

**A Proposal to Administer
An Academic Achievement Growth Model
For Accountability Reporting for
North Dakota Schools and Districts**

**Submitted by the
North Dakota Department of Public Instruction**

October 15, 2008

**North Dakota Department of Public Instruction
Dr. Wayne G. Sanstead, State Superintendent
600 East Boulevard Avenue, Dept. 201
Bismarck, ND 58505-0440
www.dpi.state.nd.us**

Table of Contents

Abstract	3
I. Analyzing the State’s Capacity to Support a Growth Model	5
1.1 Improving the Current North Dakota Accountability Model	5
1.2 Building Supportable Value in Growth	7
1.3 Instructional Supports for Growth Measures	8
1.4 Capacity to Incorporate Growth for Accountability	9
1.4.1 State Data Infrastructure	9
1.4.2 Match Rates and Analysis of Possible Bias	11
1.4.3 Current Accountability System Rules	12
1.4.4 Student Participation and Secondary Indicator Rates	14
1.4.5 Implications of Growth-Based Accountability	14
1.4.6 Assessment Quality, Data Systems and Growth	15
1.4.7 Alternate Assessments	18
II. The Growth Model Proposal	18
2.1 Retaining the Current Accountability System’s Essential Design	18
2.2 Amending Growth into the Current Accountability System	19
2.3 Assigning Student Achievement with a Growth-to-Standard Model	20
2.4 Considerations for the State’s Alternate Assessments	24
2.5 The Effect of the Growth-to-Standard Model on Adequate Yearly Progress	25
III. Adherence to the Seven Core Principles	25
3.1 Universal Proficiency by 2014 and School Accountability	25
3.2 Establishing Student-Appropriate Growth Targets	26
3.3 Separate Accountability for Reading and Mathematics	27
3.4 Inclusion of All Students	27
3.5 Compliant State Assessment Systems	27
3.6 Tracking Individual Student Progress	28
3.7 Participation Rates and Additional Academic Indicators	28
Conclusion	29
Appendix A: Confidence Intervals	30
Appendix B: Tables and Figures	36

Abstract

The North Dakota Department of Public Instruction (NDDPI) seeks authorization to incorporate a growth-to-standard model of student achievement within its current state accountability system. This growth-to-standard achievement model will provide an additional means of measuring student growth in relation to the state's challenging achievement standards and reporting overall student proficiency and growth-to-proficiency for every public school, school district, and the state.

The NDDPI submits this growth-to-standard model proposal pursuant to current peer review guidance provided by the U.S. Department of Education (ED). This proposal is consistent with the principles and practices of the State's currently approved federal accountability workbook rules. Additionally, this proposal employs student achievement and demographic data derived from the North Dakota State Assessment System, which received a *Fully Approved* status from the ED in June 2007. This proposal has been adopted based on its simple and straightforward design, its transparent and verifiable metrics, and its ready accessibility to educators, parents, and the public.

Assurance of Consistent Measures

The NDDPI proposes to administer a growth-to-standard model based on the achievement results generated from the North Dakota State Assessment (NDSA) and the North Dakota Alternate Assessment based on both alternate and modified achievement standards (NDAA) in reading and mathematics in grades 3-8 and 11. Scores generated in grades 3 and 11 will constitute baseline reference data but will not be used alone to calculate projected scores. The NDSA evidences a comprehensive vertically equated scale. The NDAA reports non-vertical scale achievement level performance. Federal peer review studies have reported that the NDSA and NDAA based on alternate achievement standards are valid and reliable measures and meet recognized quality standards. The state's NDAA based on modified achievement standards is currently undergoing federal peer review but has been previously approved by the ED for inclusion in determining annual adequate yearly progress.

Within the NDSA, the state will interpolate grade 9 and 10 reference scale scores, respectively, based on the NDSA's vertically equated scale for the purposes of optimizing growth projections. Within the NDAA, the state will credit as attained sufficient growth any student who has moved from the *novice* achievement level (the lowest below-proficient achievement level) to the *partially-proficient* achievement level (the highest below-proficient achievement level). The proposed growth-to-standard model will calculate growth projections for all students and report these student-specific projections to schools and districts. The state will apply the projections of all below-proficient students, as determined by the NDSA and NDAA, for the purposes of calculating adequate yearly progress (AYP) for every public school, public school district, and the state.

Assurance of Data Integrity

The state provides a longitudinal data system that can reliably support the administration of the proposed growth-to-standard model. The state's longitudinal data system provides student-specific identification numbering that ensures the reliable matching of student records. This identification system has been functional since 2003 and has been operational within the current assessment system's structure since the 2004-05 school year. The state evidences consistent,

reliable statewide tracking of students among various school plants and school districts for the purposes of properly employing its full academic year accountability rules. The state data system reliably manages various student achievement and demographic information, including ethnicity, disability, free/reduced lunch participation, language proficiency, among other categories. The state's methods of identifying, numbering, matching, and tracking account for an approximate 94% inter-year tracking rate, statewide, and respectable tracking rates among multiple years, grades, subjects, and demographic subgroups.

Assurance for Integrated Accountability Rules

The proposed growth-to-standard model will be integrated into the state's current accountability rules as a final phase in determining AYP for every public school, school district, and the state. The state would first conduct a status analysis, followed by a safe harbor analysis, and followed by the Title I privilege analysis, all conducted according to currently approved rules. Then the state would employ, as a final phase of the AYP processing, this growth-to-standard model. The state intends to apply a 99% confidence interval in its growth-to-standard model, in accordance with its current accountability rules, and presents the merits of this position within this proposal. The state will conduct a growth-to-standard analysis for all students, regardless of their achievement level, and all schools, regardless of their AYP status, for the purposes of student services support, school improvement activities, and future accountability reference.

The state's proposed growth-to-standard model uses actual student achievement data from the NDSA and NDAA to establish and track student performance such that every student evidences sufficient yearly gains to reach or surpass proficiency within three years. Grades 3 and 11 constitute baseline reference years for recording initial or final student achievement; however, no growth projections are calculated on grades 3 or 11 data alone. The proposed growth-to-standard model presents up to three years of historical data for demonstration purposes. The model references the previous year's scale score for each student and subtracts that scale score from the current year's scale score. Within the NDSA, if the difference between the previous and current years' scores is equal to or greater than one-third of the difference of the *proficient* cut score value in three years, then the student is recorded as *proficient* for accountability purposes. If the difference is less than one-third the difference of the *proficient* cut score value in three years, then the student's respective achievement level remains unaltered. For each subsequent annual iteration, the original three-year goal remains as the target; no annual resetting of the target occurs. Within the NDAA, the state will credit as sufficient growth any student who has moved from the *novice* achievement level (the lowest below-proficient achievement level) to the *partially-proficient* achievement level (the highest below-proficient achievement level).

The NDDPI places this proposal before peer review with full confidence that the state's fully approved assessment system, longitudinal data system, reliable accountability rules, and commitment to the incorporation of growth analysis into instructional design, provide systemic integrity.

I. Analyzing the State's Capacity to Support a Growth Model

The North Dakota Department of Public Instruction (NDDPI) seeks authorization to incorporate a growth-to-standard model of student achievement within its current state accountability system. This growth-to-standard achievement model will provide an additional means of measuring student growth in relation to the state's challenging achievement standards and reporting overall student proficiency and growth-to-proficiency for every public school, school district, and the state.

The NDDPI submits this growth-to-standard model proposal pursuant to the *Peer Review Guidance for the NCLB Growth Model Pilot Applications*, approved by the U.S. Department of Education (ED), dated January 25, 2006, and the unofficial guidance, *Growth Model Proposal Peer Recommendations for the NCLB Growth Model Pilot Applications*, released by the ED, dated September 23, 2008. Additionally, this growth-to-standard model proposal is predicated on the currently approved *North Dakota State Consolidated Application Accountability Workbook* and the North Dakota State Assessment System's *Fully Approved* status, awarded by the ED in June 2007.

The valid and reliable measurement and reporting of student academic growth has been a long-standing challenge within education. The adoption of a status achievement model marks a first step in clearly delineating the public policy goal of defining and ensuring the attainment of literacy (i.e., proficiency and above) for all students. The incorporation of a growth principle provides an additional, recognized means of tracking and crediting students, schools, and districts along an appropriately challenging path to attaining recognized proficiency for every student. When students are measured for growth against valid and reliable assessments and schools and districts are credited for supportable gains, professional acceptance of accountability measures improves, educational institutions adopt meaningful incentives to build further value into student growth, public understanding of educational goals increases, and the public's educational mission becomes more transparent and matures.

1.1 Improving the Current North Dakota Accountability Model

The adoption of any growth model within a state accountability system should minimally address three concerns: (1) the reliable measurement and tracking of student growth against a common measure; (2) the crediting of supportable growth for a school's or district's adequate yearly progress (AYP) reporting, and (3) the use of student-specific growth projections to guide future instruction toward meeting proficiency goals. Growth models are designed to measure, recognize, and foster further academic gains.

A status accountability model establishes clearly defined targets for student achievement. A growth model accountability provision provides a defensible path to meet an achievement target, for students who have not already done so, while providing credit for sufficient growth. A growth model also provides a conceptual aid to educators to attend to the unique growth patterns of all students so that each student can optimize their educational experience.

North Dakota's current accountability model complies with the current status accountability model prescribed within the *Elementary and Secondary Education Act of 1965*, as reauthorized within the *No Child Left Behind Act of 2001*. The state has invested substantial human and financial resources into developing a valid and reliable assessment system, a stable longitudinal data system, and a defensible and compliant accountability system. The North Dakota State

Assessment (NDSA) and the North Dakota Alternate Assessment (NDAA) based on alternate achievement standards have been awarded a *Fully Approved* status from the ED in June 2007 for fully meeting all requirements of the federally prescribed peer review process. The state's NDAA based on modified achievement standards is currently undergoing federal peer review but has been previously approved by the ED for inclusion in determining annual adequate yearly progress. Henceforth, for the purposes of brevity, the NDAA will refer to both forms of the alternate assessment, i.e., based on alternate achievement standards and modified achievement standards. The NDDPI has developed a dependable longitudinal data system with integrated file sharing and extensive SAS-based analysis tools to perform the various procedures required of a statewide accountability system and reflective of recognized quality assurance measures. The NDDPI has developed and administers an accountability system whose AYP rules are fully compliant with federal law and regulations. The state's federal accountability system has demonstrated itself to be stable and has undergone compliance monitoring by the ED for validation.

With the status provisions of the state's accountability system in place and performing in a proper manner, the NDDPI seeks to access the additional educational value of a growth model provision. The NDDPI has studied growth model options since the approval of growth model pilot projects in 2005 by the ED. The NDDPI became eligible for a growth model submission in June 2007 with the *Fully Approved* status of the state assessment system. The NDDPI has consulted with its assessment and accountability technical advisory committee for assistance in analyzing the benefits of various growth model options within the currently defined regulatory constraints. The NDDPI has engaged the resources of its state assessment and accountability state-level advisory committee, the *Standards, Assessment, Learning, and Teaching (SALT)* committee, a 12 member committee comprised of local school administrators and teachers, to review the merits and foundational designs of an appropriate statewide model.

Within the past several years, since the introduction of a growth model option from the ED, various state-level committees have expressed support for the inclusion of a growth model provision within the state's federal accountability plan. The state's association of school administrators has endorsed the adoption of a growth model. A working group of the state's sixteen largest school districts, the Study Council, has requested the adoption of a meaningful growth model option. The Governor's Commission on Education Improvement, Assessment and Accountability Committee, has explored the inclusion of an appropriate growth model within the state's adequacy study. The North Dakota Legislative Council's interim Education Committee has discussed the proper placement of a growth model within the state's overall accountability provisions. Individual school districts and citizens have encouraged the development of a viable growth model provision.

Underscoring this collective interest is a desire among schools, parents, and students to understand and monitor each student's unique growth toward proficiency and ultimately to their highest achievable level. Policymakers and educators are interested in providing a common means of defining what appropriate growth might be and supporting schools in tracking each student's academic growth toward the desired achievement target. The adoption of a meaningful growth model unleashes the resources of the state to support this goal and to add recognized value to its state accountability system.

1.2 Building Supportable Value in Growth

A variety of growth model concepts have been analyzed within recent years, including value-added indexing, value chart indexing, growth-to-standard, and percentile gain, among other models. Each model legitimately addresses a unique measure of growth that might be appropriate within a specific educational culture and with specified constraints to its application. Discerning which growth model option might be most appropriate and viable within the unique setting of a given state requires attention to foundational policy goals and educational culture.

Throughout the discussions that have surrounded the study and possible adoption of growth models within North Dakota, certain policy goals and cultural sensibilities have repeatedly emerged. Participants in these discussions have reiterated the need to adopt a growth model that is simple and straightforward in design, applies transparent metrics and lends itself to easy verification, and is readily accessible to educators, parent, and the public. Users of growth information should be able to easily understand the nature of any projections, the reasons that support how any determination is made, and can explain the method and its meaning to another individual with confidence and without conceptual error. Such a policy goal constitutes a daunting challenge.

In deliberating the merits of various models, the growth-to-standard model emerged as an appropriately designed model that sufficiently met the policy goals stated above. It provides a conceptual framework that meets the desired aims of both accountability and instructional improvement. It is visual or graphic in nature; it illustrates the student's current achievement level to the desired achievement target, as represented by a line that intersects with proficiency. Its metrics employ simple mathematics that can be easily explained to a layperson and that can be independently verified for accuracy. Its simple design allows any teacher to understand and explain its method and purpose to a parent verbally or with simple graphic aids. It allows students to understand where they currently stand and where they need to be to meet grade-level expectations now and in the future. This model advances instruction by design by demystifying the assessment and reporting process.

