to declare a diagnostic test for reading at the elementary level. The choice was either MAP or DIBELS. DIBELS NEXT has an option for the same format as AIMESweb called V-Port, which is going to be used by some districts. This would be a duplication of information. The idea is good though – if we could all be on the same page
Not fully aware of what they have to offer.
N/A
This assessment is very limited in scope for measuring the content areas of math and reading.
N/A
We don't use the AIMSweb and therefore can't comment.
Do not have any experience with this test.

Question 10c	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your "yes" and "no" answers.: • ACT for 11th graders
Comment Section	
giving the ACT with the writing component would be a valued assessment measure	
Need to add writing component	
N/A	
ACT is Not a standard based assessment! This assessment is designed to sift and sort a body of students and rank them in order of performance. It does not measure a student's performance against a fixed standard.	
All 11th graders already have to take this national test. Which has years of successful use to back it up? I think we need a uniform test across grades though.	
This would be an inappropriate use of the ACT exam. It is not aligned to the Wyoming Content and Performance Standards	
Normed assessments are not as useful as CRT.	

N/A
No
This is something all parents understand as well as the students. I also think you would need to include the Work Keys.
1/2 day nationally normed summative assessment that students take seriously
Does not currently assess Wyoming Standards, so could not be used to measure NCLB requirements. National Assessment - Summative
N/A
This is not the purpose of the ACT. It does a great good but it isn't what PAWS is supposed to be.
However, EVERY student at Shoshoni takes the ACT and we do consider ACT data as a factor when we look at the success of our HS.
N/A
Not aligned to Wyoming Standards. Not the intended purpose and design of the ACT.
ACT claims to be a predictor of College Readiness. It's purpose doesn't align with NCLB, unless we are going to require all students take a "college track."
The growth model outcomes more accurately shows the learning process within a district.
Only for college bound students. We test all kids and you cannot compare the scores to other states when they only test 40-65% of the students.
Again the State would need to be careful with the expectations form these results. The report levels for college readiness may be influenced by the state requirement that all students take the assessment. Percentages need to be captured but at the same time the college readiness levels may change based upon scores, unknown.
For the percent of students with no plans to attend post-high training at an institution that requires ACT, there is no incentive. Results are not timely.
N/A
This test is not designed to measure student performance on state standards.
I am not convinced that the ACT is a good measure of all student learning. It measures college readiness fairly well but not all students go to college and that limits its effectiveness. It is already the gatekeeper for Hathaway
ACT could be used for the state test if the National Common Core Standards were aligned to the test and our State continues to support the National Common Core Standards.
The ACT test does not assess Wyoming students on Wyoming standards of Wyoming curriculum. I do not see how the ACT could actually be aligned to meet the specifications of a state standards based test.
If averaged with MAP. Don't want one high status test.
N/A
I don't think ACT is aligned to Wyoming state standards. The purpose of ACT is different than the requirements of NCLB.
N/A
N/A
State and National comparisons needed for most college entrance requirements - less duplication of assessment time.
The ACT is designed to predict a students chances of success in post secondary education; it is not designed to assess a student's mastery of standards.
N/A
Not at this time as it is not aligned to state standards - perhaps in the future.
This would give districts a much clearer picture of career and college readiness, and students would take it more seriously since it is already required. This would eliminate extra testing with PAWS. I would also suggest that the Explore and Plan also be required to show growth over time.
Valid and reliable, normed, credible because it ties into Hathaway and college entrance.
Good test
This is a very good idea – all juniors are required to take the ACT. This would also motivate a large group of students who have an interest in the Hathaway Scholarship. I fully embrace the opportunity for students to be able to bank their scores. I don't know how this would work. There would be little motivation for some students who

choose not to look at ACT as an important assessment tool.
It has meaning. It is valid and reliable. If NCLB continues to have special education sub groups tested with the ACT type test, I worry about the impact on these children.
N/A
ACT is a possibility, however I would want to see alignment studies and cut scores established first. Just saying this is the new state assessment without the additional work would not be appropriate in my thinking. However, ACT and Work Keys does measure College and Career Readiness which is where all schools need to focus.
This has potential since all 11 th graders already take it.
The ACT test is designed for college entrance purposes and is not aligned with state standards for all students.
Better test and it is nationally recognized. Does a comparison of not just our state, but your students compared to al other states. Could eliminate the amount of time spent on testing of students, provide more instruction time.

Question 10d	There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your “yes” and “no” answers.: • DIBELS (Dynamic Indicators of Basic Early Literacy Skills)
Comment Section	
N/A	
More of a progress monitoring assessment	
N/A	
Not robust enough.	
Too short of an assessment. Helps with instruction.	
DIBELS is only appropriate for the elementary level.	
N/A	
N/A	
No	
Yes, growth over time	
Only reading literacy program, good info about but not appropriate	
Does not currently assess Wyoming Standards, so could not be used to measure NCLB requirements. – Diagnostic Assessment (classroom)	
N/A	
Same	
DIBELS is useful in finding elementary reading levels and it helps us identify our students most in need of intensive intervention.	
DIBELS is intended for early literacy skills. It has wide acceptance and is used by this district for progress monitoring and formative and summative assessment. There are defined benchmark scores that define the proficiency level of students. It is administered three times per year - beginning, midyear, and end of year. It could be part of a system (and is part of ours), but would be insufficient for a comprehensive assessment system.	
DIBELS is a reading fluency test. It measures fluency and some small amount of comprehension. It is best suited for the classroom and school level use. It is not even a comprehensive reading assessment (fluency is one of five components of reading) and, in my opinion, should not be used at any level of state reporting.	
N/A	
A good measure for having students ready to read by end of Kindergarten and also at grade level by the end of third grade. A real indicator of academic success in the future.	
N/A	
Not designed for this, a progress monitoring tool, even slightly poor for that task	
For lower grades	
N/A	
This test is designed as a progress monitoring tool.	

See AIMSWEB
This is more of a diagnostic assessment and not a formative or summative evaluation.
The DIBELS test does not assess Wyoming students on Wyoming standards of Wyoming curriculum. I do not see how DIBELS could actually be aligned to meet the specifications of a state standards based test.
DIBELS is better suited for K-1 students.
N/A
These are not aligned to Wyoming state standards and the purpose of DIBELS is much different. It is a formative assessment.
N/A
N/A
Mostly for the progress monitoring with shorter intervals between sessions- Hard to use for state and national comparisons.
DIBELS probes are screeners only; they are not in any way a measure of a student's mastery of standards, and in fact in the most recent training by the authors they assured legal action would be taken against any LEA/SEA that uses the probes for a purposes other than for which they were designed.
It's a screener
This is also a screening test, not an actual measure of student reading.
This is an inauthentic and simple measure of literacy.
N/A
Quick and dirty at lower levels
DIBELS (and DIBELS NEXT) is used as a diagnostic tool that affords monitoring. It could work as well.
Now that they include reading comprehension, it could work for the appropriate age groups.
N/A
This would not be appropriate for all of the grade levels not does it have the scope needed to determine if a student is graduation ready.
This is not a gauge of proficiency.
We choose not to use DIBELS as it is not aligned with our instructional strategy, and therefore can't comment.
For younger students this would be a great option for a valid and reliable test for them.

Question 10e	10e. There have been suggestions that the following assessments could be used in lieu of PAWS to meet NCLB requirements. In your opinion, could Wyoming currently use the following assessments in place of PAWS? Please explain your "yes" and "no" answers.: • NAEP (National Assessment of Educational Progress)
Comment Section	
N/A	
Supposed to be the best test....however we don't get school level feedback of the test.	
N/A	
This has not been aligned to Wyoming's standards as they now exist.	
National test, very organized, not a lot of room for error in administration. However this test a limited number of grades that goes to compare how we're doing against other states, not how the individuals did.	
NAEP does not provide individual student or district level data to districts. It would be an inappropriate use of the NAEP assessment.	
N/A	
Can they get it right?	
N/A	
Once again parents and students understand this assessment.	
No certainty that interestingly normed test would tell us the information for US standards	
Does not currently assess Wyoming Standards, so could not be used to measure NCLB requirements. National Assessment/Summative	
N/A	

Same
The purpose of NAEP is to look at national levels of progress. No individual student data is ever received by districts.
We do not get student, classroom, building, or district results.
NAEP has been clear that it is not designed to be a state assessment and, last I knew, would not allow the federal government to use the assessment for purposes of NCLB.
NAEP is a good assessment but reports are limited to the state-level and urban school districts. Individual results would be needed for school improvement efforts.
A good comparison to the rest of the country. Wyoming does very well.
It does not provide student or school scores only state level reports. The test would definitely give comparisons to other states. Might not be a bad idea. Not sure how aligned it is to our curriculum. If the feds mandated a national curriculum along with a national exam we could move forward and quit wasting time on PAWS.
Potentially but may be stretched when used as a data collection tool and not a benchmark calculator. The norm group is from across the US and this may be difficult to relate with Wyoming expectations.
If MAP is not available. There is still the problem of student motivation at the upper grades.
N/A
This test is not aligned with state standards. We have never seen student or school level results. This would move us into the realm of a national assessment.
Local districts don't even receive NAEP tests. We have no idea how our local students score. That would have to change dramatically for it to even be considered.
This test does not provide individual student results and is not aligned with the State standards and probably will not consider the National Common Core Standards as part of their test.
It is my personal opinion that this nation should simply establish a national curriculum that would be required for each student to learn in any school in the nation, no matter where they live. Currently, with high stakes testing and the desire to have students appear to be performing well in given states, many students are being shortchanged by state identification of low level performance standards, which give the appearance of high level student performance. The establishment of a national standards based curriculum and a national assessment such as NAEP would create a level playing field and a national standardized measuring stick by which to determine student's performance in any school in the nation, based on the same curricular expectations and assessments. To the best of my knowledge, the United States is the only nation in the industrialized world that allows for 50 different curricula and 50 different assessments. It makes no sense, even though I fully understand and appreciate state versus nation sovereignty in this particular area.
Need a test to measure students on annual growth not groups.
N/A
At this point NAEP only tests certain subjects each year and only in grades 4 and 8. We don't get individual or school or district results.
N/A
It would be interesting to compare our school or measure our school against others in other states. How do the common core standards align to the NAEP?
If the Feds would allow in lieu of state assessment in 4th, 8th, and one time in H.S. - For that year use the NAEP results as a alternate assessment for elementary, middle, & H.S. use your state results for comparisons.
The standards on which NAEP is built is vague and it is a norm referenced assessment and is not a measure of a student's mastery of standards.
Could be a possibility
The administration of NAEP does not test all students
N/A
N/A
We know little about this test, since we neither administer it nor see results.
The NAEP assessment is an excellent assessment and provides very good information, but a concern would be access to data at the district, school, teacher, and student level. At the present, there is no way a district can access student performance at the local level – it is only available at the state level. Districts need their own data to drive instruction.

My understanding of this test and the usage will not match.
N/A
Limited grade levels are tested, and now district or student results are provided. Students only take one content area each year and the cut scores for proficiency are not overly realistic.
N/A
NAEP does not report individual scores and does not test all students as required by NCLB.
We are not provided with the detailed results and it only tests math and science.

Question 11a	For Wyoming, what are the pros and cons of the following options pending reauthorization of the Elementary and Secondary Education Act (ESEA): <ul style="list-style-type: none"> • Using the current contractor to amend PAWS after reauthorization ESEA.
Comment Section	
The test itself is not a bad with future administration assurances and qualifications.	
Why would we use the same contractor we just had significant troubles with?	
The current contractor for PAWS has not delivered a reliable valid test. PAWS should be replaced with MAP which is reliable and valid and used by all districts now anyway.	
The current contractor has not been able to deliver an acceptable product. On the other hand, the WyCas contractor (measured progress) delivered a great product.	
Pros - not breaking the contract, mending issues. Cons- really need to go a different route than PAWS	
The current vendor has demonstrated an inability to deliver on the assessment needs for Wyoming	
All of us are tired of explaining to our constituents that we really have no valid baseline due to annual change to our state assessment.	
N/A	
?	
I think we can do better. I know they are one of the largest testing company, but what they did with the administration of PAWS was inexcusable	
Pro- would save a huge amount of state money; Con - Current contractor doesn't have a good reputation and even repairing PAWS would be suspect	
PROs: We know what the current contractor is capable of completing; CON's: Same as the Pro.	
N/A	
Depends on what reauthorization looks like. If a growth model becomes acceptable (my preference) then we should go with one of the two or three companies that are already developed in this field.	
If we can't build a growth model assessment, using Pearson would keep us from "throwing the baby out with the bath water". Consistency builds reliability.	
N/A	
Supportive with regard to amending PAWS after reauthorization. As far as the contractor, there are only a limited number of companies that have the capacity to provide this service to the state. Continuing to walk out instead of working through the issues is only to our own detriment. Bid when multi-year contracts come up.	
Based on NCS's 2010 performance, we would say no to this suggestion. Tech support was totally unacceptable.	
I do not think Pearson can meet the needs of Wyoming. They area too large for our market. The assessment should be more in line with our improvement model. I think that MAP aligns with the improvement model being proposed.	
We need a state audit on our assessments in place now and then plan a new assessment strategy with input from all stakeholders on what a transparent Wyoming assessment package might look like in the future. What we have now is too much testing and nothing is connected. Every time the legislators fund something new they add an assessment (i.e.. Elementary reading, summer school, etc.) This just duplicates or partially measures next to nothing. WE need one plan that covers the basics that patrons, staff members and students understand and can accept. Legislators are not policy makers for education. Most don't have the expertise to make such decisions. The one size fits all mentality from the feds is not the answer. Wyoming and other rural states need to create a testing plan on their own that meets federal guidelines.	