The growth-to-standard model, or any other legitimate model, provides a conceptual and graphic means to demonstrate that *all* students are capable of evidencing improvement, and significant improvement at that. Its graphic qualities allow for educators, parents, and students to visualize better the learning process, what will be required for a non-proficient student to narrow any achievement gap, and what the trending shows for an already proficient student. The growth-to-standard model offers a means to measure the relative value of education provided or the relative effectiveness of any school or school district based on the incidence of growth within its institution. Schools that have a population of disadvantaged students who have not yet reached proficiency but who have evidenced measurable gains will be credited for the educational value they offer. Such a model provides a safeguard against making false identifications of AYP deficiencies, when, in fact, growth patterns demonstrate that true academic gains are occurring.

A growth-to-standard model offers a supportable means of reporting true value, both status and growth, within every public school, school district, and the state.

This true measure of academic achievement integrates seamlessly with the state's current accountability system. The state's system applies to all public schools and school districts and assesses all students in grades 3-8 and 11 in reading and mathematics for the purposes of AYP determinations and overall reporting. Achievement results are generated for composite and

subgroup reporting. Student achievement reports are made available to parents in paper and electronically for every school, school district, and the state. State reports include achievement level reporting occurs by subgroup, subject, standard, and benchmark. The state issues reports on the incidence of highly qualified teachers, including by subject, course, demographic subgroup, school size, and experience. The state issues additional reports regarding graduation, retention, dropout, and attendance.

The state provides school choice and supplemental services options for students in program improvement schools. Schools on program improvement develop improvement plans that detail the initiatives designed to improve overall school operations and student academic gains. The NDDPI with the advisory assistance of its Title I Committee of Practitioners supervises the development and implementation of all program improvement activities. Schools within extended program improvement must undergo a selective series of initiatives, including administrative restructuring, curricular restructuring, or other prescribed activities, to build internal capacity for student academic gains.

Under North Dakota state law, any teacher who teaches in a core subject area must meet fully the provisions of a highly qualified teacher as a component of his/her teaching license. The state supports this zero-tolerance policy by removing the approval and accreditation designation of any school that does not meet the full highly-qualified-teacher requirement. Such a loss of an approval or accreditation designation may result in the loss of certain state foundation aid financial payments. A teacher must be highly qualified to legally teach in North Dakota.

The inclusion of a growth-to-standard model complements the many current policy initiatives that ensure that every student meets grade-level expectations. These initiatives exist to build the capacity of every school and school district to provide the quality of instruction and student supports that optimize the students' opportunities to meet the state's challenging standards.

1.3 Instructional Supports for Growth Measures

The growth-to-standard model provides, in addition to a means for appropriately determining adequate yearly progress for schools and districts, a meaningful tool to incorporate a uniform measure of student achievement and growth into school curricular and instructional improvements.

Since the growth-to-standard model applies to all students with evidence of past performance on the NDSA and NDAA, it effectively becomes a vehicle for tracking the achievement pattern of students over time within every school, district, subject, subgroup, and the state. Growth for every student is projected based on past years' performance. Growth is projected up to a three-year target goal, dependent on the grade level of the student, and subsequent growth is measured against that target goal and the calculated projection.

For a student who might be below-proficient, schools can use projections to determine an appropriate instructional plan to foster the student's achievement of the ultimate target goal. For a student who might be proficient already, schools can use projections to determine if the student is on target to remain proficient and to reach advanced proficiency, or if the student is evidencing a deterioration of the projection and thus at risk for falling to a less than proficient level in the future. The state's growth-to-standard model will produce a summary growth report for every student at the school level that will itemize the student's historical scale scores, the student's projected target goal and timeline, the calculated trajectory for remaining on course

with the target, and the intermediate scores that must be met to evidence sufficient growth each year. This information will aid school personnel in independently verifying the determination and in establishing a meaningful instructional plan.

The growth-to-standard model is designed to report progress on the NDSA and NDAA, which are aligned to the state's challenging standards. The growth-to-standard model reconnects teachers to the state's challenging achievement descriptors and content standards, which specify the content-level of the assessments. This reconnection to the achievement descriptors and standards reiterates the instructional foundations of the assessments and their role in ensuring that every student has sufficient access to and an opportunity to reach proficiency on the state's standards. This reconnection to standards underscores the foundational principles of the state's accountability system.

The state supervises a long-standing school accreditation procedure that requires all approved schools to conduct ongoing school improvement efforts based on supportable student achievement data. Schools build school profiles based on available instructional impact data to apply best-practices instructional initiatives that are designed to optimize student performance. The growth-to-standard model offers a uniform growth metric based on the state's challenging standards that will fill a void in current data profile efforts. Growth trending supplies insight into student performance that might not otherwise be observed on status measurements alone.

The growth-to-standard model evidences the potential to further disaggregate student achievement data into more clearly discernable patterns, which offer further potential for more insightful analysis and improvements.

1.4 Capacity to Incorporate Growth for Accountability

The NDDPI stipulates that the state has the infrastructure and technical capacity to implement and administer a growth-to-standard model that will meet all recognized quality assurance measures. The NDDPI is prepared to demonstrate this capacity for certain critical program components, including the following:

- State Data Infrastructure
- Match Rates and Analysis of Possible Bias
- Current Accountability System Rules
- Student Participation and Secondary Indicator Rates
- Implications of Growth-Based Accountability
- Assessment Quality, Data Systems, and Growth
- Alternate Assessments

The following narrative addresses each of these critical components.

1.4.1 State Data Infrastructure

The NDDPI has developed a statewide student, school, and district web-based data management system, titled the State Automated Reporting System (STARS). The NDDPI supervises the management of data submission from all schools and school districts within the state according to the established data system protocols. The STARS system collects, compiles, analyzes, and reports data gathered on a variety of student-, school-, and district-specific

information. The following are included among the student and programmatic support data files managed by the STARS system: student enrollment, including student demographic and achievement information; homeless; immigrant; limited English proficient; refugee; Section 504; migrant; special education membership; summer school; federal consolidated grant applications; federal consolidated program budgeting; Title I targeted and schoolwide participation; school district financial reporting; special education unit reporting; vocational education center reporting; school district fall membership and structural organization reporting; school directory listings; school plant membership and structural organization reporting; school plant directory listing; school licensed personnel reporting; school course identification and registration; school non-licensed personnel reporting; state assessment demographic and program reporting; school district professional development reporting; student suspension and expulsion reporting; school fall calendar reporting; school spring calendar reporting; and school district transportation reporting.

STARS is fully capable of managing the records of the state's approximately 190 school districts, 470 schools, and 95,000 students, and the compilation of summative data across multiple years.

The NDDPI supervises the management of extensive student-specific information gathered at the school and school district level on STARS. Since 2004 STARS, and its precursor data management system titled the Online Reporting System (ORS), have built student reporting around a student specific identification numbering system. Each student at the time of enrollment within any North Dakota school receives a unique, dedicated student identification number. The supervision of the issuance of all student identification numbers resides within the NDDPI, as a quality assurance measure to ensure the dedicated data management of each student's academic records. The STARS system automatically reviews each student entry to ensure that no student identification has been replicated, altered, or omitted. NDDPI's Management Information System Unit manually verifies the integrity of the STARS student identification system as another means of validating integral data. The STARS student identification system performs in a dependable and reliable manner.

The NDDPI conducts ongoing analyses with data collected by STARS. The NDDPI uses a variety of software, dependent on the designated task, to perform its many functions. Among the software used for data analysis are Excel, Access, SAS, and SQL Server. The NDDPI also contracts with outside data analyst specialists to conduct various data analysis tasks. Among the tasks that the NDDPI contracts to conduct is the annual AYP determination process. NDDPI's Management Information System Unit, Standards and Achievement Unit, and Data Driven Enterprises, the NDDPI external contractor, collaborate to compile, analyze, verify, and report all AYP determinations, school and district profile reports, and other related reports. The NDDPI and Data Driven Enterprises use SAS as the selective software to manage and analyze all AYP determinations and student performance reports.

As a manner of conducting annual adequate yearly progress, the NDDPI references up to three years of inclusive student demographic and achievement data to produce valid and reliable reports. The NDDPI compiles information from various independent data files to perform an AYP analysis, including student demographic files, student attendance files, student achievement files, student graduation files, school plant organization files, school plant program files, school district organization files, and school district program files, among others to conduct a thorough and accurate analysis of AYP at the school, district, and state levels. The NDDPI assembles and integrates these respective files across three years into its SAS code to conduct the AYP analysis. The state's unique student identification, school identification, and district identification coding ensure that all reports are appropriately specific and allow for year-to-year

stability. The NDDPI routinely tracks students, through its full academic year rules and data management safeguards, as they move from school to school and from district to district. The NDDPI is likewise able to monitor the exiting of students from the system to ensure the proper transfer of records among schools and out-of-state.

STARS has proven itself to be a dependable longitudinal data system and has demonstrated through the annual administration of AYP determinations, based on multi-year records, its capacity to produce valid and reliable reports. STARS and the state's SAS-based coding is fully capable of managing the detailed analysis required of a growth-to-standard model.

1.4.2 Match Rates and Analysis of Possible Bias

The state provides a longitudinal data system that can reliably support the administration of the proposed growth-to-standard model. The state's longitudinal data system provides student-specific identification numbering that ensures the reliable matching of student records. This identification system has been functional since 2003 and has been operational within the current assessment system's structure since the 2004-05 school year. The state evidences consistent, reliable statewide tracking of students among various school plants and school districts for the purposes of properly employing its full academic year accountability rules. The state data system reliably manages various student achievement and demographic information, including ethnicity, disability, free/reduced lunch participation, and language proficiency, among other categories. The state's methods of identifying, numbering, matching, and tracking account for consistently high inter-year tracking rates, statewide, and comparable tracking rates among multiple years, grades, subjects, and demographic subgroups.

Tables 1.4.2(a-g), presented in the Appendix, provide various data matching rates based on multiple-year, grade level, achievement level, and subgroup classifications. Table 1.4.2(a) presents the number and percent of students whose fall 2007 and fall 2006 test records match, by grade level. Table 1.4.2(b) presents the number and percent of students whose fall 2007 and fall 2006 test records match by subgroup. Table 1.4.2(c) presents the number and percent of students whose fall 2007 and fall 2006 test records match by reading achievement level. Table 1.4.2(d) presents the number and percent of students whose fall 2007 and fall 2006 test records match by mathematics achievement level. Table 1.4.2(e) presents the number and percent of students whose fall 2007 and fall 2004 test records match by grade level, representing a three-year match. Table 1.4.2(f) presents the number and percent of students whose fall 2007 and fall 2004 test records match by reading achievement level, representing a three-year match. Table 1.4.2(g) presents the number and percent of students whose fall 2007 and fall 2004 test records matched by mathematics achievement level, representing a three-year match.

These comparative match rates indicate that the state's longitudinal data system demonstrates a high match rate between contiguous years and a respectable match rate among three-year comparisons. The student identification system performs appropriately to account for the high-to-respectable match rates. Non-matching rates indicate students who were not in the state's assessment system in one or the other test administrations. The state's longitudinal data system demonstrates a fundamental capacity to administer a growth-to-standard model.

1.4.3 Current Accountability System Rules

The North Dakota accountability system conducts annual AYP determinations based on clearly defined protocols for every public school, school district, and the state. AYP determinations are based on student academic achievement rates and participation rates that result from valid and reliable state assessments in reading and mathematics in grades 3-8 and 11. AYP determinations also report student graduation rates for all public high schools and student attendance rates for all public elementary and middle schools. AYP determinations report student achievement rates and participation rates for composite results and for the following subgroups: (1) ethnicity, including white, Native American, black, Asian, Hispanic, and other; (2) economically disadvantaged; (3) students with disabilities; and (4) students with limited English proficiency. AYP determinations are reported separately for each public school plant, every school district, and the state.

Table 1.4.3(a): Indicators Referenced in Determining AYP

<i>I. Primary Indicators: Student Achievement on State Assessment</i>				
<i>Population</i>	<i>Math Achievement</i>	<i>Math Student Participation (95%)</i>	<i>Reading Achievement</i>	<i>Reading Student Participation (95%)</i>
<i>Total</i>	*	*	*	*
<i>White students</i>	*	*	*	*
<i>Native American students</i>	*	*	*	*
<i>Black students</i>	*	*	*	*
<i>Asian students</i>	*	*	*	*
<i>Hispanic students</i>	*	*	*	*
<i>Other ethnicity students</i>	*	*	*	*
<i>Economic disadvantaged</i>	*	*	*	*
<i>Disabilities</i>	*	*	*	*
<i>Limited English Proficient</i>	*	*	*	*
<i>II. Secondary Indicators</i>				
<i>A. Elementary/Middle Schools: attendance rates</i>			*	
<i>B. High Schools: graduation rates</i>			*	

AYP determinations compile the number of students who have reached proficiency on the state assessments and establish respective proficiency rates for composite and subgroup reporting. These proficiency rates become the basis for calculating school plant, school district, and state AYP determinations.

AYP achievement determinations are made relative to the established annual measurable achievement targets set by state rules. A school or district must attain composite and subgroup achievement rates that meet or exceed the established achievement targets that are effective during that academic year. Below is the chart of established achievement targets established by the state for the purposes of determining AYP.

**Table 1.4.3(b): State Intermediate Achievement Targets, 2002-2014
Goals for Percentage of Students Scoring Proficient and Advanced in Reading/Math, Graduation
and Attendance Rate Targets**

Subject/ Grades	School Years*												
	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14
Reading													
4	65.1%			73.8%			82.6%			91.3%			100%
8	61.4%			71.1%			80.7%			90.4%			100%
11	NA			57.2%			71.5%			85.7%			100%
12	42.9%			NA			NA			NA			NA
Math													
4	45.7%			59.3%			72.9%			86.4%			100%
8	33.3%			50.0%			66.7%			83.3%			100%
11	NA			43.1%			62.1%			81.0%			100%
12	24.1%			NA			NA			NA			NA
Graduation	89.9%			73.09%**									
Attendance	93.0%												

• During 2001-04 the state assessment was administered at the 12th grade. Since 2004-05 the state assessment has been administered at the 11th grade, to replace the 12th grade assessment.

**In 2005-06 the graduation target was recalculated according to rules set forth within the State's Accountability Workbook. This new graduation goal will remain, unless amended by the state at a later time.

The state intermediate achievement targets identify what percentage of students in a given subject (i.e., reading and mathematics) and grade (i.e., 4, 8, and 11) must be proficient/advanced during a specified year. A school's AYP report presents actual student performance separately for reading and mathematics across all its tested grades. To arrive at an appropriate AYP determination for an entire school, the NDDPI equates or weights the effects of different grades with their different intermediate targets. If a school tests more than one grade, then the state intermediate target for each respective grade is equated, or weighted, based on the proportion of students within each respective grade. This weighting produces a single composite goal that the school must make. In effect, a school's weighted goal equates the state intermediate target across all tested grade levels for that school.

A school's weighted target, i.e. achievement target, is calculated based on the school's grade configuration (e.g., K-6, K-8, 6-9, 9-12). If a school has a grade that coincides with any grade that has an identified state intermediate achievement target, as specified for grades 4, 8, and 11 in the table above, then the school adopts that achievement target for its target. If, however, a school has more than one grade that coincides with any grades that have a designated state intermediate achievement target, then the school adopts a weighted achievement goal, based on the relative number of students in each respective grade. For instance, a K-6 school contains solely the grade 4 reading intermediate goal of 82.6%; therefore, that school adopts an achievement goal of 82.6%. However, a K-8 school includes both the grade 4 (82.6%) and the grade 8 (80.7%) intermediate goals; therefore, a weighted calculation results in a rate that lies between the two percentages. Assuming equal numbers of students in each grade, the weighted goal, or achievement goal, will equal 81.65%.

The state has established a minimum number of 10 students, to ensure proper student identification safeguards as required under the *Family Education Rights and Privacy Act*. There exist no additional minimum number rules for composite or subgroup reporting. The state applies a confidence interval, set at 99 percent, to provide for reliable determinations among settings of different sampling sizes. Safe Harbor is determined as prescribed within federal law without the application of a confidence interval.

The state's AYP rules conduct a sequence of calculations before issuing a final AYP determination. An overview of this determination process is presented below.

Table 1.4.3(c): Present State AYP Determination Process

In calculating AYP determinations, schools, districts, and the state receive the benefit of a five-tiered review, which includes the following:

1. Review of current year achievement data with a reliability test;
2. Review of two years combined achievement data with a reliability test;
3. Review of three years combined achievement data with a reliability test;
4. Determination of Safe Harbor;
5. Review of targeted Title I student achievement data with a reliability test;
 - a. Review of one year achievement data with a reliability test;
 - b. Review of two years combined achievement data with a reliability test;
 - c. Review of three years combined achievement data with a reliability test;
 - d. Determination of Safe Harbor.

If a school passes any of these five levels of review, the school is identified as "*Met adequate yearly progress*." If a school does not pass all five levels of review, as is applicable to the school, then the school is identified as "*Did not make adequate yearly progress*." This multi-step determination process ensures that any designation of "*Did not meet adequate yearly progress*" is reliable and constitutes a warranted identification.

The state's proposed growth-to-standard model would retain the core components of the state's current accountability rules for AYP determinations. The state proposes to amend the current rules to place the growth-to-standard model determinations as a new, sixth step in the process presented above for any schools that would not have met AYP at that point. The state will report, nevertheless, the growth rates for all schools, regardless of AYP determination status, through separate reports, including school and district profile reports and school-based, student-specific growth reports.

1.4.4 Student Participation and Secondary Indicator Rates

The state will continue to calculate AYP determinations for participation, graduation, and attendance according to current rule. The NDDPI issues annual AYP determinations that report the secondary indicator rates for every public school, school district, and the state. The NDDPI issues annual school and district profile reports that detail the secondary indicator rates for composite and subgroup analysis. The state will not apply any growth-to-standard model to the graduation or attendance secondary indicators.

1.4.5 Implications of Growth-Based Accountability

A growth-to-standard model provides a constructive metric to assess a school's or district's ability to add value to instruction and student learning. A growth-to-standard model provides data that reveal learning trends, both positive and negative, for every individual student in addition to composite and subgroup analysis. The ND growth-to-standard model requires any

student to evidence sufficient annual achievement growth to reach proficiency within a three-year period, without annual resets. Such evidence of growth is commendable and worthy of meritorious recognition. Such recognition for an individual student's growth is awarded with a designation of sufficient growth equivalent to proficiency for that year. The NDDPI forbids the use of growth rates of any already proficient students to be used to compensate or mitigate for other below-proficient students.

All schools, regardless of their respective status or growth determinations, are still accountable for meeting the annual measurable achievement targets established by the state. Any application of the growth-to-standard model for a school or district must still comply with meeting the state's defined target. The state makes no exemption or recognizes no distinction during the AYP determination process.

All schools, regardless of their respective status or growth determinations, are still accountable for complying with the school improvement or corrective action sanctions established by the state. Any school that does not make AYP following the application of the state's growth-to-standard model will be held to the same state school improvement or corrective action sanctions, as set forth within Section 1116 of the ESEA and the state's consolidated application accountability workbook, as approved by the ED. A school designated as not making AYP during a given year, even following the application of the growth-to-standard model, will not be sanctioned further aside from those sanctions already prescribed by law or the state's accountability workbook.

All schools, regardless of their respective status or growth determinations, will receive the benefit of a full accounting of student-specific growth rates for the purposes of reporting to students and parents and for school improvement activities.

By including a growth-to-standard model, it is the expressed intent of the state to (1) add additional relevance to the results generated by the state's assessment system, (2) add value to the AYP determination process by awarding measurable achievement gains, (3) recognize schools for accelerating the learning of below-proficient students, and (4) expand school-based improvement efforts based on growth-to-standard data.

1.4.6 Assessment Quality, Data Systems, and Growth

The state's growth-to-standard model references student achievement results generated through the NDSA and the NDAA. These two assessment instruments alone, and no other, constitute the recognized measurement of student achievement aligned to the state challenging content standards. The NDSA and the NDAA based on alternate achievement standards have been awarded a *Fully Approved* status from the ED in June 2007 for fully meeting all requirements of the federally prescribed peer review process. The state's NDAA based on modified achievement standards is currently undergoing federal peer review but has been previously approved by the ED for inclusion in determining annual adequate yearly progress. Henceforth, for the purposes of brevity, the NDAA will refer to both forms of the alternate assessment, i.e., based on alternate achievement standards and based on modified achievement standards. The NDAA is specifically addressed in section 1.4.7 below.

The federal peer review process has reviewed and approved the NDSA for adherence to the state's content standards, fidelity of the state's achievement standards, coherence within the state's overall assessment system, integrity of technical quality, fidelity to content alignment,

inclusiveness of all students, and comprehension of public reporting. The NDSA is recognized as a valid and reliable measure of student achievement against the state's standards.

The NDDPI contracts with CTB/McGraw-Hill, LLC, to develop, administer, score, analyze, and report annual statewide assessments that are fully aligned to the state's content standards and that meet all recognized technical and quality assurance standards. Achievement data are compiled and reported at the student-, classroom-, school-, district-, and state-levels. CTB/McGraw-Hill, LLC, produces annual technical reports that present comprehensive impact data regarding the performance of all tests, including the validity and use of test scores, test construction, student demographics and performance, construct-related validity and reliability, scaling and linking, test and item statistics, fairness, inter-rater reliability of constructed-response items, anchor item performance, and field test statistics.

The NDSA reports achievement scores on a vertically aligned scale. CTB/McGraw-Hill, LLC, researchers have linked the NDSA to the CTB/*TerraNova* scale, which allows the NDSA to take advantage of the vertical properties of the *TerraNova* scale. The underlying vertical scale enables comparisons to be made across grades directly on the test scale, especially to contiguous grades on the scale. The vertical scale measures a single content domain (e.g., mathematics) and allows for continuity across the grades.

North Dakota State Assessment							
Achievement Standards Cut Scores*							
May 2005							
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Reading							
Partially Proficient	579	606	619	632	644	648	679
Proficient	610	630	645	655	666	670	700
Advanced	650	670	690	696	707	714	738
Mathematics							
Partially Proficient	547	586	609	627	641	654	702
Proficient	572	610	632	653	670	686	739
Advanced	632	655	672	692	706	725	775

* Presented in scale scores on the North Dakota State Assessment

Evidence of the validity of the NDSA growth scale is provided by the increase of the scale score at selected percentiles as grade level increases. Figures 1.4.6(a) and 1.4.6(b) presented in the Appendix display the scale scores for several points on the score distributions for each grade of the reading NDSA and the mathematics NDSA, respectively. These scale scores indicate the growth, or change, in score by grade at various percentiles. Ideally, the scale score associated with each percentile will increase from grade to grade. Aside from the extreme percentiles, this is clear evidence of an upward progression of scale scores across all grades.

Reliability within testing refers to the consistency of the students' test scores on parallel forms of a test. A reliable test is one that produces scores that are expected to be relatively stable if the test is administered repeatedly to the same population of individuals under similar conditions. It is helpful to analyze reliability across time and within a specific test form.

Table 1.4.6(a) presented in the Appendix provides a comparison of the NDSA across a four-year period, specifically the mean and standard deviation of scale scores for all students. These data indicate substantial stability from form to form, which indicate that the NDSA provides sufficient reliable performance from year to year. Such a characteristic is required to provide confidence in the administration of a growth-to-standard model. Table 1.4.6(d) and Table 1.4.6(e) presented in the Appendix provide a multi-year comparison of overall student achievement on the NDSA categorized by achievement level. These data reveal respectable comparable performance from year to year and support further the evidence that the NDSA is a stable, reliable measure of student achievement across time.

Within the NDSA's measurement of reliability, the stratified alpha statistic measures the internal consistency of a test comprising multiple-choice and constructed-response items. The NDSA's average 60 item length raises the prospect of a more reliable assessment. The reliability coefficient is a ratio of the variance of true scores to observed scores, with the values ranging from 0 to 1. The closer the value of the reliability coefficient is to 1, the more consistent the scores. Generally, reliability coefficients that are equal to or greater than 0.85 are considered acceptable or defensible for tests of moderate lengths. Table 1.4.6(b) and Table 1.4.6(c) listed in the Appendix present the reliability coefficients for the NDSA reading and mathematics assessments, respectively. The statistics range from 0.85 to 0.92 and are determined to be defensible.

Within the NDSA, the reliability of reported test scores can be characterized by the standard error associated with the scores. The standard error of measurement (SEM) may be used to determine the range within which a student's true score is likely to fall. An observed score is not a student's true score; rather, it is an estimate of a student's true score. Sixty-eight percent of the time a student's observed score obtained from a single test administration would fall within one SEM of the student's true score were the test administered on other occasions. Figure 1.4.6(c) and Figure 1.4.6(d) presented in the Appendix display the conditional SEM curves for each grade and content area. SEM statistics tend to be higher at the low and high ends of the scale score range, since there are fewer observations at these levels. These figures demonstrate that the tests are designed so that measurement error is minimized in the middle of the scale range where the majority of students are located. Complete NDSA validity and reliability statistics can be accessed from the Fall 2007 NDSA Administration Technical Manual, <http://www.dpi.state.nd.us/testing/assess/07final.pdf>.

These reliability statistics indicate that the NDSA demonstrates desirable reliability. These data further indicate that the NDSA provides comparable information as students move from one grade to the next. This is a critical issue that supports the vertically-equated design of the NDSA, which further supports its use within a longitudinal analysis of growth. The NDSA has performed in a stable manner since the state moved into its current grade-level assessment regimen with the 2004-05 academic year. The NDSA provides reliable information and has done so with a sufficient number of annual administrations, such that it can legitimately be referenced for the purposes of conducting a growth-to-standard model.

1.4.7 Alternate Assessments

The NDDPI has developed and administers two independent forms of the NDAA: (1) the NDAA-1, based on alternate achievement standards, and (2) the NDAA-2, based on modified achievement standards. Federal peer review studies have determined that the NDAA-1 based on alternate achievement standards is a valid and reliable measure and meets recognized quality standards. The state's NDAA-2 based on modified achievement standards is currently undergoing federal peer review but has been previously approved by the ED for inclusion in determining annual adequate yearly progress. Henceforth, for the purposes of brevity, the NDAA will refer to both forms of the alternate assessment, i.e., the NDAA-1 based on alternate achievement standards and the NDAA-2 based on modified achievement standards.

The NDAA reports student achievement according to the state's recognized achievement levels: *novice*, *partially proficient*, *proficient*, and *advanced*. The *novice* and *partially proficient* achievement levels correspond to below-proficient. The *proficient* and *advanced* achievement levels correspond to proficient and above. Since the NDAA does not incorporate vertically aligned scale scores, the state's growth-to-standard model does not recognize scale score movement within the NDAA. Instead, within the growth-to-standard model, student growth will be defined as the reported movement from one achievement level to another, either up or down.