I believe that the Pearson product has been tainted with doubt. Regardless of the findings of the third party reviewer most members of the Wyoming community are feeling a fear that the PAWS assessment is not the best tool for the student. The alignment of assessment questions are in doubt. The writing prompts are considered bias. The reporting timeline and data return do not help the districts. As such the entire data flow is suspect.
The current contractor has no credibility.
It's hard to make a judgment of the pros and cons of any program until the reauthorization takes place
Con: lack of trust in current contractor to deliver and manage the project.
This would probably be the cheapest and maybe the easiest fix. It still is not a growth measure and doesn't let us measure growth of individual students in an adequate manner.
We think any contractor would be ok but a true assessment company and not a multipurpose company might be more appropriate.
I have no opinion other than to say that performance of any given school and district should be based on an improving performance model rather than a static performance line that is arbitrarily raised periodically. The variables in given schools and given student bodies makes it virtually impossible for some schools to ever meet proficiency standards; however, showing improvement is likely possible in all schools.
Get permission from feds and start over
Pros - familiarity with the assessment; Cons lack of timely response, technology issues, poor service
A concern is that Pearson's track record is not very good. How can we be assured that there will be improvement? It will also depend on the requirements of ESEA.
Why use a contractor with a record of failure.
We should wait until ESEA is reauthorized; only do the work once.
No pro's come to mind; con's may have failed Wyoming's students and they deserve better.
Con: Pearson consistently over promises and under delivers their product. Their training and communication is poor
Someone needs to make the assessment shorter in time and valid.
Not qualified to comment.
Pros -Psychometric ability; Cons - Still too much focus on multiply choice items
Credibility has been destroyed this past year.
We prefer NWEA over Pearson.
Pro - Traffic Signal Report is valuable to adjust instruction design and delivery. Individual student reports are comprehensive and easy for everyone to understand. PAWS provides a lexile and quantile score to identify reading and math levels, respectively. Grade 10 students can take the test and "bank" their scores. Con - Timeline for delivery of data prevents effective use/planning. Pearson gives the impression that technology concerns as the norm, rather than issues that need to be addressed immediately. Assessments are extremely time consuming – way too long. The need for field test items for Pearson is very demanding and requires a student to spend much more time on assessments. For example, a field test on a writing prompt extends the test time 1/3rd beyond the actual test. PAWS is merely a "snapshot" of student performance. PAWS assumes that student should be "proficient" at a certain grade level. There is no way to measure student academic growth during the school year. Pearson provided a number of field test items in the test – some of that is understandable, but a field test writing prompt was included in the test, the test became rather cumbersome. I do not feel that Pearson has the commitment to take the assessments seriously, which was quite evident when the technology issues began to plague many schools. An e-mail was sent to our district indicating that Pearson would solve the major technology issue during the Summer."
The horse is dead; get off and get a new one. Save some dollars and pride. Take a serious look at Map
It hasn't been a positive experience why would we keep the current contractor for.
The pros would be the work has already begun and we have an assessment to begin the work. The Cons would be we are moving to the common core standards and this assessment is not aligned with that, so, we will have to rewrite the assessment anyway.
We are small fee Pearson. I do not think that they provide us what we need because they have much more lucrative contracts to focus on.
The legislature appears to want more accountability from educators. We should cancel the contract with NCS Pearson and get an assessment system that legislators and educators can agree is a meaningful assessment of

student achievement.
 Pearson seems to have some issues with the testing process and is not a reliable testing source. We need to develop a good testing system and stick with it for more than 1 year.

Question 11b	For Wyoming, what are the pros and cons of the following options pending reauthorization of the Elementary and Secondary Education Act (ESEA): • Using the current contractor to begin amending PAWS in anticipation of ESEA reauthorization (e.g. beginning to consider growth models, formative type components to the assessment, adaptive items, and including Race to the Top criteria).
---------------------	---

Comment Section

N/A
See comments above...Need timely results.
See above comments.
No. See Above.
Same
We do not believe the current vendor could deliver the high quality process and product that we expect in Wyoming.
Can we stay with a proven "test" for a while.
N/A
?
We are already on the hook with them so I think we need to move ahead with them and finish our contract. Then after that we need to start looking for another company.
Pro- Growth models actually reflect what students individual growth is. Con- Not certain of the ability of the contractor to administer this type of test.
We should not try to be "in front" of the pack in anticipation what the law "could be". We need to be patient and work towards an assessment that will give us valuable information in the classroom as a state.
N/A
How many more millions are we going to throw in this hole. I don't know where the problem is - the contractor on WDE but I fear that we are financially indebted to this company to such a extent we may never break away.
We need to move toward a growth model assessment. We would be able to determine the effectiveness of teachers based on the growth of their students. When low performing students make adequate growth, teachers wouldn't necessarily be penalized because of their lower proficiency levels. Also, a growth model would provide meaningful and valid data that would inform instruction and then create a model of continuous improvement.
N/A
I have concerns with acting in anticipation of the reauthorization of ESEA. If states continue to respond to what they think will be required in reauthorization, which is greater federal control and loss of state control, we pave the way for the reauthorization to impose regulation on states that might otherwise have been negotiated out in the federal policy making process. Federal involvement in formative assessments and further regulations that come with Race to the Top criteria will be an easier sell in reauthorization if federal legislators can claim states are already doing it anyway.
Again, no to NCS. It might make sense to explore those options but we need to know the guidelines tied to reauthorization before we sign multi-year/multi-million dollar contracts. Two wrongs (NCS) don't make a right, right.
This is not their area. MAPS already does this very well.
Maybe
Again, regardless of the way it is painted, doubt is in the air.
The current contractor has no credibility.
It's hard to make a judgment of the pros and cons of any program until the reauthorization takes place
For growth models and adaptive items – type assessments with formative components, this current contractor would be a poor choice. NWEA would work for these types

This seems to be the best option in my opinion. Look at what is available and develop our test for Wyoming students and make it meaningful by attaching to Hathaway or something similar
The current contractor could meet this expectation but the test would need a major overhaul to become a growth model. It might be more appropriate to rebuild the assessment to better meet the needs of a growth model than to remodel the current test.
I am all for growth models
Wait until the reauthorization is done, then do it right.
Same as above
Their track record is not very good. How can we be assured that there will be improvement? We should wait to see what ESEA will require.
How much additional time and Wyoming dollars would this cost? Is there a better contractor?
We should wait until ESEA is reauthorized; only do the work once.
No Pro's; Con's they are beginning to consider? We need to go with a proficient company in this area such as NWEA and stop pouring funds into Pearson's.
Again, Con: Pearson consistently over promises and under delivers their product. Their training and communication is poor
N/A
N/A
Pros - Familiarity with the state.; Cons - Would need more state input and clearer expectations for service.
PAWS is not formative and I question if simply amending it would make it formative. Can the feedback from PAWS ever be timely?
We prefer NWEA over Pearson. We don't believe Pearson will get us our results fast enough to use for formative.
Using Pearson as the vendor for PAWS would be a mistake. When Pearson purchased Harcourt Assessment there was the feeling that the mistakes that Harcourt had been making with PAWS would be addressed since Pearson was known for quality work. This was not the case – the mistakes continued – the same employees were “running the show,” and problems continued to mount.
No
No
I am not sure what the experience level of the current contractor is in the area of growth model assessment. I believe that is where we want to be, I am just not sure the current contractor can get us there.
N/A
The last thing we need from NCS Pearson is a new edition of PAWS. This contractor has consistently demonstrated an inability and unwillingness to listen to Wyoming educators. We need to switch to an assessment system that educators trust.
No, MAP already has a growth model with formative components.

Question 11c	For Wyoming, what are the pros and cons of the following options pending reauthorization of the Elementary and Secondary Education Act (ESEA): • Joining a consortium such as the Smarter Balanced Assessment Consortium in anticipation of ESEA reauthorization.
Comment Section	
No	
No Comment	
Not needed since we already have MAP	
This would be worth exploring.	
Pros - Consistency	
We are not familiar with this consortium.	
Worth Investigating	
N/A	
?	

No Comment
Not aware of what this consortium does but anticipating ESEA be a tricky situation.
It concerns me that Wyoming may resort to joining a consortium of this nature. Once this occurs we "may" be working towards a National assessment, (which we already have in place in the assessments of NAEP & ACT) Our state assessment should reflect the state goals of what we want our education to look like.
N/A
This could be one of those alternatives we look into. It would be nice to collaborate with others for essential testing - could be cheaper - more consistent - could ultimately provide ongoing formative assessments of standards of mastery instead of just one annual summative assessment.
I have no knowledge of this consortium
N/A
I have spent considerable time reading the Smarter Balanced Assessment Consortium documents and MOU, as well as portions of the 1200+ page Race to the Top grant they have submitted to the USDE. While the summary of the philosophy and work of the consortium sounds appealing, I have a list of concerns when reading the details described in the documents mentioned above. I have emails from WDE staff stating Wyoming is joining the consortium, and I responded with sincere disappointment that this direction would be taken without discussions with districts, local school boards, and the Wyoming public. Having read these documents, a decision of this kind takes Wyoming education in a substantially different direction than what the people of Wyoming have known and to do so without their knowledge or the knowledge of their locally elected education representatives is wrong.
Wyoming should seriously consider joining SBAC
I would be in favor of this if we have equal say.
Maybe
N/A
N/A
It's hard to make a judgment of the pros and cons of any program until the reauthorization takes place
Pro: Economy of scale and possibility of an innovative assessment being developed. Con: We would lose our focus on big, important skills that we currently measure; we lose local (Wyoming) control.
National consortiums scare me to death. One size shoe does not fit everyone and that is what happens in those kinds of things. We are a very unique state in terms of population, geography and diversity. We need to do what is best for us and not for someone else and hope we fit in.
This has merits but we need to have a more acting voice in whatever the consortium is. All states involved should be able to pilot the test and build it.
I am not versed well enough to have an opinion regarding this option.
Wait so we are not doing double work.
Pro - better assessment Cons - Time involved
If we radically change our state standards by adopting the Common Core it might make sense to join a consortium that is looking at assessments for the same standards.
No Opinion. Don't know what that consortium is.
N/A
Pro's - More minds working through solutions. Con's Where everybody thinks alike, no one thinks much - Lipman
Pro, Being a part of the development of the national assessment would benefit Wyoming students.
I think this would be a good idea
Would need more information
Pros - Will give a better connection to other states and a better balanced assessment system.; Cons - I can't see any.
The advantages of this approach would be it seems to mesh well with the national standards and race to the top.
We don't understand the implications - No - what's good for one state is not always good for Wyoming.
Joining a consortium might be a good option. A concern might be if the assessments address what Wyoming educators would like for students to know, understand, and be able to do. The ELL test (WIDA – ACCESS) was started this year for English language learners, and WIDA is a large consortium with the same standards for all

states involved. The test is reliable and consistent. Possibly by having a larger population taking the same test, the field test items would not be so cumbersome and the testing company might have a greater commitment to produce immediate results and quality administration protocol.
More money and more opinion. The expertise is here in Wyoming under contract - use them: Wyoming Educators.
Possibly
This could be more cost effective than the current system. It would also allow us to compare our students to other states in the consortium. Alignment would have to take place at the curricular and the assessment level.
N/A
Joining this consortium has some real appeal. First of all, we could work with a contractor in which we would have some confidence. Second, we could collaborate with other states and have a measure that would create meaningful state to state comparisons. This would allow the state legislature to have confidence in the significant investments they have in public schooling in Wyoming. This would be a test of accountability and Wyoming educators would rise to the challenge. But the devil is in the details. Staff from the Wyoming Department of Education has indicated this consortium has already been joined without meaningful consultation with district representation. If this membership is to be genuinely favored by Wyoming educators they should be represented in the discussion process.
For larger districts that have the time and money to hire people to attend and participate it would be ideal. However, for smaller districts this just adds more work to already thin stretched staff.

Question 11d	For Wyoming, what are the pros and cons of the following options pending reauthorization of the Elementary and Secondary Education Act (ESEA): • Canceling the current contract for cause, and reissuing an RFP to meet current NCLB requirements.
Comment Section	
Wait for National Test	
N/A	
Se comments below	
Absolutely!	
Pros - Cut our losses early and move on to something that will work the 1st time and won't keep changing.	
This should be a strong consideration.	
Yes	
N/A	
?	
It would be very hard to do as we have so much time and money invested with Pearson. It would only create more confusion within K-12 education.	
There is no certainty of where the federal government is going with this one	
In my opinion this is probably the best scenario we could work towards. AT the present, Pearson and Paws do not have a good reputation in the state. Starting fresh with a new RFP, and stakeholders would (more than likely) give our state assessment a "fresh start" in the eyes of the people.	
N/A	
Waste of time and money at this point. Finish what we have in spite of its problems until the new field is identified.	
This is a frightening concept to me. Starting over we would lose everything we've gained. I would support it if we could move toward the development of a growth model.	
N/A	
I don't have enough information to take a position on this but would ask, in the long run, what purpose would it serve students in Wyoming?	
If the contract is cancelled how will Wyoming meet NCLB requirements in 2011-2012? Is there really time enough to change vendors?	
I agree with the canceling of contract but feel it would be a waste of resources if we did not wait for	

reauthorization.
Probably the next move.
I don't think you will find cause for canceling. The issues weren't always with Pearson but they will hold much of the blame. I just believe that a pre-read of the assessment was completed someplace and the gaps were always there.
Yes
It's hard to make a judgment of the pros and cons of any program until the reauthorization takes place
Pro: good idea in terms of holding contractors accountable Con: could a new contractor have a quality test in time?
Seems to be a big waste of time and money and puts us behind the eight ball for data for this school year. Much like shutting the door after the cow gets out
No – this would be a step backward. To meet the future makes sense but to remodel for the current NCLB knowing that there are major changes coming does not make any sense.
I do not see how starting over is a viable alternative to fixing the current problems. It would seem to me that the resources available could be better used in fixing the problem rather than beginning again. I was not in favor of dumping the previous assessment provider. I still believe that WyCAS was a better measure of student performance than PAWS, although I do understand the issue of banking test scores and the overall costs of the WyCAS.
Schools need stability.
Same as above
I can't think of a reason why this would make much sense. It would take 2-3 years to get back to where we are.
Is NCLB still the standard to be met? Didn't President Obama say that whatever new reform to place " it won't be called NCLB? My point; other than his "blue print" for success and the push for common core standards, what's been done?
N/A
Cancel for cause and negotiate with feds an exemption in order to RFP for the absolute minimum needed to meet the reauthorization ESEA - then put resources into an assessment system which serves Wyoming's students, parents, schools, and communities.
Con, it is not feasible to have another vendor attempt to develop the federally required state standards based assessment by next spring, and rushing to production would compromise reliability and validity.
N/A
Too disruptive
This has to be done
Would we be able to have anything in place by spring of 2011.
We should cancel
There are a number of reasons for canceling – Pearson cannot keep up. Having been involved with the original RFP a number of years ago, it seems as if many of the criteria in the RFP have disappeared. For example...student data was to be available and useable immediately. The 6-month delivery seems far too long. By the time data is available, teachers have had to resort to data from other testing instruments and PAWS data from previous years. The idea of online testing was to get "quick" and useable results.
Possible. Do we have to go buy something?
Possibly
I do not believe there is time to do this. By the time an RFP is completed and the contract awarded, then the new assessment developed would probably put us to where our current contract expires.
N/A
Absolutely
They have not proven themselves to be reliable. We need to look at a better system or use one that is currently in place. Teaching time is being wasted.