Within the AYP determination process, the state will only credit a school when a student records an upward movement from *novice* to *partially proficient*. No other movement up will be recognized, since such movement will amount to a proficient or above score. In the event that a student were to take the NDAA in year-one and then take the NDSA in year-two (or vice versa), a school will only be credited for growth if the student were to progress upward one achievement level from *novice* to *partially proficient*, from year-one to year-two. The NDDPI will not equate actual scale scores or performance scores between the NDSA and the NDAA.

II. The Growth Model Proposal

The state's proposed growth-to-standard model will apply to all public schools, school districts, and the state in a manner parallel to the accountability rules that are currently operational within the state. The growth-to-standard model provides an additional means of measuring student growth, both composite and by subgroup, in relation to the state's established achievement targets. The proposed model will look at the progress individual students make toward proficiency from one year to the next. Student progress is charted against each student's own projected growth-to-proficiency. Students who equal or exceed their unique growth-to-proficiency projections will be credited as proficient for the purposes of accountability reporting. Student growth achievement is compiled in the aggregate and by subgroup and applied to the state's AYP determination process. What follows is a detailed summary of the process.

2.1 Retaining the Current Accountability System's Essential Design

Section 1.4.3 above presents an overview of the state's current accountability structure. The proposed growth-to-standard model is designed to complement the foundational principles established within the current accountability structure and not to introduce confounding metrics that might obscure policy aims or confuse the educational community or the public.

The proposed growth-to-standard model will retain the current accountability system's essential design. The model will be integrated into the state's current accountability rules as a final phase

in determining AYP for every public school, school district, and the state. The state would continue to conduct a status analysis, followed by a safe harbor analysis, followed by the Title I privilege analysis, all conducted according to currently approved rules, including the use of a minimum N value of 10 and the state's approved confidence interval, set at 99 percent. If following this current series of analyses a school or district did not meet adequate yearly progress, then the state would employ this growth-to-standard model. As a further instructional administrative aid to schools, the state will conduct a growth-to-standard analysis for all schools, regardless of their AYP status, for the purposes of student services support, school improvement activities, and future accountability reference. This model provides additional value to the state accountability system by enhancing the state's currently viable accountability system and by further supporting the instructional supports to schools and districts.

2.2 Amending Growth into the Current Accountability System

The state proposes to amend its current accountability system by integrating current status measures with rigorous student growth measures into a unified metric of school and district performance. The state will retain its current core design, as outlined in section 1.4.3; however, the state will amend the system's design by introducing a new, final phase to the AYP determination process that performs comparably to current design.

The state's current design, in steps 1-5, references only student data based on actual student proficiency rates, according to the current status rules. The state proposes to introduce a sixth step to the AYP process that replicates the first five steps, but does so with student achievement data that recognizes sufficient student growth. The following chart illustrates this amendment.

Table 2.2: Summary of the Proposed State AYP Determination Process

<p>In calculating AYP determinations, schools, districts, and the state receive the benefit of a six-tiered review, which includes the following:</p> <ol style="list-style-type: none"> 1. Review current year status achievement data with a reliability test; 2. Review two years status combined achievement data with a reliability test; 3. Review three years status combined achievement data with a reliability test; 4. Determination of Safe Harbor; 5. Review targeted Title I student achievement data according to rules that replicate steps 1-4; 6. Review Growth-to-Standard achievement data <ol style="list-style-type: none"> a. Review current year achievement data incorporating growth benefit, with a reliability test; b. Review two years combined achievement data, including current-year-only growth benefit, with a reliability test; c. Review three years combined achievement data, including current-year-only growth benefit, with a reliability test; d. Determination of Safe Harbor, including current-year-only growth benefit; e. Review targeted Title I student achievement data according to rules that replicate steps 6a-6d above. <p>If a school passes any of these six levels of review, the school is identified as "<i>Met adequate yearly progress.</i>" If a school does not pass all six levels of review, as is applicable to the school, then the school is identified as "<i>Did not make adequate yearly progress.</i>" This multi-step determination process ensures that any designation of "<i>Did not meet adequate yearly progress</i>" is reliable and constitutes a warranted identification.</p>
--

The state stipulates that all the various procedural principles and rules contained within the state's current accountability workbook would remain unaltered. The state will faithfully administer the core accountability principles and indicators, including the inclusion of all schools and districts into all determinations; the uniform application of criteria to all schools and districts; the incorporation of valid and reliable achievement standards; the timely reporting of all information; the issuance of report cards; the management of rewards and sanctions; the inclusion of all students; the consistent application of a full academic year; the inclusion of mobile students; adherence to the goal of universal proficiency by 2014; the reporting of aggregate and subgroup results; adherence to intermediate measurable achievement targets; the issuance of annual determinations; the reporting of student achievement results; the reporting of high school graduation results; the reporting of student attendance results for elementary and middle schools; the dedication to reading and mathematics results; the validation of valid and reliable determinations; the minimal 95% participation rate of students on the state assessments; and other special circumstances rules contained within the state's workbook.

The proposed amendments to the state's accountability system restrict themselves to the administration of a growth-to-standard model for determining student achievement in terms of the state's prescribed achievement standards and against the identified intermediate achievement targets for schools and districts. The following section presents a detailed analysis regarding the operations of this model.

2.3 Assigning Student Achievement with a Growth-to-Standard Model

The state's proposed growth-to-standard model uses actual student achievement data from the NDSA and NDAA to establish and track student performance such that every student evidences sufficient yearly gains to reach or surpass proficiency within three years. The state does not annually reset the projected intersection with proficiency for a student. The student is expected to evidence annual gains to meet or exceed proficiency within the three years allowed.

Annualized growth to proficiency targets. The growth-to-standard model establishes a projected growth projection for every individual student in reading and mathematics, respectively. Grades 3 and 11 constitute baseline reference years for recording initial or final student achievement; however, no growth projections are calculated on grades 3 or 11 data alone. The proposed growth-to-standard model herein presents up to three years of historical data for demonstration purposes. As the model's administration matures into the future, it is expected that up to three years active student achievement data may be accessed for determination purposes. A student's initial projected proficiency target and the student's calculated annual required growth in scale scores will be retained for any future annual determinations. A student must maintain their progress to proficiency within the allowable time to be credited as proficient-by-growth.

The calculation process is simple, straightforward, and requires basic mathematics. The proposed model references the previous year's scale score for each student and subtracts that scale score from the current year's scale score. Within the NDSA, if the difference between the previous and current years' scores is equal to or greater than one-third of the difference of the *proficient* cut score value in three years, then the student is recorded as *proficient* (i.e., proficient-by-growth) for accountability purposes. If the difference is less than one-third the difference of the *proficient* cut score value in three years, then the student's respective

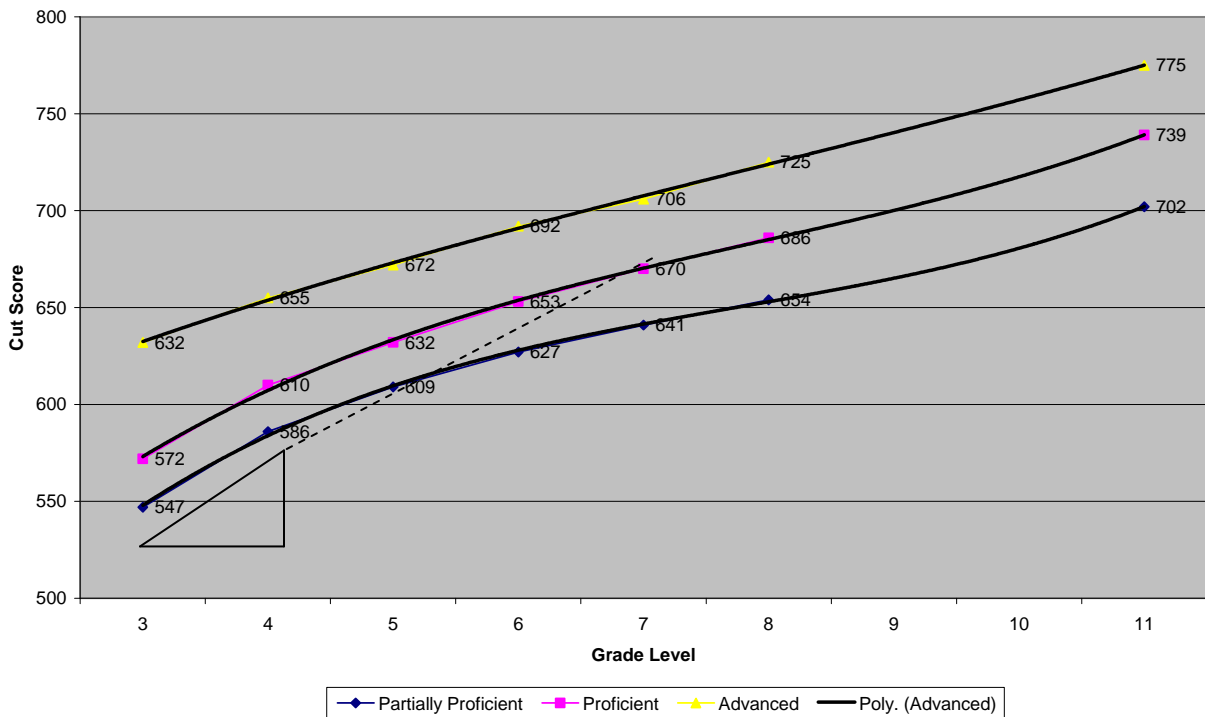
achievement level remains unaltered. For each subsequent annual determination, the original three-year goal remains as the target; no annual resetting of the target occurs.

Within the NDAA, the state will credit as sufficient growth any student who has moved from the *novice* achievement level (the lowest below-proficient achievement level) to the *partially-proficient* achievement level (the highest below-proficient achievement level).

Figure 2.3 below illustrates the principles underlying the state’s growth-to-standard model.

An illustrated example. If a student, for instance, were to score 525 on the NDSA mathematics assessment in grade 3 (designated as below-proficient) and were to score 570 on the NDSA in grade 4 (designated as below-proficient), then the NDDPI would project a three-year trajectory, from grade 4 to grade 7, to meet an expected 670 scale score, the *proficient* cut score. The student must minimally demonstrate annual gains of one-third the difference from the grade 4 score (570) and the grade 7 score (670) to remain on the trajectory’s path. In this case, the student must increase their score 100 points over three years, with a minimal annual expected gain of 33 points ($100/3=33.3$). Since the student gained 35 points from the previous year, that gain would be determined to be sufficient and the student would be identified for accountability purposes as *proficient* (i.e., proficient-by-growth).

Figure 2.3: NDSA Cut Scores by Achievement Level, Mathematics



The student’s actual growth from grade 3 to grade 4 exceeds the annualized three-year projected growth to the proficiency target in grade 7. The student is credited as proficient (i.e., proficient-by growth). If the actual student growth were less than the annualized projected growth, then the student would not be credited as proficient and would remain at their below-proficient standing.

Illustrated Example Expanded: As this student progresses each year, subsequent growth analyses might reveal a steady record of movement toward proficiency by grade 7. To remain on target for meeting proficiency and receiving credit for sufficient growth (i.e., proficiency-by-growth), based on the expected annualized increase of 33 scale points, this example student must meet or exceed the scale scores identified below.

Table 2.3(a): Expected Annual Growth for an Individual Student Example Across Grades 4-7, NDSA Mathematics Scale Scores

Grade	Expected Scale Score	Achievement Level
3	525 (base actual score)	Novice
4	560	Novice
5	593	Novice
6	626	Novice
7	660	Proficient

Table 2.3(b), presented in the Appendix, additionally provides a variety of student examples that illustrate how the growth-to-standard model might determine sufficient growth, based on students' unique performance on the state assessments, their enrollment status, or their apparent achievement projection.

Certain policy statements guide the design and operation of the growth-to-standard model. Some of these statements follow.

Full academic year. A student must be enrolled in their school district for a full academic year to be included in any growth-to-standard analysis for the purposes of accountability reporting. The state will provide, however, a growth analysis for all students, regardless of the longevity of their enrollment status to the school and district for school improvement purposes. The state's full academic year rules reference only school district enrollment for one year; it does not make any reference to enrollment within a specific school. For a student who changes residence from one school district to another, the state's full academic rule is employed; the student's growth projection would not be considered for accountability purposes, but the student would receive growth projection information for the purposes of instructional support. Any student who has been retained in their previous year's grade level will not be included in any growth projections for accountability purposes; however, the student will receive growth information compared to their previous year's achievement at the same grade level for instructional support.

Establishing baseline growth targets. The NDDPI has prepared this proposal based on multi-year historical student achievement data and has presented the 2007-08 state assessment data as the basis for this proposal. However, the NDDPI proposes to initiate the growth-to-standard model during the 2008-09 academic year. The state will calculate student growth targets based on evidenced growth from the 2007-08 state assessments to the 2008-09 state assessments. Student growth targets will be set according to the following format.

Table 2.3(f): Projected Proficiency Target Grades, Effective 2008-09

Enrolled Grade, 2008-09	Must Reach Proficiency by this Grade
3	Baseline Achievement Set
4	7
5	8
6	9
7	10
8	11
11	11

These prescribed grade targets standardize the growth determination process statewide and apply to all students in all schools. Such a prescribed target scenario provides a common reference point, eases the task of calculating targets among educators, allows for simpler explanations among parents, and removes any prospect of interruptions to the determination process. New targets will be established for any incoming grade 3 students and for students who might leave the state educational system and then return at a later time. Targets remain in effect for students who move among schools within the same school district, since the state's full academic year references district enrollment, not a specific school's enrollment. A new target is set for any student who might move to another school district to allow for the full academic year rule. This model of target determination will include the most students statewide and optimize the growth model's impact.