Question 11e	For Wyoming, what are the pros and cons of the following options pending reauthorization of the Elementary and Secondary Education Act (ESEA): • Canceling the current contract for cause, and reissuing an RFP in anticipation of ESEA reauthorization.
Comment Section	
N/A	
How do you anticipate?	
Since it is unknown when ESEA will be re-authorized MAP should be used now.	
Yes	
Pros - switching to a more valid and reliable test; Cons - money lost	
This should be a strong consideration.	
Yes	
N/A	
?	
I think we need finish the current contract and get ready for the RFP in anticipation of ESEA reauthorization	
Again anticipating ESEA action so a gamble at best.	
Again, in my opinion, we shouldn't be working to "anticipate" what the federal government is "going to do".	
N/A	
Same as above.	
See previous answer.	
N/A	
Opposed. Please see previous comments.	
Same response to scenario provided above—a real concern as to how we meet NCLB requirements.	
I agree	
N/A	
Again, how?	
Yes	
It's hard to make a judgment of the pros and cons of any program until the reauthorization takes place	
Pro: offers opportunities to focus on innovative elements in ESEA Con: We are subjecting Wyoming school accountability to the whims of state and national politics.	
Same	
We need to start working on the next RFP tied to National Common Core Standards and the new vision of the ESEA reauthorization.	
I have not opinion as to the viability of this option	
Wait until reauthorization	
Yes	
I can't think of a reason why this would make much sense.	
Probably a better option than previous question.	
N/A	
See Above	
Con, again it is not feasible to have another vendor attempt to develop the federally required state standards based assessment by next spring, and rushing to production would compromise reliability and validity.	
N/A	
N/A	
This should be done	
This timeline seems more reasonable	
We favor this	
The concept of PAWS is good. Educators are beginning to realize that the use of student data, both aggregated and disaggregated, is extremely valuable to analyze individual student results and examine instructional	

effectiveness. Whether ESEA mandates a state (or national) assessment or not...the data from these tests are quite helpful to improve student achievement.
I would wait and see.
N/A
Before any funds are used to develop a new assessment I believe we need to have a clear understanding of what ESEA looks like. Make changes mid-stream end up costing the state additional dollars that could go for other areas of education
N/A
Absolutely
See Above.

Question 12	In your opinion, should Harcourt Assessments, Inc., have been monetarily penalized for problems in 2006 and 2007 for problems with reporting and scoring?
--------------------	---

Comment Section	
N/A	
They have always been breakdowns in the information and accuracy.	
N/A	
N/A	
There's bound to be little problems	
N/A	
We are also responsible for better oversight - why did we allow an untested platform to be administered.	
What purpose would it serve	
N/A	
I'm not fully abreast of that situation and therefore can not comment.	
If a company does inadequate work they need to fix it or be fined	
I would not be a good judge to decide what occurred in 2006 & 2007. I do not know exactly "what" problems occurred in 2006 & 2007, so therefore do not know if they should be penalized. However, IF there were major problems in 2006 & 2007 I am wondering why we didn't address these then. Maybe if that would have occurred we wouldn't have had the problems we did this last year with PAWS.	
N/A	
I think Wed was equally to blame.	
I don't care, this is a legal issue that should be decided by the WDE's attorneys. It should have been dealt with long ago.	
N/A	
Uncertain. The 2007 administration of the PAWS provided for a January testing window with the ability to "bank" scores. This plan was highly deliberated by the original assessment task force throughout the many months of drafting recommendations to the Wyoming legislature and the dual administration was approved by the feds for meeting NCLB requirements. In Harcourt's response to the original RFP, they ascertained they could handle the design and scoring elements. In the end, there were scoring and other issues related to the early versions of the PAWS and the mid-year administration. However, instead of working through those issues and preserving the intent to build a more instructionally supportive assessment (as envisioned by the legislature) the WDE reacted by eliminating the January opportunity and changing vendors. With the changes made by both the WDE and Harcourt, I don't have enough information on what transpired to have an informed opinion on monetary penalties.	
If the perception is that Harcourt's work from 06/07 was sub-standard, it should have been addressed at that time, not four years later.	
N/A	
N/A	
But so should have others. The overall assumption that the assessment was being completed within an acceptable time frame was erroneous. We need the information by the completion of the school year. Where was that expectation communicated? In addition the misalignment of prompts and questions being incorrectly labeled	

or classified were always noted by teachers. In general and from day one, PAWS was in trouble. Who was the lead for the overall operation, or who mandated this be the product the state would use?
We've been watching and wondering what someone has to do in their dealings with WDE to risk breach of contract. The following question has been asked for five years " Does Wyoming have so much money that we can afford to throw away millions every year on a poor product?"
I'm not in a position to make that judgment
Contractors should be held accountable for mistakes they make.
Those problems were minor and there is no perfect system. Instead of trying to penalize people, lets work with them to get things done right.
There were mistakes, but just as many by the state, so we move on and leave it at that. Who found the mistakes? If them, we give them credit for stepping up. If we found it, why didn't it happen sooner?
I lack an understanding of these problems so I do not have an opinion.
They need to be accountable
N/A
I don't know the specifics of the problems in 06 and 07 so I can't respond.
N/A
N/A
Did they fulfill their obligations with due diligence
N/A
N/A
N/A
N/A
They did not fulfill the contract.
Mistakes caused incredible problems
One of the errors that Wyoming made with Harcourt Assessments was not Harcourt's fault. PAWS was criticized before it was ever given. Our Wyoming school districts have difficulty with "change." Elementary teachers thought students couldn't adjust to the use of computers for online testing and wanted the assessment to be "pencil/paper" design. Some districts had not given an accurate analysis of their technology – bandwidth, etc. Taking the test more than once – winter and spring was intended for students who were "ready" for the test in the winter to take it and bank their scores so they did not have to take it in the spring. Superintendents and principals didn't seem to understand this and had all students take the winter assessment, followed by all students taking the spring assessment as well (even if students had established proficiency). This was not Harcourt's fault – it seemed like lack of understanding and poor planning by districts. Disaster resulted when many "squeaky wheels" forced WDE to change things around – the "state of the art" designed assessment, under the guidance of Jim Popham, became a mediocre test to satisfy uninformed squeaky wheels. The squeaky wheels couldn't accept change and wanted things to fit their already established routines. If we had left things alone, we might have been on top (which at that time we were according to NAEP). Instead WDE gave in, and we settled for mediocrity – this was not Harcourt's fault.
Wyoming is viewed nationally as a small group with a lot of money, full of nice people. (I hate when the outsiders are correct) Again, we will hold students accountable when too many adults are not being held accountable.
That is what they are paid to do.
Perhaps to some extent, however, it is hard to determine now knowing how the contract was written.
They are paid for a product that they did not deliver.
There needs to be some real fact-finding regarding the direction provided by the Wyoming Department of Education. It is not clear whether there is a single member of the WDE assessment staff with any real experience in statewide assessment matters. There needs to be staff with credible experience and education in statewide psychometrics. This is sorely lacking.
We do not remember the problems or issues.

Question 13	In your opinion, how should ESEA be reauthorized with respect to requirements within NCLB;
Comment Section	
Growth based Assessment	
Growth Model	
Using a growth model with MAP for AYP with graduation rates.	
The new approach of identifying persistently low achieving (PLA) schools overtime is worthy of study.	
I need to look at more than 4 years & taking a look at IEP's.	
The AYP model is one that is grounded in failure. Eventually, every school will be unable to make AYP. We need a fair, consistent assessment with appropriate targets. Not all students will be proficient on the statewide assessment. We believe that NCLB is designed to rescue students from the most underperforming schools in America....those schools do not exist in Wyoming. Let's focus on the continual improvement of student achievement, celebrate the great schools in the state of Wyoming, and work to ensure a cohesive partnership between the legislature, WDE, and districts that promotes and positively portrays the excellence of Wyoming schools.	
Incorporate growth model, special needs students should be assessed based on IEP students.	
N/A	
?	
We need a growth model.	
I believe a growth model is good, graduation rates (but old model does not work for alternative high schools)	
The biggest aspect of reauthorization and measuring student achievement would be to use a growth model. Our current assessment does not take into account "where" a student starts and where they end up during an academic year. Many of our students enter the system 2 – 3 years behind in their academic preparedness. Our staff can work towards getting the student on track given the 175 contact days, however they are (not necessarily) going to achieve 2-3 years of academic achievement in that time. However our students DO gain in their learning, but are not given the chance to be recognized for this gain.	
N/A	
I think we need a growth model so we can get back to preparing students for life and not just to take a test once a year.	
Again, judge the performance of schools by the growth we make. Allow districts that are successful the flexibility and the autonomy to do what they need to do. The Dept. of Ed at both the fed. And state lever should take on a supportive role rather than a punitive one.	
AYP is a fundamentally flawed model. It guarantees the eventual failure of all schools.	
I have always supported the philosophy of a component of individual student growth in making AYP determinations but that is not to say I support making the PAWS a growth model assessment or replacing PAWS with the MAP assessment. People need to realize that a quality growth assessment that measures any level of deeper cognitive learning other than vertical skills (such as those measured by MAP) means more hours testing and, depending on the consequences for not making AYP, may not be in the best interest of individual students. The formula for determining graduation rate for AYP needs to be revised. A fifth year graduate is not a drop out. Overcoming whatever issues or barriers that contributed to the student attending a 5 th year needs to be acknowledged as a success both for the individual and the school system.	
N/A	
An improvement model should be used. MAP can achieve accurately tell us. How much a student should grow within a year. This should be used as our basis.	
Comments growth model for students. Quit measuring different groups of kids. It does not help.	
I think that some consideration needs to be tied to schools that are working with high "at-risk" populations. If the growth component with an assessment shows a growth factor over a year then regardless of AYP calculations the school or the district are moving forward. An arbitrary set percentage isn't going to help the actual individual for years to come and only after they actually exit the school. All things grow and become better and beautiful. This	

is a human we are working with, not a product that is made from parts of wood, concrete or steel.
N/A
I'm not in a position to make that judgment
ESEA reauthorization should consider status, improvement across cohorts, and growth over time for individual students. Attention should be paid to state-to-state comparability. Using effect sizes instead of performance levels would help with this. We want AYP to be more valid by identifying the proper schools so that schools that do well on AYP really are performing well. And schools that do not do well on AYP really are the schools that need more help and get more help.
WE have very few students who will ever be 100% proficient in all areas. Statistically that is not possible and yet that is what AYP says has to happen. No one wants to leave a child behind, but the reality is that not every child is going to be 100% proficient in every subject. We need to remember that we are dealing with people and not widgets. Lets reach the whole child and I think the country would be better off.
We need to consider a growth model so we can evaluate all students and better utilize the assessment as an evaluation tool for instructional purposes.
I stated in my first comment that I am not opposed to concept of school and district wide accountability. I do believe that there does need to be reasonable expectations regarding student performance for every school in this nation. I have also stated that AYP has little validity when the standards and performance levels are not the same from state to state. I would also clearly state that expecting all students to be proficient in 2014 is an absolutely ludicrous goal; it cannot be achieved when all students regardless of ability or disability are expected to reach such a unreachable goal.
Focus on all students, work toward individual growth.
N/A
Student growth; flexibility, trends over time; move to a support model not just a status model.
It would be great if the entire ESEA process was less complex.
I would like to see a growth model. Flexibility needs to be provided for students with IEP's (Elementary). I think many school improvements can be directly linked to NCLB and holding schools accountable for achievement. Schools should demonstrate growth toward proficiency. I think the National common core standards need to be supported by state exams aligned to the standards if we truly want norm referenced results (High School).
The whole law needs to be revised. Local communities need to serve their youth, states serve their communities and the Fed's serve the states. All in a relationship in lieu of the fed's putting the responsibility on local people to grow their children.
Any new model should include both measures of status and growth
N/A
N/A
This should be done a more sophisticated growth model built on the idea of progress toward the career and college readiness.
Graduation rates should be calculated over a 5 year span when determining AYP. A 4 year span makes no sense. It took me 5 years to get through college and things have turned out well. We have students who take 5 years to get through high school and they do well. What is so logical about a four year span.
Use a growth model
Since we know more about the value of growth model assessments now, it would be fitting to structure AYP around student growth rather than assuming that age and grade are contiguous. When one recognizes that physical attributes of an individual are directly related to brain development and achievement, then individual student achievement will be accurately measured. If teachers design instruction according to student cognitive needs, teachers will be much more effective in the classroom. For example, a student in the 4 th grade might not be as academically mature as his 4 th grade peer sitting beside him at a certain point in time. This does not mean that the student will not achieve – it means that it might take more time for him to achieve or his brain might be ready to achieve a few months later. Instead we assume that at a certain date and time, every student must be at the same landmark level (or higher). If the growth model is used, educators would get a better understanding of what the student actually did achieve and the next steps to continue to get better. This would also give the teacher some direction in developing lessons to meet individual student needs and improve whole group instruction.

My example would be MAP three to four times a year; Measure - intervene if needed and look if needed and look at overall student performance and growth of all children.
N/A
I would support a growth model for AYP. Students come in spread across a large curve in kindergarten. Our job should be to grow them to be college and career ready when they leave high school. If schools are closing the gap and moving to that level of learning then they should not be penalized because they were not there by grade three.
N/A
Of course, we need to move to a balanced system that has elements of proficiency status and growth. The legislature (and parents) have a vested interest in knowing whether Wyoming students have reached proficiency or are on a path to proficiency. Only a growth model could do this. In addition, there needs to be a value added component to gauge the effectiveness of teachers in the classroom. The reauthorization will doubtless have language moving from "highly qualified teachers" to "highly effective teachers and leaders". We should be designing an assessment system to accomplish this mandate before it occurs.
NCLB needs to be adjusted and more realistic.

Question 14	What advice, would you give to the Superintendent of Public Instruction and the Wyoming Department of Education with respect to anticipated changes to PAWS, as Wyoming waits for the pending reauthorization of the Elementary and Secondary Education Act (ESEA)?
Comment Section	
Do nothing	
Use MAP in the mean time.	
Use MAP now and get rid of PAWS	
Bring back Wycas	
Go to MAP's	
Support Wyoming educators and recognize the excellence in our school systems.	
Be absolutely sure of its delivery before authorizing its use. Shorten the assessment. Get results to the district quicker. Ensure each district has data that can be used to inform constituents.	
N/A	
Make sure districts have no problems with administration	
Push the feds for a growth model.	
My understanding is the PAWS will be paper and pencil next year - shorten it and eliminate the computer part so that doesn't play into it and wait to see what happens	
To the Superintendent of Public Instruction, the WDE, members of the legislator and the community: We need to get it right the next time. We need to involve all of the stakeholders in the development of the assessment. It needs to be transparent, truly assess the learning which takes place and gives accurate information back to the people in the state. When the scores are published we need to believe in them. When mistakes are made, they need to be corrected so a true picture of what is occurring is reflected in the scores. I don't believe we have done a good job in this area. Schools need to be accountable, and a good assessment will reflect that accountability. Bad assessments don't tell the entire story and give the picture that schools aren't doing their jobs. Enough is enough.	
We need credibility and reliability in our state assessment.	
lets stay the course - quit making changes	
Get more collaborative! Let the school people in our state be involved in the decisions. Come prepared with parameters that are based in current law and let us help find the solutions. Don't do it to us – do it with us. At the very least....keep us informed of the process.	
Regardless of the pressures brought to bear on the Wyoming education system, there is a need for valid and reliable test data. Given that the Feds are apparently incapable of defining a direction or know what effective assessment should look like. We need to design an assessment system that serves the students and the larger education community. Accountability as part of a feedback system that helps us improve and get better would be	

welcome. The current punitive model with sanctions and heavy handed compliance mandates has been historically and repeatedly proven to be dysfunctional.
Carefully assess what is already in place and is working in Wyoming. In conjunction with Wyoming educators and leading experts in the field of measurement, work to improve the PAWS assessment. Do not lower expectations for student achievement by designing an accountability system that lacks rigor. However, do work to educate the public and lawmakers on the appropriate use, limitations, and pitfalls of misinterpretations of assessment data. Consider the use of MAP in conjunction with the PAWS but do not mandate a system that requires MAP testing and retesting of students who have already demonstrated proficiency on the PAWS.
N/A
Do not change the format to paper and pencil. Fix the problems with Pearson until we know exactly where we have to go.
Use technology it is the way of the world. Reduce test time by half. Ask for a one month turnaround on scores. The test has good questions!!! They are aligned to our curriculum.
Let's see who that will be before making this request. Both candidates are similar but very different. As for the Department, get copies of future assessments form Pearson, make sure they are aligned to the standards and benchmarks of Wyoming. Listen to individuals and if an assessment item is questioned contact Pearson to remove the item from the scoring matrix. In short heed your own requirement when working with district on their common core assessment expectations.
N/A
I have no idea what the reauthorization is going to look like. I would support an assessment that focuses on a growth model and provides results in a timely manner. I do not believe the assessment needs to consume the amount of time that it does; taking time away from teaching and learning.
Keep the design construct the same. Focus on effective delivery and reporting to increase the confidence of the assessment. This is about customer service and increasing credibility.
Quit using PAWS as a political lever and work to make it more useable and student friendly. Fix the tech problems that seem to plague large rural districts because of connectivity and get on with the process.
We need to consider moving back to a test with the rigor of the WyCAS. That test represented more challenge and raised the bar for success of students in Wyoming. Setting the bar low is not acceptable.
I stated in my first comment that I am not opposed to concept of school and district wide accountability. I do believe that there does need to be reasonable expectations regarding student performance for every school in this nation. I have also stated that AYP has little validity when the standards and performance levels are not the same from state to state. I would also clearly state that expecting all students to be proficient in 2014 is an absolutely ludicrous goal; it cannot be achieved when all students regardless of ability or disability are expected to reach such a unreachable goal.
Keep it the same and plan after the reauthorization
N/A
Address some of the credibility issues associated with PAWS and keep it in place while beginning the discussion of developing a growth model or accepting a national test. Have a point person in the department monitor the proposed ESEA changes and share those on a regular basis with educators in Wyoming. Convene a "design think team" to meet regularly to do scenario building of "what if... we go to a growth index, an adaptive test, a national test..." That way there can be some preliminary work done at the state level. Collaborate with Wyoming teachers, superintendents, principals, local Boards of Trustees in designing a coordinated and comprehensive accountability and assessment system that flows to and from the various levels: state, district, school, classroom, and student. Begin building a robust data system that is capable of gathering and reporting student achievement on a wide range of measures.
No opinion on this question.
I would like to see the State wait for the ESEA reauthorization before making any significant changes to PAWS. Let's not put ourselves in a state of perpetual change.
"Talk less and say more" "know where to draw the line" and "ride for the brand" when it comes to Wyoming's children. Let's focus and serve their needs first and less on mandates which only serve to distract from learning and serve and political and ideological masters.
Join a consortium working on the new national test, and reconsider using the 2010 administration of PAWS for

AYP calculation now that the third party evaluators have determined the results to be valid and reliable.
N/A
Make PAWS meet the minimum federal standards
I believe PAWS should be thrown out. We should start anticipating a focus on college and career readiness and how assessment system from K-12 can tell us that.
Think big picture/ long term.
move forward anticipating reauthorization, switch to NWEA
I would not continue with Pearson. Until ESEA is reauthorized, I would appeal to the feds and ask that we provide a growth model assessment (i.e. MAP) and use My Access for writing. MAP is aligned to the present Wyoming standards (which will be aligned to the Common Core State Standards soon). My Access has hundreds of researched based writing prompts at all grade levels. Student work is assessed electronically with consistency. Students could be given a pre-test (Fall) and a post-test (Spring), and data would be collected to show growth over time, during the school year. The data would provide essential information for teachers to develop plans for student achievement and to adjust instructional practices. The results would be immediate and useable for the next school year. AYP could be based on Adequate Yearly Progress – exactly what AYP means, measuring if the student made adequate progress over one year of schooling.
Look at some of the framework from Kansas!!! Keep it simple.
N/A
I would suggest they proceed with caution and not commit dollars to develop an assessment that may not meet the requirements by the time it is ready to roll out.
Any assessment that is used needs to be valid, timely in regards to length of test time and reporting results, and growth model with formative components.
We should advise the new state Superintendent of Public Instruction to cancel the contract with NCS Pearson for cause. We should ask for an expedited waiver to use MAP testing as an interim statewide assessment until the consortium assessments are developed and deployed. We should also demand that the Superintendent hires assessment staff with direct experience in statewide assessment matters.
If we are going to stay with PAWS, make it more reliable and valid test. Shorten it and make it less redundant.

Question 15	15. What is your opinion of states vying for Race to the Top Grants, as well as becoming members of consortia such as the Smarter Balanced Assessment Consortium?
Comment Section	
Does not apply, buying into the latest program in not the answer	
Not supportive of competitive grants	
Both have too many federal strings and we should not participate in either.	
Wyoming will never be successful with these types of initiatives until the legislature has the courage to reform its continuing contract laws. Wyoming needs to follow the example set by Colorado's legislature as it wrestled with the antiquated practices of granting tenure to teachers.	
Anything to help students.	
We would not recommend applying for Race to the Top Grants. Again, we are not familiar with the Smarter Balanced Assessment Consortium.	
N/A	
I am ok with this.	
?	
Not a supporter of more governmental interference	
I think rural states are highly unlikely to get money. As indicated the consortium is something I am not familiar with.	
Once we take this money, and/or become members of this consortium we have given away any say in what our education looks like. We will be tied to what the Federal Government rules education to look like. Wyoming is a rural state. If we haven't learned yet that a good many requirements of NCLB just isn't workable in this state, we never will. Wyoming people know what they want by educating their children.	

N/A
Not interested in RTTT grants except as they work together to produce common assessments for states in the CORE areas of both formative and summative assessments of content standards mastery.
I don't have a lot of knowledge about this but I have heard from other superintendents that the strings attached to the Race to the Top Grants are cumbersome. We are so lucky that our schools are funded extremely well by the state so grant funds are often not a necessity. If I took the time to know more I might be more open to it - especially if the benefits outweigh the strings.
Race to the Top is a castigatory model intended to destroy education.
As stated in earlier comments, such action without discussions with the citizens of Wyoming on the pros and cons of participation is wrong. If voluntarily taking on the regulations tied to the Race to the Top grants and/or joining either of the two consortiums that have been established (SBAC and PARCC) is good for Wyoming, then it is worth the time to properly inform and have discussions with the public and local school boards.
RTTT is not the best model for encouraging and supporting school improvement. It isn't likely that the federal education system can improve local schools more efficiently or effectively than the local educators. It probably takes both or all three levels (federal, state, local) but they need to know their roles. The federal and state system should encourage and assist with local educators in becoming better.
I am not in favor of the race to the top because we lose local control. I think Wyoming would fair well on both a regional and national level.
N/A
N/A
N/A
I have no need for the Race to the Top Grant and the strings that are attached to it.
It would be foolish not to participate in attempt that ultimately will affect us in one way or another. Being part of the conversation helps to shape that conversation.
I think it is a waste of time and money. As I said earlier, one size shoe does not fit everyone and I think Wyoming always gets beat up in those deals.
The race to the top grants were not at the level needed to be considered a contender. We missed the boat and should have gone a different direction. As for the consortium, we are fine with it if there are changes so all are represented more equally.
As to Race to the Top my response would be why bother. As to SBAC I have no opinion.
Don't do it! Too many strings.
N/A
I support the use of common core standards and therefore it makes sense to help each other design accurate assessments. If Wyoming wants to vie for grant money then the WDE needs to have the people in place who have the knowledge and skills to coordinate those efforts.
Probably politically motivated. I am skeptical of how decisions are made as to which states qualified for RTT grants, and those that did not. Why a competitive grant formula, some win/ some don't. How about the Federal gov't adequately fund all school districts on some sort of equal basis.
I worry about the strings attached to RTTT (Elementary). I believe there are challenges with the Race to the Top Grant requirements. I do not know enough about the Smarter Balanced Assessment Consortium to recommend or not to recommend joining the consortium (High School).
If that's the best solutions they can hope for then more power to them. Wyoming has more options and doesn't need to play in that arena, we can make our own way and focus our energy and resources at home to build capacities for thinking and learning which serve Wyoming.
If the funds and school turnaround options would benefit students then Wyoming should apply for the funds.
N/A
N/A
I believe we should vie for the money and be bold in our ideas. I think a larger consortium would help us not be so insulted, and force us to have more common benchmarks of success.
If we are headed towards national standards in the state of Wyoming the n SBAC makes some sense. As the leader of education in my community I believe I need to consider all funding sources, including RTTT monies.
We don't like it. Money pushes us in directions that may not be good for us.

Race to the Top grants could be quite political. In our state we have so few Charter Schools that our chances are rather nil. Plus...our state does quite well on NAEP – isn't Race to the Top targeting states who aren't doing well? The concept is good, but as with many educational programs where we want our lower scoring students to improve, we don't seem to have the time, energy, or finances to encourage advanced students to improve. A concern could be if the stakes get too high that there might be political issues to overcome.
I think the race to the top is also the race to the bottom of the budget. We are not California, Texas, and New York.
N/A
I would support the assessment consortium. Sharing the cost and the expertise is good use of resources. However, the work must still align to our state standards and not be dictated by larger states which more students.
Districts should have been consulted.
The Race to the Top grants are accountability-driven with mandates for significant changes in continuing contract status and incentive pay. While these are not necessarily bad ideas, they should emanate from Wyoming citizens and educators. As for the multistate Assessment Consortium, this could be a real improvement for Wyoming education, in terms of legitimate accountability assessment. But there needs to be far greater representation among Wyoming educators and other stakeholders in the consortium adoption process. There was precious little input in adopting the Common Core Standards.
Race to the Top is more of a teacher incentive program and could cause teachers to become more selective of students they teach and classes. Again for larger districts it could be a good thing for smaller a burden on the staff and administration.

Question 16	What is your opinion of Wyoming adopting a growth model in concert with meeting current AYP?
Comment Section	
Yes, it is the only thing that makes sense	
Supportive	
Absolutely	
This is worth exploring.	
I think growth needs to be looked at as opposed to current Yap requirements. Goal should be for kids to grow - some are never going to make AYP.	
The options for a growth model should be considered	
Much better than not	
I would prefer a growth model.	
?	
Strongly support those efforts.	
Since the lowest achieving students can demonstrate growth the (Map test) I believe this model is most agreeable for all learners.	
As stated in many of my answers above, we need to be moving towards this concept. I'm not sure why Wyoming has waited so long in looking at this type of assessment for our state. If we truly want to be on the "cutting edge" we should have been working on this years ago.	
N/A	
Should be obvious by this point that I am a very strong advocate.	
YES!! I have been talking to J. McBride about serving on a task force to look at a growth model for Wyoming for the last 5 years....it really needs to be considered and we have many educators in our state that would volunteer to be involved in the process.	
We have advocated for a growth model for years. Our appeal has fallen on deaf ears.	
Supportive – as long as the PAWS isn't vertically scaled and then called a growth assessment.	
RTTT is not the best model for encouraging and supporting school improvement. It isn't likely that the federal education system can improve local schools more efficiently or effectively than the local educators. It probably	

takes both or all three levels (federal, state, local) but they need to know their roles. The federal and state system should encourage and assist with local educators in becoming better.
In am in favor of a growth model. It accurately reflects what is happening in classrooms and schools.
ok with growth model
Interested but I am not sure of the expectations of that growth model. Current growth models work as long as outside pressures don't force and expected growth to happen. Students may only grow and incremental they can handle and then suddenly break forth. More than likely someone will expect all the students to grow and the same rate and reach the same point at the same time. Not realistic.
Yes
I support that.
It is a good idea, but only can be effective if a test is used which is designed for that purpose.
HURRY UP AND GET IT DONE! That benefits every student in Wyoming.
We think it is the best solution to the success of our instruction. If we want to see our students progress and evaluate the impact of instructions, it is necessary for our assessment to move in this direction.
That is the only solution to defining performance that makes any sense to me.
We already are in our district.
I favor a growth model
A growth model is important for students, not so much for AYP, but for their own learning.
If the growth model includes firing school administration and faculty due to some time/ occasional low scores, I would question many variables in a students life over which faculty and staff have little or no control.
I am in favor of placing an emphasis on student growth, however we want to ensure that the expected growth targets are realistic (Elementary).). I think that adopting growth models will lessen the number of schools not meeting AYP. I think it would be more accurate to compare growth for individual students in order to identify patterns and trends, than to track cohort groups of students. Using a growth model would call for individual goals rather than a universal goal (High School).
Measuring and educating growth on a continuum of learning is where we should be if we truly want to leave no child behind.
The district would support this move as long as it does not result in two accountability models, a Wyoming model and a federal model.
Good Idea
N/A
I am an advocate for growth model, if the test is worthwhile and can be vertically scaled around progress toward graduation and career, college readiness.
A growth model makes sense to teachers, parents, and students. I am in favor of it.
We like it, we look for improvement.
It would be a wise decision to move in this direction. Dr. Laurel Ballard is a phenomenal assessment person – she could make it work.
Possible; be very careful we don't sell children short on growth.
N/A
I would support the growth model for the reasons mentioned earlier. The critical goal is to graduate all students college and career ready. If students are on track to reach that goal then I believe districts and schools should be given credit for closing the gap. I believe you could do this and still maintain high graduation standards.
If the data is timely.
We should definitely move to a growth model of statewide assessment. But there is little hope that staff from the Wyoming Department of Education can manage a contract with such complexity, when they have demonstrated that they can't manage a contract with simple proficiency status determination. Neither can NCS Pearson. If we are serious about moving to a growth model we need staff and the contactor that can realistically manage such a complex assessment undertaking.
Not familiar with this concept and how it would apply to AYP.