Interpolated targets in non-tested years. The state interpolates proficiency scale score targets for grades 9 and 10, which are not otherwise assessed or where cut scores have not been established, to allow for a more even spread of proficiency targets among the grades. Interpolation of proficiency target scores are based on the best-fit regression analysis for the proficiency lines drawn within reading ($y = 0.2409x^3 - 4.1864x^2 + 31.203x + 582.59$) and mathematics ($y = 0.3394x^3 - 5.978x^2 + 49.648x + 529.1$). Interpolating results allows for students in grades 6 and 7 to retain a three-year target goal against proportional targets. Since the state does not currently assess students in grades 9 and 10, students will not ultimately be required to demonstrate actual proficiency until grade 11. Nevertheless, all students are expected to meet proficiency by the year 2014.

Intermediate achievement targets. The state's growth-to-standard model references all school achievement rates to the state's established intermediate achievement standards. As the intermediate achievement standards increase incrementally, so also do the expected achievement targets for all schools and school districts.

Unmatched students. During any growth-to-standard calculations, any student who was not in the state's assessment system in the previous year will be removed from any calculations, since there exists no previous data upon which to calculate a projection. Such students will be counted at their current performance level. Assessment reliability studies indicate that the most reliable comparison of scale scores exists between contiguous years, where reliability statistics indicate favorable comparisons, and not among more widely spread comparisons. Grade 3 scale scores will serve only for the purposes of establishing growth projections beginning in grade 4; the state will not calculate non-empirically based projections at grade 3.

Proficient and advanced students. Students whose scores meet or exceed the proficiency cut score for their respective grade level prior to any growth analysis will be recognized as proficient

outright and any growth value generated by an analysis of their scores will not be applied to another student.

Minimum group size. The state will retain 10 as the current minimum group size, which was established to protect student identification under FERPA rules. No other minimum group size limits apply for any school organizations or subgroups.

Confidence interval. Since the state has established through previous federal review a confidence interval, set at 99%, for the state's accountability system, as outlined in Table 1.4.3(c), and since the same foundational AYP process is proposed with the use of the growth-to-standard model, as outlined in Table 2.2, the state stipulates that the application of the confidence interval is entirely consistent in design and intent with the growth-to-standard model application as it is in the current status model. The state asserts its intent to apply the 99% confidence interval for the purposes of AYP accountability determination. Refer to the Appendix, *Confidence Interval*, section for the state's narrative and data in support of this policy.

Actual Student Growth Rate Impact Data. The state's growth-to-standard model calculates every student's respective growth to determine if that growth is sufficient to meet or exceed proficiency within three years. Every student's growth target and target progression is calculated. A below-proficient student's achievement level is reassigned as *proficient* if his/her actual measured growth equals or exceeds that of the student's expected annualized growth. Tables 2.3(d-f), presented in the Appendix, provide actual student growth rate impact data generated from the 2007 NDSA and NDAA. Table 2.3(d) presents the number and percent of students tested in fall 2007 whose scale score increased by at least the targeted amount, thereby making sufficient growth. Table 2.3(e) presents the number and percent of students tested in fall 2007 whose scale score increased by at least the targeted amount, thereby making sufficient growth, by achievement level. Table 2.3(f) presents the number and percent of students tested in fall 2007 whose scale score increased by at least the targeted amount, thereby making sufficient growth, by subgroup classification. These data indicate that representative students across all grades, achievement levels, and subgroups evidence growth-to-proficiency patterns and that the state's longitudinal data system can successfully track such projections. Similarly, these data indicate that the state's system can identify and track growth progressions of students both above and below proficiency. This ability builds instructional value into the system, thereby enabling teachers to accurately monitor the overall trending of specific students. If students are collectively identified as being at risk for dropping out of proficiency, preventive strategies can be employed thereby forestalling and reversing any downward movement. Table 2.3(c), presented in the Appendix, provides a prototypic school report that lists the student-specific growth progressions within a given grade 4 mathematics class, following a cross-year growth analysis. Such a tool as this offers schools additional information to build instructional plans to aid students at risk for learning setbacks.

2.4 Considerations for the State's Alternate Assessments

The NDDPI has developed and administers two independent forms of the NDAA: (1) the NDAA1, based on alternate achievement standards, and (2) the NDAA2, based on modified achievement standards. The NDAA reports student achievement according to the state's recognized achievement levels: *novice*, *partially proficient*, *proficient*, and *advanced*. The *novice* and *partially proficient* achievement levels correspond to below-proficient. The *proficient* and *advanced* achievement levels correspond to proficient and above. Since the NDAA does not incorporate vertically aligned scale scores, the state's growth-to-standard model does not recognize scale score movement within the NDAA. Instead, within the growth-to-standard model,

student growth will be defined as the reported movement from one achievement level to another, either positive or negative.

Within the AYP determination process, the state will only credit a school when a student records an upward movement from *novice* to *partially proficient*. No other movement up will be recognized, since such movement will amount to a proficient or above score. In the event that a student were to take the NDAA in year-one and then take the NDSA in year-two (or vice versa), a school will only be credited for growth if the student were to progress upward one achievement level from *novice* to *partially proficient*, from year-one to year-two. The NDDPI will not equate actual scale scores or performance scores between the NDSA and the NDAA.

2.5 The Effect of the Growth-to-Standard Model on Adequate Yearly Progress

The state's growth-to-standard model calculates every student's respective growth to determine if that growth is sufficient to meet or exceed proficiency within three years. Every student's growth target and target progression is calculated. A below-proficient student's achievement level is reassigned as *proficient* if his/her actual measured growth equals or exceeds that of the student's expected annualized growth. When the calculations of growth and the reassignment of *proficient* for below-proficient students are completed, the NDDPI proceeds to determine AYP according to the rules set forth within the state's amended accountability workbook, as outlined within Table 2.2. The growth-to-standard model solely focuses on reassigning proficiency based on growth; once reassignments are completed, the growth-to-standard model's functions are likewise completed. The state's AYP determination process takes over and calculates the AYP status of every public school, school district, and the state.

Reassigning the achievement level of below-proficient students based on actual measured growth can and does affect AYP determination rates. Table 2.5, presented in the Appendix, provides a comparison of the number of AYP failures between the state's current status model and the proposed growth-to-standard model, by individual composite and subgroup indicator for math and reading. The net effect of reassigning proficiency-for-growth reduces the number of schools that do not make AYP, from 169 to 133, a 21.3% reduction. Growth-based reassignments impact reading results more than mathematics.

The NDDPI places no relative value, importance, or goal regarding the degree of reductions generated by the growth-to-standard model. Of preeminent concern to the NDDPI are the principles upon which the growth-to-standard model is based and the integrity of its methodology. The NDDPI asserts that the AYP determination rate evidenced by the growth-to-standard model is appropriate and defensible.

III. Adherence to the Seven Core Principles

The NDDPI stipulates that the state's proposed growth-to-standard model of accountability fully meets the seven core principles put forth from the ED. The following sections address each respective principle and how the state's model specifically enacts these assurances.

3.1 Universal Proficiency by 2014 and School Accountability

The state proposes a growth-to-proficiency model based solely on student achievement results generated by the NDSA and the NDAA. The state's model places a three-year proficiency target projection for all students. It is the clear policy of the state that all students meet proficiency by

2014. The state proposes an integrated accountability system that combines status, Safe Harbor, and growth components to produce appropriate AYP determinations for all public schools, school districts, and the state. The state's growth-to-standard model complements the state's current status model, including Safe Harbor, for reading and mathematics in grades 3-8 and 11. All grades are used in some manner to establish either baseline data or growth projections. The NDDPI includes up to three years achievement data to generate the required projections or AYP determinations. Achievement data from all grade levels are referenced at some point in the projection or AYP determination process, fully accessing the established capacity of the state's longitudinal data system.

Section 2.0 presents in detail the manner in which growth targets and projections for individual students are calculated and incorporated into the sixth phase of the state's AYP determination process. The growth-to-standard model requires that actual student growth intersect with proficiency within three years; it does not credit any growth that falls short of intersection with proficiency. Any application of a growth-to-standard component within a school's AYP determination must be referenced to the state's intermediate achievement targets, as required by law. The state's intermediate achievement targets are clearly specified in section 1.4.3. The state's model requires students to continually progress through challenging grade-level expectations through to high school. The state's model actively advances such student grade level progression across all grades and does not sanction back-loading performance. Delays in achievement are not encouraged or credited; in fact, the system's design intentionally pushes students and schools to achieve proficiency within a limited time frame.

The state steadfastly advances the application of a 99 percent confidence interval to properly administer the state system's reliability functions and to sustain the integrity of the overall accountability system. Appendix A provides a detailed analysis of the merits of this provision. The state's model retains the now established and accepted components of the state's accountability system. The state attends to composite and subgroup determinations and provides student growth information to schools to advance quality school improvement activities. The accountability system, as presented in section 1, maintains the current school improvement and corrective action options provided under Section 1116 of ESEA. The state's model places high priority on generating student growth data at the school level for the purposes of improved instruction. This constitutes the model's more important value-added measures with potential to transform the use of the NDSA and the NDAA as recognized tools and references for standards-based educational enhancement.

3.2 Establishing Student-Appropriate Growth Targets

The state's growth-to-standard model bases annual growth on student progression along the vertical scale of the NDSA and the advancement of students from the novice to the partially-proficient achievement level on the NDAA, as outlined in section 1.4.6. The strong reliability between contiguous grade levels on the NDSA provides confidence in growth estimates. Scale score progression among grades on the NDSA allows confidence in appropriately measured, inter-grade progressions. All growth progressions are intentionally referenced to the NDSA's scale score and the state's approved achievement standards. The state's model establishes growth expectation measures that apply equally to all students, without regard for student demographics or school characteristics. The state's model uses these common expectation measures to calculate student-specific targets. The state's growth expectations remain fixed to grade-level achievement expectations and not to any prototypic growth patterns. The state's model does not allow the growth of certain above-proficient students to be applied to students below proficient. The state produces annual growth reports on all students to ensure that

students are progressing along their expected projections and any instructional plans are re-evaluated.

3.3 Separate Accountability for Reading and Mathematics

The NDDPI stipulates that the state's accountability system produces separate accountability decisions about student achievement in reading and mathematics. This stipulation relates to both the state's status and growth components. Section 1.4.6 and section 2 provide detailed analysis of the manner in which measurement error is accounted for and systemic validity and reliability are ensured. The state applies a 99 percent confidence interval on all AYP considerations for all schools and subgroups, regardless of size. Appendix A presents a detailed rationale for retaining this measure.

3.4 Inclusion of All Students

The NDDPI stipulates that the state's accountability system and this proposed growth-to-standard model includes all students in all tested grades, including both the NDSA and the NDAA, and that the NDDPI can account for all public schools, school districts, and student subgroups.

Section 1.4 and section 2.2 present in detail the manner in which the state's enrollment rules and longitudinal data system identify and track individual students across grades and accommodates enrollment movement across district lines. The proposal includes a variety of tables that demonstrate the capacity of the longitudinal data system to match and track individual students across multiple years. When student matching is not possible, the student will not be included into any growth determinations. The state's full academic year policies set the general method for reporting district- and school-based growth. Students who remain within a school district will have their growth data included into any accountability determinations; students who transfer to a new school district will not have their growth included into accountability reporting, but the growth will be reported for school improvement and instructional purposes.

The state's model includes and fully accounts for all student subgroups. The state's model employs a confidence interval to optimize the full-inclusion of subgroups, while ensuring systemic reliability. The NDDPI will recognize a student's most current subgroup classification, if the student's classification were to change during the tracking time period.

The state's model includes all public schools and school districts into accountability determinations. Since the growth-to-standard model represents the sixth, and final, phase in the AYP determination process, all schools and districts will be reviewed for growth and any appropriate findings will be applied to the final determination. No schools or districts will be excluded from growth determinations; the manner in which this determination process unfolds is outlined in section 2. The state's accountability workbook provides the operational rules that guide all AYP determinations.

3.5 Compliant State Assessment System

The NDSA and the NDAA based on alternate achievement standards have been awarded a *Fully Approved* status from the ED in June 2007 for fully meeting all requirements of the federally prescribed peer review process. The state's NDAA based on modified achievement

standards is currently undergoing federal peer review but has been previously approved by the ED for inclusion in determining annual adequate yearly progress.

The design and capacity of the NDSA and NDAA to meet fully the requirements of a growth-to-standard model is presented throughout section 1 and section 2 of this proposal. The NDDPI administers the NDSA and the NDAA in reading and mathematics in grades 3-8 and 11 and references multiple years of data to produce the state's current AYP determinations. The state assessment system produces achievement reports for individual students, schools, school districts, and the state. These data are disaggregated and reported for student subgroups. Within the growth-to-standard model, as specified in section 2.3 and its supporting tables, individual student growth information will be provided to the school and the school will be required to distribute the results to parents. These student reports identify each student's base achievement level, projected target, expected annualized growth, and trending data.

The growth-to-standard model references the state's documented vertical scale alignment among the various grade levels in both reading and mathematics to appropriately calculate and report student growth projections. The state's assessment system design has been stable since 2003 and is expected to remain as such based on forthcoming Request for Proposal specifications that require equating of scale score methods. Any forthcoming changes to the state's assessment system that might result from a change in vendor or testing model will be structured to allow for appropriate equating and seamless reporting. The NDDPI fully anticipates that the state's assessment system will remain effectively stable in forthcoming years. There should be nominal impact to the operation of the state's growth-to-standard model.