Question 17	What is your opinion of the Common Core State Standards that Wyoming recently adopted, as well as creation of regional or a national assessment to measure student performance against those standards?
Comment Section	
Great Start	
If were going to have national standards we should have national assessment.	
The common core does not differ substantially from what we do now, however I fear a national assessment will not be meaningful.	
I would support this.	
Win situation	
We are still forming opinions about the new adoption. We believe Wyoming students would perform well on regional and/or national assessments.	
I am in favor.	
N/A	
?	
I think Wyoming can and has developed some excellent core standards. We would need to make sure that these standards could be used in a true comparison of the regional or national standards.	
I have reviewed the standards and find them to represent what students should know. Regional and national assessments stay on the side of local control so this would not be proper for our direction	
I agree our students need to be held to the same standards as those of students in other states. I am hoping that a regional or national assessment would be only "one" of several pieces of information that we would use to assess our students and educational system. Again, one size does not fit all students. We need to have a balanced look at assessment. The sure way to leave children behind in to use a one-shot assessment to tell that student (or teacher or district) whether they are "proficient" or not.	
I personally have not yet spent the time working with the new standards to have formed an opinion	
I am supportive - see previous comments	
We need the standards so that we are able to focus our instruction on what is important to our learners. Standards need to be updated and changed based on the changing world we live in. Our common core standards are already closely aligned with the national standards so it might make sense. My question is "Can this been done within a growth model?"	
They have been adopted, we will adopt them. They reduce the creativity of exceptional teachers. They reduce the education process to compliance rather than originality, inspiration, ingenuity, inventiveness, resourcefulness, and vision.	
Districts were notified of Wyoming signing the MOU for the development of the Common Core Standards via a WDE memo a month after the MOU was signed. In my opinion, that was wrong. However, I do not have near the concern over national standards as I do the development of a national curriculum and regional or national assessments. If required to take a position today, I would be against joining either of the two consortiums being formed (SBAC and PARCC). While I support the many of the philosophies behind NCLB, the regulation and design toward urban settings has not served Wyoming students well. What would make one think that turning more authority over to federal regulation would be a good move for Wyoming? More importantly, I receive feedback from district patrons that are unhappy with the impact NCLB has made on education in Wyoming and would not support moving to a more federally standardized system without their knowledge and input. Bluntly, I would have patron groups that would be up in arms over voluntarily moving toward more federal control and regulation of the education of their children.	
The Common Core State Standards are well written with rigor emphasized. It's not clear as to the value of having regional or national assessments?	
I like the common core standards and I am fine with a criterion based regional or national assessment.	

I think it is a good idea to use regional or national standards
It actually looks like this has happened. Again I don't think a student graduating from high school is ready to be an open heart specialist. But we seem to be pushing that way. How fast, how fast how soon, and at my expectation level. The Common Core Standards are workable and they will help get the attention off the weak standards currently in place. About the national assessment, watch out, it may not work across all areas of the state and the bias needs to be carefully checked.
The concept is interesting. I would want to see a proposal.
There are many questions yet to be answered in regards to this but I do support the adoption of the common core standards and I would support any assessment that aligns itself to those standards be it state or national, bearing in the mind the comment in regards to the amount of time spent on testing.
Common Core standards have potential to increase consistency across states and to raise expectations for student outcomes (rigor). It is difficult to improve when the target (goal) keeps moving.
N/A
We see Common Core Standards as a positive move and a way to raise the bar. Any new assessment must be at the quality that PAWS is or better and proficiency must be set at a high level.
I am all for a national based curriculum and a national assessment such as NAEP.
I have only just began to look at them. I believe good teaching will out so any new standards.
The community core state standards are better than our current standards. A regional or national assessment should be no problem for Wyoming and may save money, be more consistently aligned to standards, and be more timely with a response.
I support the use of Common Core State Standards and possibly a national assessment.
Favorable if it creates a nationwide standards and assessment.
It is a positive step. Great idea.
There are some merits in that comparisons would provide larger populations to work with and the standards would be focus on the essential and timeless concepts. However caution and skepticism should be used to avoid a narrowing of the curriculum and an "all think alike" mentality.
While this will provide comparability between states, any local flavor will be lost due to the fact that the standards had to be adopted "as is" and with the number of national standards already large states will hesitate to add locally developed standards.
Good Idea
From a national perspective, it makes sense to establish common standards. The common core are more explicit than current WY standards.
I believe the are focused, usable and better than the current standards in place.
The standards seem to be very well done. Feedback from teachers has been positive thus far. However, the national standards and regional /national assessments can be viewed as contrary to local control. This doesn't always set well in rural Wyoming.
Much better quality, national assessment
It is a good idea. The assessment would be a good measure of student performance. Our district uses comparative ACT data, which provides a good indicator of achievement. Although there is data that compares our district to the state, it is valuable for us to beyond state lines.
Go with national standard
N/A
The common core was a very good decision. Assessments now need to be developed that are tightly aligned the content and the cognitive level of the standards. If we can participate in the development and save funds in the long run then I would say it is a viable approach.
Local control has lost all meaning common core standards are not load, but new moving to Natl. Standards is just additional busy work that detracts from the teaching and learning should be occurring.
The Common Core Standards are generally well done. There are some significant issues where they lack clarity. For example, the math standards lack any specificity about the high school course sequence for all students and college preparatory students. Similarly, they lack detail about the role of technology in mathematics education. The real problem with the Common Core Standards is that they were adopted by the Wyoming State Board of Education without any meaningful discussion among Wyoming educators. We recognize that they had a couple

of panels who advised the state board to adopt them, but this was not thorough. There should have been real discussion and debate about this major sea change in education and once again local educators are left scrambling trying to catch-up to this mandate without involvement. The notion of a regional or national assessment to measure the standards is quite attractive. Wyoming educators would be well served by an assessment system that is designed by a competent contractor directed by competent staff. We need to catch up as a state and need to retain staff that have the expertise to lead in this critical assessment arena.

If the national assessment will replace all the other testing that is now required, we are for it.

Question 18	What is your opinion of the National Assessment of Educational Progress (NAEP) and how it is used in Wyoming?
--------------------	---

Comment Section

I would like individual data

It is viewed a "something we have to do" because we don't get builder results.
--

It is a waste of time. We do not receive any data from this assessment on individual students.
--

As it is currently used, the data is basically useless.

NAEP is basically a way to test 4th and 8th graders on how well they compare to other states, not how they do as individuals.

NAEP does not provide individual student or district level data. NAEP is not a focus in our district. We focus primarily on our in-district assessments, MAP, and PAWS.

Wycas was more aligned to NAEP and the result was better NAEP scores - PAWS was not so why are we surprised that NAEP scores are down?
--

I would prefer not to give this test. Has little meaning to us.

?

I think it has good meaning for us as a State. It allows us to measure our students with the rest of the country.

Because of the comparison with other states I believe most publications in Wyoming do not value because it test "our kids" only. I personally believe state comparisons are a good thing if they aren't high stakes.
--

It is another piece of the assessment pie. It helps us gauge what our students are doing as compared to other students in the country, but it doesn't tell us the entire story.

We have participated in NAEP for many years. It is a well run and delivered assessment program. However we as a district do not benefit by seeing the results- which ends up meaning that we spend the time without obtaining data to use.
--

It is an occasional test that helps us see how we stand on the international stage. I just am not sure how reservation schools seem to always be selected in a pool drawing.
--

For the last 5 years, my little district and schools are always "randomly selected" to take the NAEP. It tells us how we do as a state compared to other states. I don't mind participating but the data is not very useful to us as we receive no student level results.

The NAEP takes lots of time and energy and yields no local data that helps us educate children.

Given the small student population of the state, most districts in Wyoming are selected and required to give the NAEP assessment every year. Scores provide data for tracking student achievement at the national level and can be used for state comparisons. However, the assessment is not aligned to Wyoming Content & Performance Standards. District and school level data is not provided. The test is not designed for that level of student reporting. Therefore, the NAEP serves its intended use by providing some national comparative data but provides very limited use to the state and local districts on improving teaching and learning.
--

NAEP is a good assessment but again, the reporting to very limited.

I think it is a good way to see the big picture. Wyoming is consistently in the top 10 nationally. This should be shared with everyone.

I wish our WDE and legislators would celebrate the good scores.

I like the NAEP assessment for the information. I don't like the congratulations letter we get every year when we are notified that we get to participate in the assessment, every district in the state does. The proctor oversees and a bit much because we may not follow the rules of the assessment, but I do understand the need to exact

results. The process is relatively easy and the students work through it with ease. Perhaps too much ease as I don't think they participate in this assessment with the same vigor as the PAWS, less expectation.
Strongly dislike it.
I'm not familiar with how it is currently being used.
NAEP is the only way we can measure our state against other states. However there is no way to USE it at the district or school level to improve performance.
N/A
R. J. Kost has been on the test question group nationally and he struggled with the philosophy and creators of the test. There is more theory and college philosophy than practical and application directed assessment. It basically addresses one style of student.
I believe it provides a common understanding of student performance across this nation based on a one set of assessment standards. Makes perfect sense to me.
NAEP is just another measure. The media likes it for press.
N/A
We don't get district or building or individual student results. I don't find the results very useful.
One more assessment that isn't completely aligned to local curriculum.
NAEP works well at the State level. I would like to see the data become meaningful at the school level. To accomplish this, each school should be given access to their results, and furthermore, it would be beneficial to use this data to compare ourselves to other schools in Wyoming and other states.
Give local schools the results.
NAEP results are published at the state level only, and not useable at a district level.
It's used to compare nationally. It would be more useful if we could get student level data.
N/A
I think this test is handled well in the state and follows suggested use.
The real value of NAEP is that it provides a means for Wyoming schools to compare ourselves regionally and nationally.
It's difficult to have an opinion about this test when we don't see it or get results - but Wyoming does perform well on this test.
The NAEP test is an excellent overall measure for student performance, but no higher than the 8 th grade. Since NAEP data is not available at the district, school, or individual levels, the only comparison to be made is with summary scores from other states.
It does what it is meant to do; show national ranking.
N/A
NAEP gives us a comparison to other states which can be helpful. However it is not totally aligned to what we want our kids to know and be able to do. In addition, results are only available at the state level which means the results are not easily used to make educational decisions.
We don't get specific student data or school data.
NAEP has the unfortunate problem associated with a federalized testing mandate. The results don't mean anything to students or schools, so motivation is a serious question. The timing (in midyear) is also problematic because students have not finished an academic year. This is another federal mandate that serves no real purpose for Wyoming educators or students.
At this point we see no benefit from yet another required test.

Question 19	Feel free to provide additional comments related to PAWS and its administration and scoring.
Comment Section	
N/A	
Test takes too long, results are not timely, Over the past five years - the test has gotten worse - not better which makes districts and state look foolish.	
N/A	

N/A
Sorry for the lack of information. This is my first year at central office.
N/A
N/A
Information sent to districts needs to be sent in a timely manner and be correct. The WDE should work with districts in providing accurate information, which is then reported to the state. Training given to districts must be consistent and useful and the assessment shouldn't be "changed" every year. If we are to continue assessing Writing it should be scored on its own. We shouldn't be using a language arts score where reading consists of 60% of the score and writing 40%. Each area should stand on its own. In the beginning students were allowed to take the assessment twice...if we do nothing else we need to be able to offer this option again. Sometimes the day we picked to do the assessment isn't the best day for the student (life brings its own surprises at times). We have the opportunity to have the best educational system in the United States and we have the opportunity to build an assessment system which reflects what is actually occurring in our schools. With the current assessment system I do not believe we have actually captured the true learning occurring. It's time to make it right. Thank you for the opportunity to have input on this issue!
N/A
N/A
N/A
Given our experience with PAWS, one could conclude that we have been an experimental development site for a test company who is continually remodeling the test for their own purposes. We need a test designed for Wyoming educators and children that will provide data to help with improving educational opportunities for students.
I want to express our sincere appreciate for the Management Committee's time and efforts to collect input from local districts. This survey was well prepared and asked the right questions. We would also thank the Select Committee on School Finance for their recent inquiries into assessments and accountability systems used in Wyoming. It is encouraging to see our elected legislators being persistent in getting information and answers on issues so critical to Wyoming's youth.
N/A
N/A
We need a summative assessment in Wyoming. I think PAWS has lost its edge. There is no support for it even though I think it is a pretty valid test with good alignment to our curriculum standards. Testing takes too long, the results are not returned in a timely or useful manner.
N/A
N/A
N/A
The district has put a lot of time into designing common and strategic approaches to addressing student achievement. PAWS is part of that approach. Its poor administration and lateness in scoring have been detrimental to executing our strategic plan. We have expended a lot of effort to align student learning with PAWS measurement skills. Changing the vendor and adding features to a test doesn't require a change in measurement skills. The structure we have aligned to can live in a different "body" if the targets don't change. It is difficult to improve if the targets keep changing.
N/A
PAWS may have had its faults at times, but it was a good start and focus for what we were trying to accomplish – determining proficiency to meet the requirements of NCLB. Now, the focus is in a different direction and it is our opinion that we need to change to meet that focus which is a more reasonable direction than NCLB previously had. We welcome the challenge and accountability if we can use it correctly.
I believe I have committed all of thoughts in previous questions
Any assessment system we have will have problems. I do feel we can do better.
N/A

<p>Fix it don't throw it out. It would take two years to design a new assessment. Help districts, schools, teachers, students, and parents make use of the information provided by the PAWS results.</p>	
<p>N/A</p>	
<p>N/A</p>	
<p>Where are the Wyoming parents and community members who ask for this assessment? Are we participating to be compliant with NCLB or are we trying to serve Wyoming youth, or are we trying to do both. The current PAWS is trying to serve too many masters and failing to serve any of them well. Perhaps the minimum assessment to meet the mandates, then we use NWEA MAP data for state and local comparisons and incorporate ACT explore, ACT, and NAEP data to provide complete criteria for evaluation and monitoring purposes. Increased time on task and decrease interruption by assessment.</p>	
<p>The decision to move back to a paper and pencil administration of PAWS for 2011 is regressive. It does not solve the problem of on-line testing and will increase administration, and scoring time. At a time when the RFP for the national assessment requires it to be substantially delivered on-line with immediate results, the move to a paper pencil administration is counter to the direction of assessment nationally and internationally.</p>	
<p>N/A</p>	
<p>N/A</p>	
<p>N/A</p>	
<p>N/A</p>	
<p>PAWS is a detriment to Wyoming students</p>	
<p>Whatever decisions are made regarding the state assessment, the Traffic Signal report is one of the most valuable pieces of data to inform teachers about student proficiency and improve instructional delivery. It would be wonderful if that tool would be available with whatever assessment is given. Allow 10th grade students to continue to bank their scores. This is a very strong motivational tool for students to take learning seriously. By the time those students get into the 11th grade, the ACT keeps the majority of them in line as well. Don't eliminate the state assessment. In the past 10 years, education has made huge leaps in student achievement with teachers and administrators examining aggregated and disaggregated student data. If a different assessment instrument is used, please make certain that it has a measure for individual student Lexile (reading) scores and Quantile (math) scores. DO NOT make the assessment entirely pencil paper. Students will have difficulty with this, but the "old school" teachers might not since "it's always been done that way." Thank you for the opportunity to express our opinions in this important survey.</p>	
<p>N/A</p>	
<p>N/A</p>	
<p>N/A</p>	
<p>I don't understand! If we questioned the validity of the test because of the crazy problems with testing last year; and a student was still able to test proficient or advanced in spite of that, shouldn't they get credit for it anyway? It just seems that WDE cares more about numbers (AYP) than kids. And it is certainly clear that the federal government would throw the student under the bus just to prove point. Whatever point that may be. What happened to integrity? It seems that the WDE chose to work with a company that could not fulfill its promises; and then decided that this method is also the way the WDE would deal with kids and school districts! You have made state testing a joke. Its paper and pencil. No, it's computer. Wait, no it paper and pencil again. These changes do more to mess up the validity of a standardized test than anything else we do. Remember that the term is standardized testing. It is not standardized if we keep changing the way it is administered.</p>	
<p>We hope that the results of the survey are broadly shared with the public and the legislature. All too often, Wyoming educators have not been given voice with regard to assessments and accountability systems that are designed without meaningful consultation. In point of fact, the same kind of survey (albeit shorter) should be collected from Wyoming teachers, parents, and other stakeholders.</p>	
<p>We need to limit required testing to 1 test and then allow teachers the time designated by schools to teach. At this point, we are teaching testing and not the material that is essential for students to become effective members of society.</p>	

Extra	<p>Does your district participate in the "traffic signal report" made available by WDE, which includes instructional suggestions? If you answer yes or no, please give us your opinion of</p>
--------------	---

N/A
N/A
N/A
We have had several individuals participate in the "traffic signal report" work done by Pearson. These have not been terribly productive for our district (although it's better than nothing), because there is lack of clarity with regard to the yellow performance band. Our faculty are unclear whether the yellow represents the high end of basic, the low end of proficient, or both. Indeed, there is significant legislative criticism about the lack of use of PAWS data, but there is precious little staff development on the meaning and utility of these results.

Extra Question 2	Does your district use PAWS scores as criteria for "body of evidence"? If you answer yes or no, please give us your opinion of the manner by which WDE reviews and comments on a district's " body of evidence"?
-------------------------	--

Comment Section	
------------------------	--

N/A
Please refer to our responses to questions 7 and 9.
N/A
No, peer review has validity issues.
N/A
N/A
No. I think the Body of Evidence program is seriously broke. The evaluation of the district's Body of Evidence is subjective. Three evaluation groups coming up with three different answers makes it apparent to me the process is to subjective.
N/A
N/A
We do when a student is appealing the decision of proficiency for math, language arts, or science. The review of the "Body of Evidence" plan should be a review by the state and not by "Peers". There are too many obstacles to overcome with the bias of each person on the reviews. The state should review and give it the green light or inform the district that they need to revise.
N/A
N/A
N/A

Currently we do use PAWS scores for BOE only as an alternative piece of evidence to demonstrate proficiency if students have not demonstrated proficiency on common district assessments in prior years. I'm not convinced that we need another system for graduation. I would recommend eliminating it. The system WDE uses is not a bad system, but the scoring guide is poorly written and is interpreted differently by each set of peer reviewers.

N/A

Yes, we use PAWS scores as criteria for the body of evidence. But most of our students have already earned proficient status in reading, writing, and math through our local assessments, embedded in Indicator Courses. The WDE peer review system for evaluating the "body of evidence" is an earnest undertaking by department staff. They are genuinely interested in trying to qualify, rather than disqualify, BOE plans. They are caught in a Catch-22 where they don't want to change the rubric, because that would be unfair to some districts, but they recognize that the rubric is fallible in some significant ways. In this case, the WDE is doing the best it can, but it is a severely flawed system. This legislative mandate should be stricken, especially when the legislature wants to add additional accountability measures that are more meaningful.

Extra Question 3	Do you believe the writing portion of PAWS is an objective way to measure a student's proficiency in writing? If you answer yes or no, please give us your suggestions on ways to improve this component of the assessment.
-------------------------	---

Comment Section

N/A

N/A

N/A

N/A

N/A

No, as there seems to be discrepancies in the scoring of student work. At times, the prompts are confusing and unclear. Also, consideration should be given to removing writing from AYP determinations as writing is not a requirement.
--

N/A



APPENDIX B

Selected Statutes

Title 21: Education

Chapter 2: The Administration of the State System of Education at the State Level

Article 3: State Board of Education

21-2-304: Duties of the state board of education.

(a) The state board of education shall:

(i) Establish policies for public education in this state consistent with the Wyoming Constitution and statutes and may promulgate rules necessary or desirable for the proper and effective implementation of this title and its responsibilities under this title. Nothing in this section shall give the state board rulemaking authority in any area specifically delegated to the state superintendent;

(ii) Through the evaluation and accreditation of school districts, implement and enforce the uniform standards for educational programs prescribed under W.S. 21-9-101 and 21-9-102 in the public schools of this state, including any educational institution receiving any state funds except for the University of Wyoming and Wyoming community colleges. The board shall ensure that educational programs offered by public schools in accordance with these standards provide students an opportunity to acquire sufficient knowledge and skills, at a minimum, to enter the University of Wyoming and Wyoming community colleges, to prepare students for the job market or postsecondary vocational and technical training and to achieve the general purposes of education that equips students for their role as a citizen and participant in the political system and to have the opportunity to compete both intellectually and economically in society;

(iii) By rule and regulation and in consultation and coordination with local school districts, prescribe uniform student content and performance standards for the common core of knowledge and the common core of skills specified under W.S. 21-9-101(b), and promulgate uniform standards for programs addressing the special needs of student populations specified under W.S. 21-9-101(c) that ensure these student populations are provided the opportunity to learn the common core knowledge and skills as prescribed by the uniform student content and performance standards pursuant to this paragraph. Student content and performance standards prescribed under this paragraph shall include standards for graduation from any high school within any school district of this state and shall describe required performance levels in order to achieve proficiency of the common core of knowledge and common core of skills prescribed under W.S. 21-9-101(b). The ability to prescribe content and performance standards shall not be construed to give the state board of education the authority to prescribe textbooks or curriculum which the state board is hereby forbidden to do. Graduation

standards imposed under this paragraph shall require the successful completion of the following components, as evidenced by passing grades or by the successful performance on competency-based equivalency examinations:

- (A) Four (4) school years of English;
 - (B) Three (3) school years of mathematics;
 - (C) Three (3) school years of science; and
 - (D) Three (3) school years of social studies, including history, American government and economic systems and institutions, provided business instructors may instruct classes on economic systems and institutions.
- (iv) Establish, in consultation with local school districts, requirements for students to earn a high school diploma as measured by each district's body of evidence assessment system prescribed by rule and regulation of the state board and required under W.S. 21-3-110(a)(xxiv). A high school diploma shall provide for one (1) of the following endorsements which shall be stated on the transcript of each student:
- (A) Advanced endorsement which requires a student to demonstrate advanced performance in a majority of the areas of the common core of knowledge and skills specified under W.S. 21-9-101(b) and proficient performance in the remaining areas of the specified common core of knowledge and skills, as defined by the uniform student content and performance standards promulgated by the state board pursuant to paragraph (a)(iii) of this section;
 - (B) Comprehensive endorsement which requires a student to demonstrate proficient performance in all areas of the common core of knowledge and skills specified under W.S. 21-9-101(b) as defined by the uniform student content and performance standards promulgated by the state board pursuant to paragraph (a)(iii) of this section;
 - (C) General endorsement which requires a student to demonstrate proficient performance in a majority of the areas of the common core of knowledge and skills specified under W.S. 21-9-101(b) as defined by the uniform student content and performance standards promulgated by the state board pursuant to paragraph (a)(iii) of this section.
- (v) Through the state superintendent and in consultation and coordination with local school districts, implement a statewide assessment system comprised of a coherent system of measures that when combined, provide a reliable and valid measure of individual student achievement for each public school and school district within the state, and the performance of the state as a whole. Improvement of teaching and learning in schools and fostering school program improvement shall be the primary

purposes of statewide assessment of student performance in Wyoming. The statewide assessment system shall:

(A) Measure individual student performance and progress in a manner substantially aligned with the uniform educational program and student content and performance standards imposed by law and by board rule and regulation;

(B) Be administered at appropriate levels at specified grades and at appropriate intervals aligned to the standards, specifically assessing student performance in reading, writing and mathematics at grades four (4), eight (8) and eleven (11), and effective school year 2005-2006, and each school year thereafter, assessing student performance in reading, writing and mathematics at grades three (3) through eight (8) and at grade eleven (11). In addition and commencing school year 2007-2008 and each school year thereafter, the statewide assessment system shall assess student performance in science not less than once within each grade band for grades three (3) through five (5), grades six (6) through eight (8) and grades ten (10) through twelve (12). The structure and design of the assessment system shall allow for the comprehensive measurement of student performance through assessments that are administered each school year simultaneously on a statewide basis and through assessments administered periodically over the course of the school year which are designed to provide a more comprehensive and in-depth measurement of subject areas aligned to the state content and performance standards. The assessment system may also measure the other common core of knowledge and skills established under W.S. 21-9-101(b) which can be quantified;

(C) In addition to subparagraph (a)(v)(B) of this section, measure student performance in Wyoming on a comparative basis with student performance nationally;

(D) Measure year-to-year changes in student performance and progress in the subjects specified under subparagraph (a)(v)(B) of this section and compare and evaluate student achievement during the process of student advancement through grade levels. The assessment system shall ensure the integrity of student performance measurements used at each grade level to enable valid year-to-year comparisons;

(E) Include multiple measures and item types including grade appropriate multiple choice and open-ended testing such as constructed-response, extended-response and performance-based tasks, to ensure alignment to the statewide student content and performance standards;

(F) Provide a fair and unbiased assessment of student performance without regard to race, ethnicity, limited English proficiency and socioeconomic status;

(G) Provide appropriate accommodations or alternative assessments to enable the assessment of students with disabilities as specified under W.S. 21-9-101(c)(i) and students with limited English proficiency;

(H) Provide a measure of accountability to enhance teaching and learning in Wyoming and in combination with other measures and information, assist school districts in determining individual student progress. The assessment results shall be reported to students, parents, schools, school districts and the public in an accurate, complete and timely manner and shall be used in conjunction with a school district's annual assessment to design educational strategies for improvement and enhancement of student performance. This design for improvement shall be part of each district's school improvement plan. In consultation and coordination with school districts, the board shall review and evaluate the assessment system regularly and based upon uniform statewide reports from each district, annually report to the legislature on student performance at specified grade levels and on school improvement plans.

(vi) Effective school year 2005-2006 and each school year thereafter, through the state superintendent and in consultation and coordination with local school districts, by rule and regulation establish a statewide accountability system providing annual accountability determinations for all schools and school districts imposing a range of educational consequences resulting from accountability determinations whereby:

(A) The continuous improvement of student achievement at all schools and appropriate educational interventions fostering continuous improvement serve as the basis for statewide accountability system design;

(B) Annual accountability determinations within the system are made for each school based upon adequate yearly progress measures defined by the federal No Child Left Behind Act of 2001, as may be subsequently amended, and the school's progress in improving student achievement as measured by adequate yearly progress data and by data from the district's body of evidence assessment system required under W.S. 21-3-110(a)(xxiv) and from other related sources which improve the reliability of accountability determinations as prescribed by rule and regulation of the board;

(C) To the extent possible, appropriate consequences resulting from accountability determinations are made subject to the discretion of school districts. The system shall establish a range of consequences which increase in the degree of intensity over time, with significant interventions imposed only upon repeated failure to meet school improvement and performance criteria over a consecutive period of time;

(D) Teacher and administrator quality and student remediation are the focus of consequences imposed upon schools failing to meet school improvement and performance criteria and target levels;

(E) A range of rewards is provided to schools meeting school improvement and performance criteria at levels set by the state board.

(b) In addition to subsection (a) of this section and any other duties assigned to it by law, the state board shall:

-
- (i) Repealed By Laws 1997 Special Session, ch. 3, § 302; 1994, ch. 17, § 2.
 - (ii) Enforce the uniform state educational program standards imposed by W.S. 21-9-101 and 21-9-102 and the uniform student content and performance standards established by rules and regulations adopted under subsection (a) of this section by taking appropriate administrative action with the state superintendent, including but not limited to the changing of accreditation status;
 - (iii) Repealed by Laws 1993, ch. 217, § 3.
 - (iv) Repealed by Laws 1987, ch. 190, §§ 2, 5.
 - (v) Initiate or facilitate discussions regarding the needs of and the means for improving education;
 - (vi) Repealed by Laws 1987, ch. 190, §§ 2, 5.
 - (vii) Repealed by Laws 1994, ch. 17, § 2.
 - (viii) Approve or disapprove alternative scheduling for school districts requesting to operate for fewer than one hundred seventy-five (175) days in school year, but no schedule shall be approved which reduces the pupil-teacher contact time defined by the state board;
 - (ix) Repealed by Laws 1994, ch. 17, § 2.
 - (x) Repealed by Laws 2006, Chapter 34, § 2.
 - (xi) Repealed by Laws 1994, ch. 17, § 2.
 - (xii) Repealed by Laws 1994, ch. 17, § 2.
 - (xiii) Repealed by Laws 1994, ch. 17, § 2.
 - (xiv) Establish improvement goals for public schools for assessment of student progress based upon the national assessment of educational progress testing program and the statewide assessment system established under paragraph (a)(v) of this section;
 - (xv) Promulgate rules and regulations for the development, assessment and approval of school district teacher performance evaluation systems. Rules and regulations adopted under this paragraph shall allow each district flexibility in developing an evaluation system which meets the individual needs of the district;

(xvi) Through the state superintendent, implement, administer and supervise education programs and services for adult visually handicapped and adult hearing impaired persons within the state.

(c) The state board shall perform an ongoing review of state board duties prescribed by law and may make recommendations to the legislature on board duties. In addition and not less than once every five (5) years, the board shall evaluate and review the uniformity and quality of the educational program standards imposed under W.S. 21-9-101 and 21-9-102 and the student content and performance standards promulgated under paragraph (a)(iii) of this section, and shall report findings and recommendations to the joint education interim committee of the legislature on or before December 1 of the year in which the review and evaluation was undertaken. The joint education interim committee shall report its recommendations, based upon findings and recommendations of the state board, to the legislature during the immediately following legislative session.

(d) Repealed by Laws 1994, ch. 17, § 2.

(e) In addition to subsections (a) and (b) of this section, the state board shall establish statewide goals for Wyoming public education.

Title 21: Education

Chapter 9: Courses of Study, Textbooks, Supplies

Article 1: Courses of Study

21-9-101. Educational programs for schools; standards; core of knowledge and skills; special needs programs; class size requirements; cocurricular activities.

(a) The board of trustees of each school district within the state shall cause the schools under its jurisdiction to provide an educational program in accordance with uniform standards defined under this section and rules and regulations promulgated by the state board of education pursuant to W.S. 21-2-304(a).