3.6 Tracking Individual Student Progress

The state's longitudinal data system is capable of fully managing student data files across multiple years for the purposes of administering the state's growth-to-standard model. Section 1.4 provides an overview of the longitudinal data system and its capabilities. The NDDPI manages a statewide student identification system that allows for the dependable matching of student records across data files and across multiple years. The various tables provided within section 1.4 demonstrate that student files can be matched and allow for the tracking of students by various subgroup indicators across multiple years. The state employs internal quality assurance checks and external validation opportunities by local schools to ensure that data are accurately managed. The state reports student information across the full range of student subgroups. When student matching is not possible, the student will not be included into any growth determinations. The state's full academic year policies set the general method for reporting district- and school-based growth. Students who remain within a school district will have their growth data included into any accountability determinations; students who transfer to a new school district will not have their growth included into accountability reporting, but the growth will be reported for instructional and school improvement purposes. The NDDPI has managed the state's longitudinal data system for several years for the purposes of making reliable AYP determinations, among other reporting requirements. The NDDPI is capable of administering the proposed growth-to-standard model.

3.7 Participation Rates and Additional Academic Indicators

The NDDPI will administer participation rate and secondary academic indicator reporting as stipulated within the state's current accountability workbook. The NDDPI does not foresee the prospect of any alteration in the calculation or reporting process of participation or the secondary indicator, given the design of the growth-to-standard model as a final phase in the

current determination process. The NDDPI will be able to provide a full accounting of all enrolled students in the determination process and subsequent reporting of results.

Conclusion

The NDDPI seeks authorization to incorporate this proposed growth-to-standard model of student achievement within its current state accountability system. This growth-to-standard achievement model will provide an additional means of measuring student growth in relation to the state's challenging achievement standards and reporting overall student proficiency and growth-to-proficiency for every public school, school district, and the state.

This proposal is consistent with the principles and practices of the State's currently approved federal accountability workbook rules. Additionally, this proposal employs student achievement and demographic data derived from the state's assessment system, which received a *Fully Approved* status from the ED in June 2007. This proposal has been adopted based on its simple and straightforward design, its transparent and verifiable metrics, and its ready accessibility to educators, parents, and the public.

The tables and figures presented throughout this proposal's narrative and appendices have been generated from a variety of data files. In the interest of meeting the specified space limitation within this submission, the data files have not been included. The NDDPI commits itself to provide to the ED any and all files that the ED requires to make a full and complete determination of the proposal's value. The NDDPI will accommodate any such ED requests.

The NDDPI places this proposal before peer review with full confidence that the state's fully approved assessment system, longitudinal data system, reliable accountability rules, and commitment to the incorporation of growth analysis into instructional design, provide systemic integrity and will advance the quality of education across North Dakota.

Appendix A: Confidence Intervals

It is the expressed aim of the NDDPI to build parity, in design and effect, between the state's status and growth-to-standard AYP determination processes. To accomplish this aim, the NDDPI declares its intent to apply a 99 percent confidence interval within the growth-to-standard AYP determination process that is commensurate to the application of the 99 percent confidence interval used currently within the state's status AYP determination process. Confidence intervals are not used in the calculation of Safe Harbor.

The state's current use of a 99 percent confidence interval within status AYP determinations is recognized as consistent with federal regulations, has received general acceptance for its appropriateness as a measure of assurance, and has been officially reviewed and approved by the ED for its inclusion within various states' accountability workbooks. The ED, following an extended validation process in 2003 and reaffirmed annually, has recognized the appropriate use of a 99 percent confidence interval within the North Dakota accountability plan. The use of a 99 percent confidence interval constitutes official public policy, without condition.

A Question of Legitimacy

In asserting its privilege to use a 99 percent confidence interval in calculating growth-to-standard AYP determinations, the NDDPI acknowledges the oppositional standing of the growth model peer review committee on this issue. The peer review committee has issued unofficial guidance--guidance that does not constitute official ED policy--that generally discourages the use of confidence intervals, conditionally limits any use of confidence intervals to 68 percent, and outright prohibits the use of a 99 percent confidence interval. The peer review guidance characterizes the application of confidence intervals in status and growth determinations as follows:

“The justification for employing confidence intervals around the AYP status target is based largely on reducing the impact of score volatility due to changes in cohorts being assessed from one year to another, and thus reducing the potential for inappropriately concluding that the effectiveness of the school is improving or declining. Under the growth model the issue of successive cohorts is no longer in play since we are measuring the gains over time that are attained by individual students.” *(Summary by the Peer Review Team of April 2006 Review of Growth Model Proposals, May 17, 2006)*

In issuing its unofficial guidance on the use of confidence intervals, the peer review committee presents its reference point for or its characterization of the issue at hand.

Ever mindful of the difficult task the peer review committee has assumed in establishing operational procedures and conducting individual state reviews, the NDDPI respectfully disagrees with and objects to this characterization of the use of confidence intervals within accountability systems, in general, and within the NDDPI's growth-to-standard model, in particular. The position assumed within the peer review committee's unofficial guidance inadequately represents the application of confidence intervals within North Dakota's accountability system. The NDDPI further asserts that in so delimiting the application of a confidence interval and completely rejecting the use of a 99% confidence interval, the peer review committee's unofficial guidance inappropriately intrudes into its contractually-secured

accountability rules, compromises the carefully crafted integrity of the state's accountability system, and establishes an inadequate, unexamined precedent regarding the meaning of systemic validity and reliability.

The North Dakota Context

Within the North Dakota accountability workbook, the confidence interval provides a universally applied statistical means of ensuring that all AYP determinations are objectively and proportionally applied to all public schools, school districts, and subgroups. The state selected a base N value of 10 to provide protections against student identification, pursuant to the privacy provisions of the *Family Education Rights and Privacy Act* (FERPA). The state intentionally rejected the inclusion of any additional N value at the school or subgroup level (e.g., setting a value of 30 or 40 students) for the purposes of limiting the application of accountability reporting under the guise of reliability. The state eliminated the prospects of these higher limits to allow for a complete accounting of all schools and subgroups, while honoring FERPA privacy considerations. In establishing a higher N limit, the state would have effectively removed a large number of smaller schools or subgroups from the state's accountability system. Disparity in accountability would emerge: schools or subgroups immediately above the N value would undergo reporting while those schools or subgroups below the N value would remain exempt.

To mitigate the effects of these unfair reporting disparities, the NDDPI proposed and successfully defended the use of a 99 percent confidence interval. A confidence interval uniformly applies to all settings and provides appropriate and proportional protections to all. Additionally, as the state's intermediate achievement targets incrementally increase every three years, the mathematical properties of confidence intervals commensurately reduce the degree of protections afforded to schools, districts, and subgroups. Such incremental decreases in protections allow schools and districts to prepare for the inevitable and universal impact of 100 percent proficiency by 2014. The use of confidence intervals provides greater transparency in public reporting and constitutes a widely accepted means of measuring and confidently reporting school, district, and subgroup performance rates. The confidence interval is an integral component of the state's accountability rules and has been embraced as a critical public policy assurance.

A Question of Proper Perspective and Use of the Confidence Interval

Within its unofficial guidance, the peer review committee states: "Under the growth model the issue of successive cohorts is no longer in play since we are measuring the gains over time that are attained by individual students." The NDDPI respectfully submits that this statement mischaracterizes the nature of the state's growth-to-standard model and, by doing so, unnecessarily removes the legitimate application of a confidence interval.

The core benefit of the state's growth-to-standard model occurs at the very onset of any accountability reporting, prior to the administration of the state's AYP determination rules. The state's growth-to-standard model calculates every student's respective growth to determine if that growth is sufficient to meet or exceed proficiency within three years. Every student's growth target and target progression is calculated. A below-proficient student's achievement level is reassigned as *proficient* if his/her actual measured growth equals or exceeds that of the student's targeted annualized growth. When the calculations of growth and the reassignment of *proficient* for below-proficient students are completed, the NDDPI proceeds to determine AYP according to the rules set forth within the state's amended accountability workbook, as outlined within Table 2.2.

The growth-to-standard model solely focuses on reassigning proficiency based on growth; once reassignments are completed, the growth-to-standard model's functions are likewise completed. The state's AYP determination process takes over and calculates the AYP status of every public school, school district, and the state.

In effect, the growth-to-standard model only provides a mechanism to reassign student proficiency markers based on evidence of sufficient growth. Once all students' scores are reviewed and proficiency reassignments are complete, the student achievement database is ready for AYP analysis. Then the state's currently approved AYP procedures fall into place. The growth-to-standard model has done its job and is employed no further. Students in the database are identified as either proficient or below-proficient for the purposes of determining school, district, and subgroup achievement rates. These achievement rates are then compared to the state's intermediate achievement targets and a confidence interval is applied to ensure the sufficiency of the rate. The fundamental AYP procedure is unchanged from the current status model during these determinations.

The state's AYP process still references current and multiple-years' achievement data to calculate determinations. The AYP process compares cohort populations as before, as it must. In every regard, the determination process is the same. The state is still obligated to determine if a school or subgroup, regardless of its size, has made sufficient achievement in terms of the state's intermediate achievement target. This still requires a confidence interval. To not employ a confidence interval would deny schools and subgroups and the public sufficient reliability safeguards. The use of confidence intervals is required to properly conduct any AYP determination process. It is untenable to conduct any AYP determination process, within the state's system, without employing a confidence interval.

The Impact of the Confidence Interval on the Growth-to-Standard Model

The NDDPI has prepared this submission of a growth-to-standard model mindful of the peer review committee's unofficial guidance that discourages the use of confidence intervals, conditionally limits any use of confidence intervals to 68 percent, and outright prohibits the use of a 99 percent confidence interval. The NDDPI submits this growth-to-standard model with the clear conviction that the state's over-riding accountability policy goals and provisions require its adoption.

The NDDPI has designed the growth-to-standard model to establish parity, in design and effect, between the state's status and growth-to-standard AYP determination processes. The model's procedures are outlined in Section 2 of this submission.

To demonstrate the relative impact of the use of confidence intervals within the proposed growth-to-standard model, the NDDPI prepared four separate AYP determination scenarios:

- (1) use of the current status AYP determination process, without the use of a growth-to-standard model;
- (2) use of the growth-to-standard provision, using no confidence interval;
- (3) use of the growth-to-standard provision, using a 68 percent confidence interval; and
- (4) use of the growth-to-standard provision, using a 99 percent confidence interval.

Presented below is a table that lists the relative AYP determination findings by indicator and by overall school identification numbers.

Table CI: Comparing the number of schools that did not meet a certain achievement indicator based on either the current status model or the proposed growth-to-standard model

Indicators	Primary Indicators: Achievement							
	Reading				Mathematics			
	No Growth Model	Growth Model - No CI	Growth Model - 68% CI	Growth Model - 99% CI	No Growth Model	Growth Model - No CI	Growth Model - 68% CI	Growth Model - 99% CI
Overall	90	82	82	70	22	22	22	22
White	41	38	38	33	4	4	4	4
Native American	38	36	35	34	24	24	24	24
Black	7	7	7	7	5	5	5	5
Asian	1	1	1	1	0	0	0	0
Hispanic	5	4	4	4	0	0	0	0
Other	0	0	0	0	0	0	0	0
Low Income	105	96	95	85	29	29	29	29
Disabilities (IEP)	102	99	98	88	32	32	32	31
Limited English Proficiency (LEP)	22	20	20	19	16	13	13	13
Total	411	383	380	341	132	129	129	128

	No Growth Model	Growth Model - No CI	Growth Model - 68% CI	Growth Model - 99% CI
Number of schools that did not meet AYP:	169	157	156	133

The findings of these four scenarios report the number of identified AYP failures among the various composite and subgroup indicators. The base scenario is the “No Growth Model”, which reports the findings of the state’s current status AYP model. Within the “No Growth Model”, for example, 411 different reading indicators and 132 mathematics indicators were distributed among the 169 schools that did not make AYP. Statewide, 169 public schools of the state’s 470 public schools did not make AYP during the 2007-08 academic year, a 36% identification rate.

The three growth-to-standard scenarios, each based on different confidence intervals, report relative performances. The “Growth Model—No CI” (i.e., no confidence interval applied) reports almost identical results (157 school identifications) compared to the “Growth Model—68% CI” (i.e., 68 percent confidence interval applied), which reports 156 identified schools. Both the “Growth Model--No CI” and the “Growth Model—68% CI” evidence an approximate 33% school identification rate. The “Growth Model—99% CI” (i.e., 99 percent confidence interval applied) produces 133 schools that do not make AYP, a statewide 28.3% school identification rate.

It is not unexpected that a 68 percent confidence interval applied on the currently high intermediate achievement levels (approximating 80 percent proficiency targets) evidences negligible impact on AYP identifications. The “No CI” and the “68% CI” are effectively the same in impact. These data would indicate that a 68 percent confidence interval produces no effect and constitutes an ineffectual and disingenuous option to states.

These data report the relative impact of various confidence interval scenarios on any AYP determination policy. The NDDPI places no relative value, importance, or goal regarding the degree of reductions generated by the growth-to-standard model. Of preeminent concern to the NDDPI are the principles upon which the growth-to-standard model is based and the integrity of its methodology. The NDDPI assesses the difference between the current status model's identification rate (36% statewide) and the 99 percent confidence interval growth-to-standard model's identification rate (28.3%) as not unreasonably disparate or evidencing inordinate protection to schools at the expense of other public policy concerns. The growth-to-standard model's identification rate appears proportional, measured, and reasonable.