(b) Each school district within the state shall provide educational programs sufficient to meet uniform student content and performance standards at the level established by the state board of education in the following areas of knowledge and skills:

(i) Common core of knowledge:

(A) Reading/language arts;

(B) Social studies;

(C) Mathematics;

-
- (D) Science;
 - (E) Fine arts and performing arts;
 - (F) Physical education;
 - (G) Health and safety;
 - (H) Humanities;
 - (J) Career/vocational education;
 - (K) Foreign cultures and languages;
 - (M) Applied technology;
 - (N) Government and civics including state and federal constitutions pursuant to W.S. 21-9-102.
- (ii) For grades one (1) through eight (8), reading, writing and mathematics shall be emphasized under the common core of knowledge specified under paragraph (b)(i) of this section;
- (iii) Common core of skills:
- (A) Problem solving;
 - (B) Interpersonal communications;
 - (C) Keyboarding and computer applications;
 - (D) Critical thinking;
 - (E) Creativity;
 - (F) Life skills, including personal financial management skills.
- (c) In addition to subsection (b) of this section, each school district within this state shall provide programs designed for the special needs of those student populations specified within this subsection. Programs under this subsection shall be provided and shall identify special student populations in accordance with rules and regulations of the state board of education. The state board shall monitor the proportion of students in each special needs category, compared to available regional averages. Special needs student populations include:

(i) Children with disabilities evaluated in accordance with rules and regulations of the state board as having intellectual disability, hearing impairments including deafness, speech or language impairments, visual impairments including blindness, serious emotional disturbance, orthopedic impairments, autism, traumatic brain injury, other health impairments, specific learning disabilities, deafness and blindness or other multiple disabilities, and who, because of the impairments, need special education and related services; and

(ii) Gifted and talented students identified by professionals and other qualified individuals as having outstanding abilities, who are capable of high performance and whose abilities, talents and potential require qualitatively differentiated educational programs and services beyond those normally provided by the regular school program in order to realize their contribution to self and society.

(d) In addition to subsections (b) and (c) of this section, each school district within this state shall endeavor to maintain when practicable, in kindergarten through grade three (3) within the district, an average class size of no more than twenty (20) students per teacher, excluding children with disabilities who spend more than fifty percent (50%) of their time outside of regular classroom instruction.

(e) Nothing in this section shall be construed to prohibit school districts from establishing cocurricular activity programs which have as their purpose to provide educational experiences not otherwise provided by the local district. The legislature specifically encourages school districts to establish programs of this type.

(f) It is the intent of the legislature that the funding mechanism established by law for schools encourage school districts to achieve the goal of reduced class sizes.

(g) Not later than the 2002-2003 school year, all school districts shall provide instruction in foreign languages to students in kindergarten through grade 2 in accordance with standards promulgated by the state board.

Title 21: Education

Chapter 3: School Districts in General

Article 1: In General

21-3-110. Duties of boards of trustees.

(a) The board of trustees in each school district shall:

(xxiv) Establish a student assessment system to measure student performance relative to the uniform student content and performance standards in all content areas for which the state board has promulgated standards pursuant to W.S. 21-2-304(a)(iii). To the extent required by the state board under W.S. 21-2-304(a)(v) and (vi), the district assessment system shall be integrated with the

statewide assessment system and the statewide accountability system. A component of the district assessment system required by this paragraph shall include a body of evidence assessment system designed and used to determine the various levels of student performance as described in the uniform student content and performance standards relative to the common core of knowledge and skills prescribed under W.S. 21-9-101(b).

No Child Left Behind Act (2001)

Title I: Improving the Academic Achievement of the Disadvantaged

Part A — Improving Basic Programs Operated by Local Educational Agencies

Subpart 1 — Basic Program Requirements

SEC. 1111. STATE PLANS.

b) ACADEMIC STANDARDS, ACADEMIC ASSESSMENTS, AND ACCOUNTABILITY-

(3) ACADEMIC ASSESSMENTS-

(A) IN GENERAL- Each State plan shall demonstrate that the State educational agency, in consultation with local educational agencies, has implemented a set of high-quality, yearly student academic assessments that include, at a minimum, academic assessments in mathematics, reading or language arts, and science that will be used as the primary means of determining the yearly performance of the State and of each local educational agency and school in the State in enabling all children to meet the State's challenging student academic achievement standards, except that no State shall be required to meet the requirements of this part relating to science assessments until the beginning of the 2007-2008 school year.

(B) USE OF ASSESSMENTS- Each State educational agency may incorporate the data from the assessments under this paragraph into a State-developed longitudinal data system that links student test scores, length of enrollment, and graduation records over time.

(C) REQUIREMENTS- Such assessments shall--

(i) be the same academic assessments used to measure the achievement of all children;

(ii) be aligned with the State's challenging academic content and student academic achievement standards, and provide coherent information about student attainment of such standards;

(iii) be used for purposes for which such assessments are valid and reliable, and be consistent with relevant, nationally recognized professional and technical standards;

(iv) be used only if the State educational agency provides to the Secretary evidence from the test publisher or other relevant sources that the assessments used are of adequate technical quality for each purpose required under this Act and are consistent with the requirements of this section, and such evidence is made public by the Secretary upon request;

(v)(I) except as otherwise provided for grades 3 through 8 under clause vii, measure the proficiency of students in, at a minimum, mathematics and reading or language arts, and be administered not less than once during--

(aa) grades 3 through 5;

(bb) grades 6 through 9; and

- (cc) grades 10 through 12;
 - (II) beginning not later than school year 2007-2008, measure the proficiency of all students in science and be administered not less than one time during--
 - (aa) grades 3 through 5;
 - (bb) grades 6 through 9; and
 - (cc) grades 10 through 12;
- (vi) involve multiple up-to-date measures of student academic achievement, including measures that assess higher-order thinking skills and understanding;
- (vii) beginning not later than school year 2005-2006, measure the achievement of students against the challenging State academic content and student academic achievement standards in each of grades 3 through 8 in, at a minimum, mathematics, and reading or language arts, except that the Secretary may provide the State 1 additional year if the State demonstrates that exceptional or uncontrollable circumstances, such as a natural disaster or a precipitous and unforeseen decline in the financial resources of the State, prevented full implementation of the academic assessments by that deadline and that the State will complete implementation within the additional 1-year period;
- (viii) at the discretion of the State, measure the proficiency of students in academic subjects not described in clauses (v), (vi), (vii) in which the State has adopted challenging academic content and academic achievement standards;
- (ix) provide for—
 - (I) the participation in such assessments of all students;
 - (II) the reasonable adaptations and accommodations for students with disabilities (as defined under section 602(3) of the Individuals with Disabilities Education Act) necessary to measure the academic achievement of such students relative to State academic content and State student academic achievement standards; and
 - (III) the inclusion of limited English proficient students, who shall be assessed in a valid and reliable manner and provided reasonable accommodations on assessments administered to such students under this paragraph, including, to the extent practicable, assessments in the language and form most likely to yield accurate data on what such students know and can do in academic content areas, until such students have achieved English language proficiency as determined under paragraph (7);
- (x) notwithstanding subclause (III), the academic assessment (using tests written in English) of reading or language arts of any student who has attended school in the United States (not including Puerto Rico) for three or more consecutive school years, except that if the local educational agency determines, on a case-by-case individual basis, that academic assessments in another language or form would likely yield more accurate and reliable information on what such student knows and can do, the local educational agency may make a determination to assess such student in the appropriate language other than English for a period that does not exceed two additional consecutive years, provided that such student has not yet reached a level of English language proficiency sufficient to yield valid and reliable information on what such student knows and can do on tests (written in English) of reading or language arts;

- (xi) include students who have attended schools in a local educational agency for a full academic year but have not attended a single school for a full academic year, except that the performance of students who have attended more than 1 school in the local educational agency in any academic year shall be used only in determining the progress of the local educational agency;
- (xii) produce individual student interpretive, descriptive, and diagnostic reports, consistent with clause (iii) that allow parents, teachers, and principals to understand and address the specific academic needs of students, and include information regarding achievement on academic assessments aligned with State academic achievement standards, and that are provided to parents, teachers, and principals, as soon as is practicably possible after the assessment is given, in an understandable and uniform format, and to the extent practicable, in a language that parents can understand;
- (xiii) enable results to be disaggregated within each State, local educational agency, and school by gender, by each major racial and ethnic group, by English proficiency status, by migrant status, by students with disabilities as compared to nondisabled students, and by economically disadvantaged students as compared to students who are not economically disadvantaged, except that, in the case of a local educational agency or a school, such disaggregation shall not be required in a case in which the number of students in a category is insufficient to yield statistically reliable information or the results would reveal personally identifiable information about an individual student;
- (xiv) be consistent with widely accepted professional testing standards, objectively measure academic achievement, knowledge, and skills, and be tests that do not evaluate or assess personal or family beliefs and attitudes, or publicly disclose personally identifiable information; and
- (xv) enable itemized score analyses to be produced and reported, consistent with clause (iii), to local educational agencies and schools, so that parents, teachers, principals, and administrators can interpret and address the specific academic needs of students as indicated by the students' achievement on assessment items.

(e) PEER REVIEW AND SECRETARIAL APPROVAL-

(1) SECRETARIAL DUTIES- The Secretary shall—

- (A) establish a peer-review process to assist in the review of State plans;
- (B) appoint individuals to the peer-review process who are representative of parents, teachers, State educational agencies, and local educational agencies, and who are familiar with educational standards, assessments, accountability, the needs of low-performing schools, and other educational needs of students;
- (C) approve a State plan within 120 days of its submission unless the Secretary determines that the plan does not meet the requirements of this section;
- (D) if the Secretary determines that the State plan does not meet the requirements of subsection (a), (b), or (c), immediately notify the State of such determination and the reasons for such determination;
- (E) not decline to approve a State's plan before—

- (i) offering the State an opportunity to revise its plan;
- (ii) providing technical assistance in order to assist the State to meet the requirements of subsections (a), (b), and (c); and
- (iii) providing a hearing; and

(F) have the authority to disapprove a State plan for not meeting the requirements of this part, but shall not have the authority to require a State, as a condition of approval of the State plan, to include in, or delete from, such plan one or more specific elements of the State's academic content standards or to use specific academic assessment instruments or items.

(2) STATE REVISIONS- A State plan shall be revised by the State educational agency if it is necessary to satisfy the requirements of this section.

(f) DURATION OF THE PLAN-

(1) IN GENERAL- Each State plan shall—

(A) remain in effect for the duration of the State's participation under this part; and

(B) be periodically reviewed and revised as necessary by the State educational agency to reflect changes in the State's strategies and programs under this part.

(2) ADDITIONAL INFORMATION- If significant changes are made to a State's plan, such as the adoption of new State academic content standards and State student achievement standards, new academic assessments, or a new definition of adequate yearly progress, such information shall be submitted to the Secretary.

(g) PENALTIES-

(1) FAILURE TO MEET DEADLINES ENACTED IN 1994-

(A) IN GENERAL- If a State fails to meet the deadlines established by the Improving America's Schools Act of 1994 (or under any waiver granted by the Secretary or under any compliance agreement with the Secretary) for demonstrating that the State has in place challenging academic content standards and student achievement standards, and a system for measuring and monitoring adequate yearly progress, the Secretary shall withhold 25 percent of the funds that would otherwise be available to the State for State administration and activities under this part in each year until the Secretary determines that the State meets those requirements.

(B) NO EXTENSION- Notwithstanding any other provision of law, 90 days after the date of enactment of the No Child Left Behind Act of 2001 the Secretary shall not grant any additional waivers of, or enter into any additional compliance agreements to extend, the deadlines described in subparagraph (A) for any State.

(2) FAILURE TO MEET REQUIREMENTS ENACTED IN 2001- If a State fails to meet any of the requirements of this section, other than the requirements described in paragraph (1), then the Secretary may withhold funds for State administration under this part until the Secretary determines that the State has fulfilled those requirements.

No Child Left Behind Act (2001)**Title IX-General Provisions****Part C — Coordination of Programs; Consolidated State and Local Plans and Applications****SEC. 9302. Optional Consolidated State Plans or Applications.****(a) GENERAL AUTHORITY-**

(1) **SIMPLIFICATION-** In order to simplify application requirements and reduce the burden for State educational agencies under this Act, the Secretary, in accordance with subsection (b), shall establish procedures and criteria under which, after consultation with the Governor, a State educational agency may submit a consolidated State plan or a consolidated State application meeting the requirements of this section for —

(A) each of the covered programs in which the State participates; and

(B) such other programs as the Secretary may designate.

(2) **CONSOLIDATED APPLICATIONS AND PLANS-** After consultation with the Governor, a State educational agency that submits a consolidated State plan or a consolidated State application under this section shall not be required to submit separate State plans or applications under any of the programs to which the consolidated State plan or consolidated State application under this section applies.

(b) COLLABORATION-

(1) **IN GENERAL-** In establishing criteria and procedures under this section, the Secretary shall collaborate with State educational agencies and, as appropriate, with other State agencies, local educational agencies, public and private nonprofit agencies, organizations, and institutions, private schools, and representatives of parents, students, and teachers.

(2) **CONTENTS-** Through the collaborative process described in paragraph (1), the Secretary shall establish, for each program under this Act to which this section applies, the descriptions, information, assurances, and other material required to be included in a consolidated State plan or consolidated State application.

(3) **NECESSARY MATERIALS-** The Secretary shall require only descriptions, information, assurances (including assurances of compliance with applicable provisions regarding participation by private school children and teachers), and other materials that are absolutely necessary for the consideration of the consolidated State plan or consolidated State application.



APPENDIX C

Alleged Liquidated and Alleged Default Damages – Professional Services Contract between the State of Wyoming, Department of Education and NCS Pearson, Inc., executed January 12, 2009 (Amendment No. 1, executed October 14, 2009).

Alleged Liquidated Damages

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Liquid. Damage
1) Test Materials	Production of Test Administrator Training Materials (p. 4 ¹)	Test Administrator Training Materials are delivered on time to training sites without errors.	The <i>PAWS 2010 Directions for Administration</i> manual had to be split into two volumes, Paper/Pencil and Online due to the incompleteness of the online platform functionality. The <i>PAWS 2010 Online Directions for Administration</i> manual was posted as a PDF on Feb. 23, 2010. Printed copies were not delivered to schools until February 26, 2010. In addition, the Header Sheets, which provide pre-printed school information and used when shipping test materials back to Pearson, had to be re-printed due to school code errors. These too were shipped out the week of Feb. 23, 2010.	January 11, 2010 (See 2009-2010 Prg. Plan – p. 15).	Paper/ Pencil: Feb. 9, 2010; On-line: Feb 26, 2010	\$ 70,000 (35 days)

¹ Unless otherwise denoted, page numbers referenced refer to Attachment A (PAWS Action Plan) to the Original Professional Services Contract between the State of Wyoming, Department of Education and NCS Pearson, Inc., executed January 12, 2009.