Of greater paramount importance is the state's need to retain unity of purpose and design between the status component and the growth-to-standard model component of the accountability system. The system must apply uniform procedures to provide the profession and the public with a clear and consistent sense of the safeguards in place. The integrity of the overall system demands comparable statistical methods. No assurance is gained by introducing differing statistical measures for different components of the system. Such public policy assurances supersede any arbitrary alterations in the state's statistical methods.

The NDDPI asserts that the AYP determination rate evidenced by the growth-to-standard model, with a 99 percent confidence interval applied, is entirely appropriate and defensible.

Good Faith, Legal Standing, and the Public Good

The use of confidence intervals within the AYP determination process represents a tangible procedure to address the law's mandate to safeguard the validity and reliability of a state's accountability system. Concerns for validity and reliability underscore the importance of securing the integrity of measurement and reporting, and containing or mitigating any unintended consequences that might otherwise compromise the law's intent.

The carefully constructed principles and procedural safeguards contained within a state's accountability workbook, whose various provisions were negotiated and validated into a binding contract between the ED and the state, constitute foundational and binding public policy for the measurement and reporting of a state's educational accountability system. As such, these negotiated and contractual provisions must supersede any unilateral, externally imposed conditions by an unconstrained, unaccountable peer review committee, acting on unofficial guidance that lies outside the constraints of law or regulation and that undercut the contractual nature of a state's negotiated accountability workbook. No external party to a contract should redefine any terms of a contract, conditionally constrain elements of the contract's operation, or otherwise trifle with the provisions of a duly enacted contract of a sovereign state.

The unofficial peer review guidance on confidence intervals does not present extensive defense for its position and, as such, represents a general impression, not a studied analysis. The NDDPI asserts that such a general impression should never rise as an impediment to the state's over-arching policy goals, which have been negotiated and agreed to. There exists no legitimate, documented analysis that should impede the state's interests to sustain the integrity of its accountability system's methodology.

The NDDPI has faithfully administered all provisions of the state's accountability workbook and, in so doing, the law's directives. The NDDPI has placed before the peer review committee evidence of a good faith proposal to administer a reasoned and responsible growth-to-standard model. The NDDPI has been forthright in explaining the operation of the model and its relative

impact on student achievement rates and public school AYP identification rates. The model displays simple and straightforward design, transparent and verifiable metrics, and ready accessibility to educators, parents, and the public. The model blends with the state's policy goals and adheres to the seven principles of accountability espoused by the ED.

We welcome the prospect of implementing this model and sharing the results of its operation with the ED. This is a reasonable and responsible model worthy of approval.

**Appendix B:
Tables and Figures**

**Table 1.4.2(a): Number and percent of students whose fall 2007
and fall 2006 test records could be matched, by grade level**

Grade	Number of students tested in fall 2007	Number of fall 2007 students whose records could be matched to 2006 records	Percentage of fall 2007 students whose records could be matched
Grade 3	6840	N/A	N/A
Grade 4	6995	6609	94.48%
Grade 5	7035	6671	94.83%
Grade 6	7071	6695	94.68%
Grade 7	7305	6851	93.79%
Grade 8	7475	7136	95.46%
Grade 11	7723	N/A	N/A

**Table 1.4.2(b): Number and percent of students whose fall 2007
and fall 2006 test records could be matched, by subgroup***

Subgroup	Number of students tested in fall 2007	Number of fall 2007 students whose records could be matched	Percentage of fall 2007 students whose records could be matched
Asian American	332	275	82.83%
Black	718	598	83.29%
Hispanic	681	555	81.50%
Native American	3220	2897	89.88%
White	30728	29514	96.05%
Other Ethnicity	202	126	62.38%
Females	17378	16447	94.64%
Males	18486	17504	94.69%
Non-IEP	30592	28972	94.70%
IEP	5289	4990	94.35%
Non-LEP	34569	32779	94.82%
LEP	1312	1183	90.17%
Non-NSLP	24020	23006	95.78%
NSLP	11861	10956	92.37%
Non-Migrant	35782	33897	94.73%
Migrant	99	65	65.66%

*Students in grades 3 and 11 are excluded from subgroup analyses

**Table 1.4.2(c): Number and percent of students whose fall 2007
and fall 2006 test records could be matched, by reading achievement level**

Achievement Level	Number of students tested in fall 2007	Number of fall 2007 students whose records could be matched	Number of fall 2007 students whose records could be matched
Novice	2644	2429	91.87%
Partially Proficient	6698	6339	94.64%
Proficient	19878	18939	95.28%
Advanced	6222	5997	96.38%

**Table 1.4.2(d): Number and percent of students whose fall 2007
and fall 2006 test records could be matched, by mathematics achievement level**

Achievement Level	Number of students tested in fall 2007	Number of fall 2007 students whose records could be matched	Number of fall 2007 students whose records could be matched
Novice	2161	1936	89.59%
Partially Proficient	5935	5580	94.02%
Proficient	18286	17474	95.56%
Advanced	9145	8786	96.07%

**Table 1.4.2(e): Number and percent of students whose fall 2007
and fall 2004 test records could be matched, by grade level**

Grade	Number of students tested in fall 2007	Number of fall 2007 students whose records could be matched	Number of fall 2007 students whose records could be matched
Grade 3	6840	N/A	N/A
Grade 4	6995	N/A	N/A
Grade 5	7035	N/A	N/A
Grade 6	7071	5913	83.62%
Grade 7	7305	6072	83.12%
Grade 8	7475	6287	84.11%
Grade 11	7723	6546	84.76%

**Table 1.4.2(f): Number and percent of students whose fall 2007
and fall 2004 test records could be matched, by reading achievement level**

Achievement Level	Number of students tested in fall 2004	Number of fall 2004 students whose records could be matched	Percentage of fall 2004 students whose records could be matched
Novice	1696	1240	73.11%
Partially Proficient	5639	4561	80.88%
Proficient	16700	14373	86.07%
Advanced	5246	4569	87.09%

Table 1.4.2(g): Number and percent of students whose fall 2007 and fall 2004 test records could be matched, by mathematics achievement level

Achievement Level	Number of students tested in fall 2004	Number of fall 2004 students whose records could be matched	Percentage of fall 2004 students whose records could be matched
Novice	1782	1265	70.99%
Partially Proficient	5500	4367	79.40%
Proficient	16508	14264	86.41%
Advanced	5510	4858	88.17%

Figure 1.4.6(a) Cross-Grade Articulation of Scale Scores at Selected Percentiles, NDSA Reading 2007

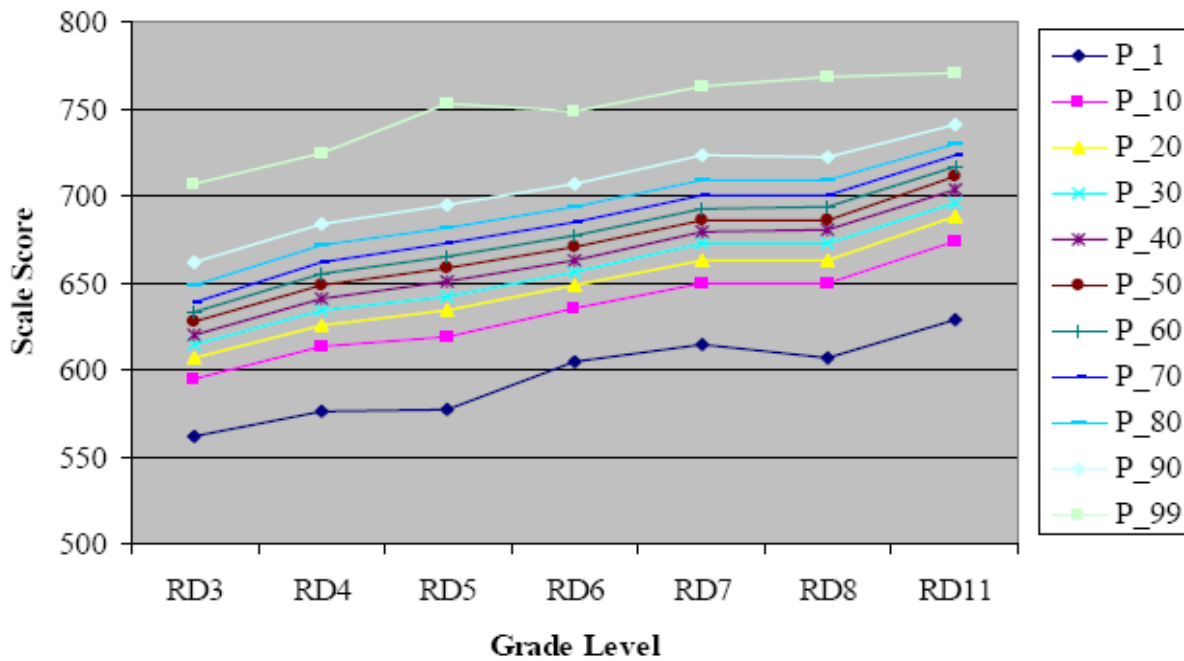


Figure 1.4.6(b) Cross-Grade Articulation of Scale Scores at Selected Percentiles, NDSA Mathematics 2007

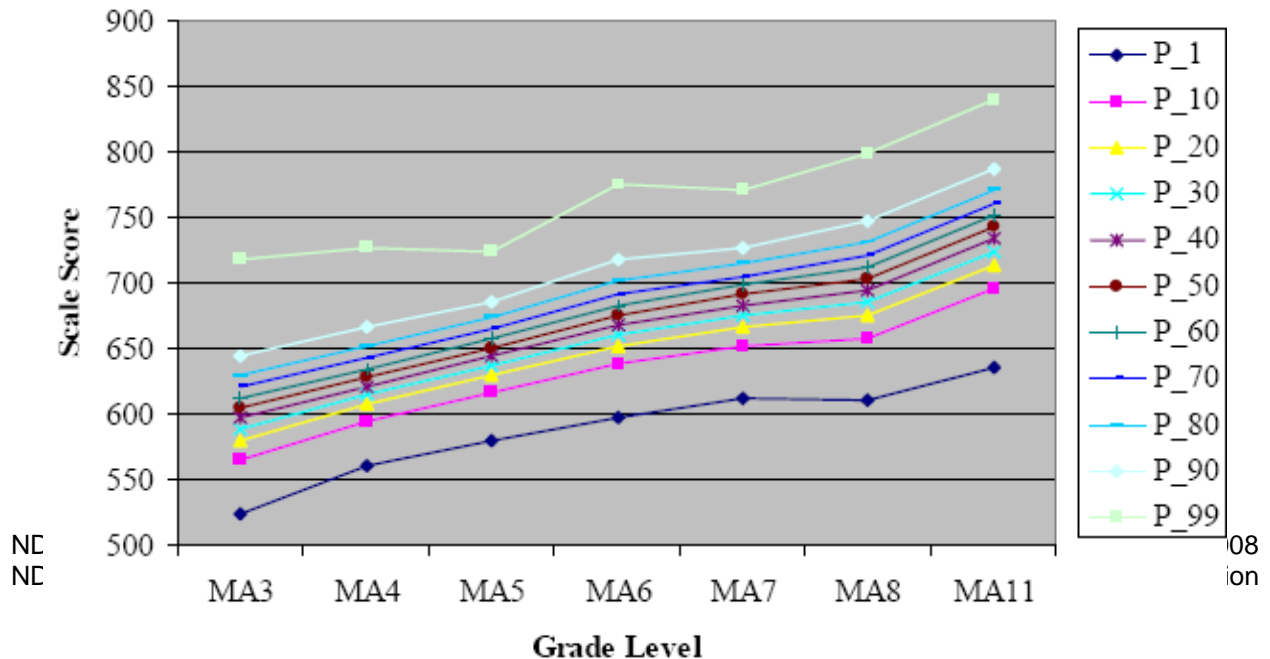


Table 1.4.6(a): Mean and Standard Deviation of Scale Scores for 2004-2007 by Grade/Content Area – Census (excluding home education students)

Content Area	Grade	2004		2005		2006		2007	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Reading	3	629.86	27.49	631.75	27.55	629.33	26.53	627.36	27.14
	4	648.04	28.14	649.58	29.84	652.74	31.57	648.04	29.55
	5	663.42	30.36	662.59	33.86	660.01	27.90	657.89	32.72
	6	672.55	28.10	671.70	30.67	674.02	28.10	670.82	29.23
	7	682.84	27.54	685.24	31.36	687.71	28.51	685.99	29.64
	8	685.20	29.36	684.51	28.58	687.40	29.77	685.98	31.42
	11	713.13	28.43	713.11	25.18	714.08	29.48	708.10	28.54
Math	3	605.13	36.06	607.18	35.63	607.66	37.39	604.78	34.44
	4	633.29	31.66	633.84	31.19	633.67	33.15	629.14	30.72
	5	649.54	30.60	653.02	31.54	656.86	30.91	650.65	28.62
	6	667.86	33.06	673.99	33.55	677.53	29.80	676.77	33.37
	7	680.76	32.13	682.98	32.76	690.75	33.73	689.93	31.80
	8	697.40	34.76	698.45	34.51	698.28	35.93	702.43	36.86
	11	741.12	40.13	744.34	39.71	743.85	39.01	741.47	38.73

Table 1.4.6(b): Reading Test Reliabilities

Grade	Number of Items	Number of Score Points	Stratified Alpha	Coefficient Alpha	Coefficient Alpha Male	Coefficient Alpha Female
3	60	63	0.87	0.88	0.88	0.87
4	63	67	0.88	0.88	0.88	0.87
5	60	63	0.87	0.86	0.87	0.85
6	60	69	0.89	0.89	0.89	0.88
7	60	68	0.90	0.89	0.90	0.88
8	62	67	0.89	0.90	0.90	0.89
11	60	63	0.90	0.90	0.90	0.89