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Liquid. Damage
2) Test Administration	Shipping (p. 5)	Ancillary Materials delivered on time to schools.	See explanation for Liquidated Damages #1) above.	On or before Feb. 20, 2010	February 26, 2010	\$ 10,000 (5 days)
3) Technology	TestNav Student Tutorial and ePat (p. 8)	Student test tutorial accessible to schools prior to test administration for full demonstration of online assessment tools; and online platform is accessible throughout the test administration window.	<p>Online platform failed to perform throughout the window when the tutorial was available; technical issues remained during the duration; problems were never completely rectified.</p> <p>For purposes of calculating liquidated damages the Department utilized the tutorial and testing windows; arguably liquidated damages could continue to be imposed to date.</p>	On or before Feb. 15, 2010 (See 2009-2010 Prg. Plan – p. 10).	To date satisfactory product has not been delivered; contract amendment required as a result of failure.	\$ 86,000 (43 days: Feb 15 – April 14, 2010)
4) Technology	Online administration platform (p. 8)	Online platform is accessible throughout the test administration window.	The Pearson TestNav 7.0 online assessment platform continued to malfunction throughout the test administration window. This is supported by Pearson's PAWS Communication #37 (April 5, 2010), "Please continue to report any technical challenges you may experience to Pearson's CSC, even though you have resolved the situation. Pearson is aware that some schools are still challenged by random	March 8, 2010 (See 2009-2010 Prg. Plan – p. 12).	To date satisfactory product has not been delivered; contract amendment required as a result of failure.	\$ 56,000 (28 days: March 8 – April 14, 2010)

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Liquid. Damage
			<p>incidents with the online testing platform. Pearson will notify WDE to introduce new software coding in the event we identify any specific root cause for these issues.”</p> <p>Testing window had to be extended an additional three (3) days due to the technical difficulties of the on-line platform.</p> <p>For purposes of calculating liquidated damages the Department utilized the testing window; arguably liquidated damages could continue to be imposed to date.</p>			
5) Reporting Services	Real-Time, online multiple-choice individual student results (p. 12)	PAWS results and reports delivered on-time, accurate and without error	On April 9, 2010, the WDE declined to accept the Real-Time (On-Demand), online multiple-choice individual student results due to multiple significant errors in appearance and accuracy. Satisfactory reports were delivered on April 23, 2010.	March 31, 2010 (Note: 2009-2010 Pgm. Plan p. 21 indicates reports due March/April TBD, 2010 – but because a date was never identified the date in the Action Plan was utilized).	April 23, 2010	\$ 36,000 (18 days: March 31 - April 23, 2010).

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Liquid. Damage
6) Reporting Services	Final On-line Report (p. 12)	PAWS results and reports delivered on-time, accurate and without errors.	Due to the deficiencies of the TestNav 7.0 Platform, analysis by an independent 3 rd party was necessary prior to processing of data and information and compilation of any reports; all reporting was delayed significantly as a result of the failure of the TestNav 7.0.	May 31, 2010 (See 2009-2010 Prg. Plan - p. 21)	Sept. 27, 2010.	\$ 172,000 (86 days: May 31, 2010 – Sept. 27 2010).
7) Reporting Services	Student, School, District, Demographic, and State PDF score reports (p. 12)	PAWS results and reports delivered on-time, accurate and without errors.	See explanation for Liquidated Damages #6) above.	May 31, 2010 (See 2009-2010 Prg. Plan - p. 21 -22)	To-date satisfactory reports have not been delivered; anticipated delivery Oct. 8, 2010.	\$ 190,000 (95 days: May 31, 2010 – October 8, 2010).
8) Reporting Services	Student Paper Reports (p. 12)	PAWS results and reports delivered on-time, accurate and without errors.	See explanation for Liquidated Damages #6) above.	May 31, 2010 (See 2009-2010 Prg. Plan - p. 21-22)	To-date satisfactory reports have not been delivered; anticipated delivery Oct. 15, 2010.	\$200,000 (100 days: May 31, 2010 – October 15, 2010).
9) Reporting Services	Interpretive Guide (p. 12)	PAWS results and reports delivered on-time, accurate and without errors.	See explanation for Liquidated Damages #6) above.	August 5, 2010 (See 2009-2010 Prg. Plan - p. 22)	To-date satisfactory reports have not been delivered; anticipated delivery Oct. 15, 2010.	\$ 104,000 (52 days: August 5, 2010 – Oct. 15, 2010).
10) Program Management	Complete and	Program is fully staffed to	Contractor failed to staff call center during the tutorial and test	Dec . 2, 2009 – Dec.	During the test administration	\$ 78,000 (39 days:

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Liquid. Damage
	consistent staffing of all program components (p. 14)	meet Agency's needs and expectations.	administration window with staff with sufficient technical expertise to assist districts and the Agency is resolving technical issues.	1 2010	and tutorial the call center remained understaffed and without sufficient expertise	Feb 19 – April 14, 2010).
11) Program Management	Quarterly Progress Reports (p. 15)	Reports shall include: information on the progress and status of the program; and, issues encountered and solutions for issues.	The Agency was not accurately informed of status of the program and the issues with TestNav which affected many of the deliverables. Quarterly reports lacked detail and did not include call center information. Agency repeatedly asked for call center reports and the Contractor failed to provide it in a timely manner.	Dec. 1, 2009 Mar. 1, 2010 June 1, 2010 Sept. 1, 2010	Dec. 1, 2009 Report was delivered on time but failed to identify the TestNav Platform issue and technical difficulties regarding implementation of the Spring 2010 PAWS Administration; March and June Reports delivered June 9, 2010; September 1, 2010 Report: to-date not received.	\$ 128,000 (64 days: Dec. 1 – Feb. 28). \$ 132,000 (66 days: Mar. 1 – June 1, 2010) \$ 14,000 (7 days: June 1 – June 9, 2010) \$ 50,000 (25 days: Sept. 1 – October 5, 2010)
12) Program Management	Quarterly Invoice (p. 15 -16)	Quarterly invoices shall be delivered in a timely manner in	The Agency was not provided with the quarterly invoice in a timely manner; when quarterly reports were submitted, lacked critical information.	Sept. 1, 2010	To date (Sept. 9, 2010) Invoice due Sept. 1, 2010 has not been	\$ 16,000 (8 days: Sept. 1 – Sept. 9, 2010)

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Liquid. Damage
		accordance with Section Four of the Professional Services Contract Between the State of Wyoming, Department of Education and NCS Pearson, Inc.			delivered to Agency	
TOTAL ALLEGED LIQUIDATED DAMAGES						\$ 1,342,000

Alleged Default Damages

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Number of Days Late
1) Test Development	Test Form Construction (p. 2)	All test forms construction tasks completed on time; and all test forms meet expectations and requirements of the Agency.	Acceptable and accurate test forms for the 2010 TestNav Platform were not provided by the Contractor to the Agency until February 26, 2009.	Dec. 2008 - Nov. 2009	Feb. 26, 2010	64 days: Dec. 1 2009 - Feb. 26, 2010
2) Test Materials	Production of Ancillary Materials (p. 4 - 5)	Ancillary Materials are developed and produced on time and without errors.	See explanation for Liquidated Damages #1) above.	Sept. 2009- Nov. 2009	Feb. 26, 2009	64 days: Dec. 1 2009 - Feb. 26, 2010
3) Events and Committees	English Language Arts (ELA) Standard Setting (p. 7)	Materials for all events and committees developed and delivered on time for successful completion of tasks; and, agenda, sign-in sheets and applicable results sent to Agency following event	Pearson's Proposal included a detailed description of the creation of a Language Arts score for Adequate Yearly Progress (AYP) which in turn promised to stabilize the PAWS writing results by relying on a combination of links to multiple-choice Language Arts items, changing the writing rubrics, changing the number of score points possible and placing the PAWS reading and writing scores on a single scale. To assist with the creation of the vertical scale from grades 3 - 8, students responded to an "off-grade" writing prompts at grades 3 - 8. The contract called for the creation	Mar. - Nov. 2010	To date has not been delivered and delivery of an acceptable product is not anticipated.	N/A: acceptable standard setting for ELA cannot be completed by Contractor.

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Number of Days Late
		or committee completion.	of approximately 100 multiple-choice Language Arts items to be reviewed, edited, field tested, scored and taken through a data review process in the 2010 assessment cycle. These tasks were completed. However, upon the failure of the online TestNav platform to function, several steps were taken to report PAWS writing scale scores as the entire test had changed. When Pearson's lead psychometrician attempted to implement the writing equating model as proposed, it was discovered to be flawed. As a result, all of the multiple-choice language arts items have been removed from any further use on a future PAWS assessment and the time and resources spent on their development now represent a loss to the Department. Finally, all student responses and subsequent hand scores of the "off-grade" writing prompts used at grades 3-8 were unusable due to the failure of the equating model.			
4) Data Management Service	State Student Data File (p. 10)	Student data file delivered accurately and on time.	See explanation for Liquidated Damages #6) above.	May 31, 2010 (See 2009-2010 Prg. Plan - p. 22).	To date has not been delivered; anticipated delivery Oct. 15, 2010.	92 days: May 31, 2010 – October 5, 2010
5) Data Management	District Student Data File (p.	District data file delivered	See explanation for Liquidated Damages #6) above.	August 5, 2010 (See	To date has not been delivered;	49 days: August 5,

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Number of Days Late
Service	10)	accurately and on time.		2009-2010 Prg. Plan - p. 23).	anticipated delivery Oct. 15, 2010.	2010 – October 5, 2010
6) Psychometric Services	Validity studies and preparation of technical manuals (p. 11)	Technical Manual shall meet Agency's expectations and delivered in time to meet Federal Accountability Guidelines and Federal Peer Review of State Assessment Systems Guidelines.	Due to the deficiencies of the TestNav Platform, Technical Manuals could not be prepared.	On or before Sept. 15, 2010	To date has not been delivered; delivery is not anticipated	20 days: Sept. 15, 2010 – October 5, 2010
7) Program Management	Complete and consistent staffing of all program components (p. 14)	Program is fully staffed to meet Agency's needs and expectations.	On-line portion of Spring 2010 PAWS Administration was not adequately staffed; indicated 6 full-time staff employed by Pearson for <i>all</i> on-line assessments nationwide contracted by Pearson; insufficient staffing as evidenced by TestNav 7.0 failure. Customer Service Center was inadequately staffed as well; evidenced by experience of State and District Staff.	Dec. 2, 2009 - Dec. 1, 2010	N/A	N/A
8) Program Management	Meetings and Communication (p. 17)	Minutes are provided to the Agency upon completion of planning, meetings,	Communication from Contractor to the State and for communication to the District routinely has numerous spelling, grammatical and typographical errors; basic proof-reading is not taking place prior to publication of communication.	Dec. 2, 2009 – Dec. 1, 2010	N/A	N/A

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Number of Days Late
		<p>management meetings, and weekly content and management calls;</p> <p>Attendance and support when requested at presentations involving the Wyoming State Assessment System;</p> <p>Availability for communication including, correspondence by email and phone as necessary to discuss progress and next steps to complete various projects and conference calls when necessary to discuss issues that arise; and,</p>				

Requirement Section	Deliverable	Measurable Indicator	Description of Events	Date due	Date Delivered	Number of Days Late
		Respond to questions and issue that arise in a timely manner.				



APPENDIX D

Committee on Health, Education, Labor, and Pensions





WYOMING LEGISLATIVE SERVICE OFFICE
DAN J. PAULI, Director

December 23, 2010

David Miller, Chairman
John Hines, Vice Chairman
Management Audit Committee
213 State Capitol
Cheyenne, WY 82002

Tom Harkin, Chairman
Mike Enzi, Ranking Member
Committee on Health, Education, Labor and Pensions
428 Senate Dirksen Office Building
Washington, DC 20510

RE: Request for Government Accountability Office (GAO) review of the United States Department of Education (USDOE) Peer Review process re: No Child Left Behind (NCLB) administered by the Office of Elementary and Secondary Education (OESE)

Messrs. Harkin and Enzi,

At its 12/15/10 meeting, the Management Audit Committee voted to release an audit of the Proficiency Assessments for Wyoming Students (PAWS). The Management Audit Committee conducts audits pursuant to W.S. 28-8-107 (b) (i).

During the PAWS audit, staff discovered what appears to be significant problems and inconsistencies associated with the peer review process used by USDOE related to the No Child Left Behind Act (NCLB). The peer review process, pursuant to NCLB (Title I, Part A, Section 1111, (e)), is used to determine the validity and reliability of statewide assessments to measure Adequate Yearly Progress (AYP).

According to our audit findings, the Wyoming Department of Education (WDE) has expended a tremendous amount of resources to meet the requirements of an enigmatic, laborious, and inconsistently applied peer review process. Education experts in the field concur, and one expert fully blames the bureaucracy within the USDOE for the inability of PAWS to be used to its fullest extent as an instructionally supportive assessment. From what we understand, only 31 states have received full peer review approval from USDOE.

On behalf of the Management Audit Committee, we would like to formally request that the Committee on Health, Education, Labor, and Pensions request that the Government Accountability Office (GAO) review the peer review process pending the reauthorization of the Elementary and Secondary Education Act (ESEA).

We have attached a copy of Chapter 5 of the PAWS audit, which provides staff's finding and recommendation. Should you require additional information, please contact Gerald W. Hoppmann, Program Evaluation Manager, at 307-777-7881.

Sincerely,



Gerald W. Hoppmann
Program Evaluation Manager

Pp: David Miller, Chair

Pp: John Hines, Vice Chair

Cc: Dan Pauli, Director, Legislative Service Office

Enclosure: 1

Recent Program Evaluations

Wyoming State Archives	May 2000
Turnover and Retention in Four Occupations	May 2000
Placement of Deferred Compensation	October 2000
Employees' Group Health Insurance	December 2000
State Park Fees	May 2001
Childcare Licensing	July 2001
Wyoming Public Television	January 2002
Wyoming Aeronautics Commission	May 2002
Attorney General's Office: Assignment of Attorneys and and Contracting for Legal Representation	November 2002
Game & Fish Department: Private Lands Public Wildlife Access Program	December 2002
Workers' Compensation Claims Processing	June 2003
Developmental Disabilities Division Adult Waiver Program	January 2004
Court-Ordered Placements at Residential Treatment Centers	November 2004
Wyoming Business Council	June 2005
Foster Care	September 2005
State-Level Education Governance	December 2005
HB 59: Substance Abuse Planning and Accountability	January 2006
Market Pay for State Employees	July 2006
Wyoming Drug Courts	July 2006
A&I HRD Role in State Hiring	December 2006
Kid Care CHIP: Wyoming's State Children's Health Insurance Program	June 2007
Wyoming Retirement System: Public Employee Plan	August 2007
WYDOT and General Fund Appropriations for Highways	May 2008
Wyoming Child Protective Services	September 2008
Department of Fire Prevention and Electrical Safety	December 2008

Office of Health Care Licensing and Surveys	July 2009
Victim Services Division: Phase I	August 2009
Victim Services Division: Phase II	February 2010
Reading Assessment and Intervention Program	February 2010
Office of State Lands & Investments: Management of State Trust Lands	June 2010

Evaluation reports can be obtained from:

Wyoming Legislative Service Office
 213 State Capitol Building Cheyenne, Wyoming 82002
 Telephone: 307-777-7881 Fax: 307-777-5466
 Website: <http://legisweb.state.wy.us>