Table 1.4.6(c): Mathematics Reliabilities

Grade	Number of Items	Number of Score Points	Stratified Alpha	Coefficient Alpha	Coefficient Alpha Male	Coefficient Alpha Female
3	60	63	0.86	0.86	0.86	0.86
4	60	64	0.87	0.85	0.85	0.85
5	60	64	0.86	0.86	0.86	0.85
6	60	64	0.88	0.88	0.88	0.87
7	60	65	0.90	0.90	0.90	0.90
8	60	65	0.90	0.90	0.90	0.89
11	60	65	0.92	0.92	0.93	0.91

Figure 1.4.6(c): Standard Error of Measurement Curves for Reading

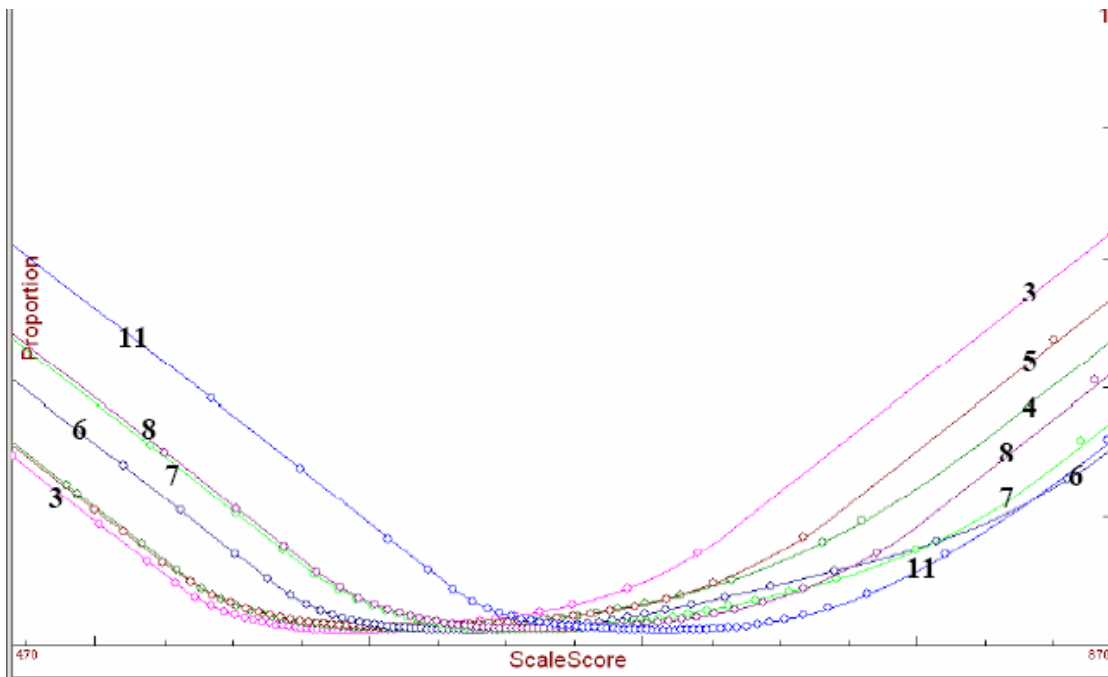
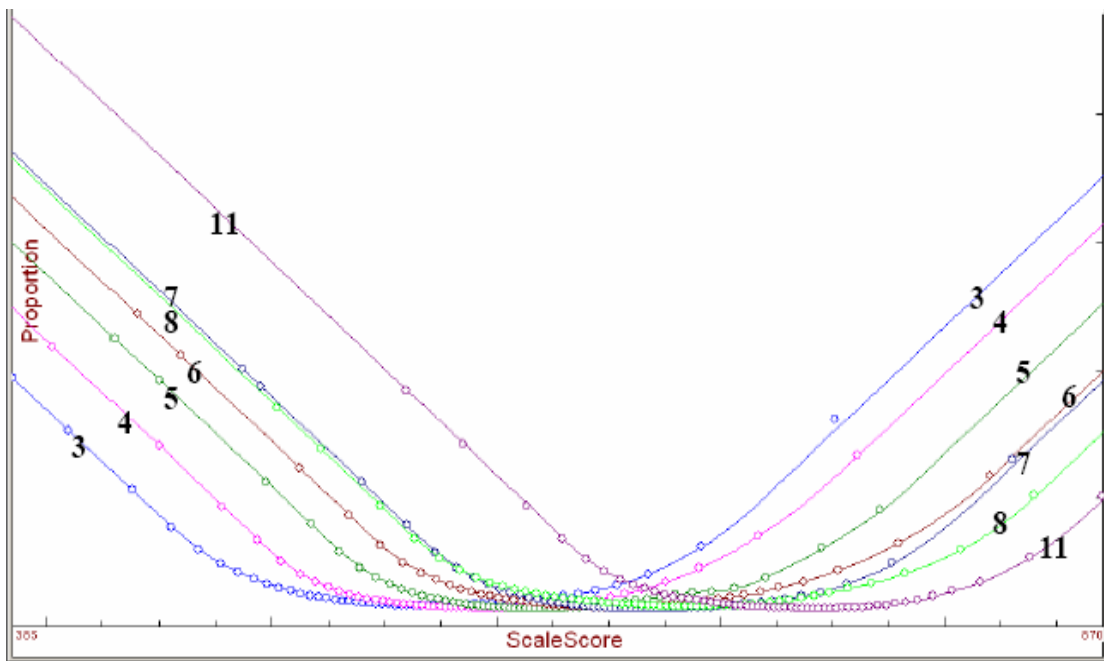


Figure 1.4.6(d): Standard Error of Measurement Curves for Mathematics



Complete NDSA validity and reliability statistics can be accessed from the Fall 2007 NDSA Administration Technical Manual, <http://www.dpi.state.nd.us/testing/assess/07final.pdf>.

**Table 1.4.6(d): Percent of Student Achievement by Achievement Level
2004-2007, Reading**

Grade	Year	N Count	Performance Levels				
			Novice	Partially Proficient	Proficient	Advanced	Proficient & Above
3	2004	7481	2.69	19.84	57.06	20.41	77.47
	2005	7345	1.92	17.05	55.98	25.05	81.03
	2006	7331	2.11	17.31	62.54	18.03	80.57
	2007	7136	2.90	21.50	56.92	18.68	75.60
4	2004	7485	6.31	18.12	56.86	18.72	75.58
	2005	7455	5.39	17.32	56.22	21.07	77.29
	2006	7309	4.76	14.86	58.65	21.73	80.38
	2007	7243	7.14	17.41	51.90	23.55	75.45
5	2004	7751	6.32	18.68	57.15	17.84	74.99
	2005	7489	6.92	20.63	53.40	19.05	72.45
	2006	7419	4.53	21.22	62.65	11.61	74.26
	2007	7298	9.21	23.66	56.06	11.07	67.13
6	2004	8160	6.90	18.60	54.58	19.91	74.49
	2005	7700	8.36	20.42	50.79	20.43	71.22
	2006	7423	5.48	17.07	54.55	22.90	77.45
	2007	7334	8.19	19.58	55.81	16.42	72.23
7	2004	8270	7.57	18.63	56.93	16.87	73.80
	2005	8134	7.23	17.11	52.14	23.52	75.66
	2006	7734	5.29	14.46	58.37	21.89	80.26
	2007	7447	6.46	16.26	56.86	20.42	77.28
8	2004	8393	8.98	19.08	55.73	16.22	71.95
	2005	8230	8.42	19.26	60.32	12.00	72.32
	2006	8097	7.57	17.27	58.02	17.14	75.16
	2007	7653	8.08	18.46	56.49	16.97	73.46
11	2004	8038	8.67	20.73	52.41	18.19	70.60
	2005	7868	7.41	19.20	59.29	14.10	73.39
	2006	7712	9.41	18.39	53.33	18.87	72.20
	2007	7596	12.81	22.22	51.65	13.32	64.97

**Table 1.4.6(e): Percent of Student Achievement by Achievement Level
2004-2007, Mathematics**

Grade	Year	N Count	Performance Levels				
			Novice	Partially Proficient	Proficient	Advanced	Proficient & Above
3	2004	7477	4.34	12.77	64.14	19.05	83.19
	2005	7347	3.70	10.17	64.15	21.98	86.13
	2006	7349	4.20	10.21	63.44	22.15	85.59
	2007	7188	3.44	10.28	69.32	16.96	86.28
4	2004	7498	5.04	15.46	60.60	18.90	79.50
	2005	7441	4.64	15.75	57.87	21.74	79.61
	2006	7323	5.82	14.82	57.04	22.33	79.37
	2007	7290	5.69	17.05	60.32	16.94	77.26
5	2004	7751	6.45	19.29	54.53	19.73	74.26
	2005	7493	5.30	16.16	59.18	19.36	78.54
	2006	7414	4.07	15.34	52.78	27.81	80.59
	2007	7361	6.55	14.39	56.26	22.81	79.07
6	2004	8176	8.57	21.47	49.94	20.02	69.96
	2005	7715	6.49	18.17	48.89	26.44	75.33
	2006	7425	4.18	13.98	52.24	29.60	81.84
	2007	7376	5.29	14.59	50.94	29.19	80.13
7	2004	8273	9.16	24.31	48.13	18.40	66.53
	2005	8144	7.53	24.31	47.84	20.32	68.16
	2006	7740	5.28	18.04	48.76	27.92	76.68
	2007	7488	5.37	17.20	45.78	31.65	77.43
8	2004	8374	9.60	25.36	46.87	18.16	65.03
	2005	8210	8.75	24.28	47.93	19.05	66.98
	2006	8080	9.37	24.59	43.30	22.74	66.04
	2007	7639	8.76	21.72	42.95	26.57	69.52
11	2004	8025	13.48	31.58	35.19	19.75	54.94
	2005	7859	11.97	31.28	36.19	20.56	56.75
	2006	7673	10.99	32.59	36.53	19.89	56.42
	2007	7577	13.71	30.90	38.99	16.40	55.39

Table 2.3(b): Various student examples of growth-to-standard determinations of sufficient student growth

Grade level in Fall 2007	Grade level in Fall 2008	Performance Level in Fall 2008	Would growth be calculated for this student?	A. Proficiency Cut-Score in Grade 7	B. Scale Score in Fall 2008	C. Difference between A. and B.	D. Total Scale Points Student Must Gain Each Year	E. Scale Score in Fall 2007	F. Actual Gain from Fall 2007 to Fall 2008	Did Student Make Sufficient Growth (is F equal to or greater than D?)	Is student considered proficient in growth-to-standard model?
3	4	Novice	Yes	670	580	90	30.0	529	51	Yes	Yes
3	4	Novice	Yes	670	585	85	28.3	562	23	No	No
3	4	Partially Proficient	Yes	670	608	62	20.7	585	23	Yes	Yes
3	4	Partially Proficient	Yes	670	602	68	22.7	646	-44	No	No
3	4	Proficient	Yes	670	645	25	8.3	617	28	Yes	Yes
3	4	Proficient	Yes	670	622	48	16.0	612	10	No	Yes
3	4	Advanced	Yes	670	657	13	4.3	637	20	Yes	Yes
3	4	Advanced	Yes	670	665	5	1.7	671	-6	No	Yes
4	4	Proficient	No - student was retained	670	616	54	18.0	N/A	N/A	N/A	Yes
N/A	4	Novice	No - student does not have a test record from fall 2008	670	578	92	30.7	N/A	N/A	N/A	No

Table 2.3(c): Prototypic report to schools presenting student growth progression data generated from cross-year growth analysis, student-specific report

Student Growth-to-Proficiency Report

Hettinger 01-013-3599 0K06

Grade 4 Math

Student ID	A. Proficiency Cut-Score in Grade 7	B. Performance Level in Fall 2008	C. Scale Score in Fall 2008	D. # Scale Points Student Must Gain Each Year	E. Scale Score in Fall 2007	F. Actual Gain from Fall 2007 to Fall 2008	Did Student Show Adequate Growth?
1284055847	670	A	665	1.67	646	19	Yes
1293302949	670	N	580	30	529	51	Yes
1489563826	670	PP	595	25			N/A
2986890736	670	P	645	8.33	646	-1	No
2986910212	670	P	645	8.33	617	28	Yes
2986942077	670	P	632	12.7	596	36	Yes
2987006709	670	P	616	18	603	13	No
2987026868	670	A	665	1.67	657	8	Yes
2987042899	670	PP	605	21.7	565	40	Yes
2987180577	670	P	613	19	608	5	No
2987238063	670	P	619	17	637	-18	No
2987300210	670	P	651	6.33	603	48	Yes
2988506970	670	P	636	11.3	696	-60	No
2988572456	670	P	619	17	623	-4	No
2988962684	670	A	657	4.33	623	34	Yes

Table 2.3(d): Number and percent of students tested in fall 2007 whose scale score increased by at least the targeted amount, thereby making sufficient growth*

READING	Number of students tested in fall 2007 with a growth score	Percentage of fall 2007 students who made sufficient growth
Grade 4	6427	67.78%
Grade 5	6472	49.26%
Grade 6	6484	60.90%
Grade 7	6580	66.26%
Grade 8	6898	36.85%

MATH	Number of students tested in fall 2007 with a growth score	Percentage of fall 2007 students who made sufficient growth
Grade 4	6469	59.78%
Grade 5	6528	58.15%
Grade 6	6536	63.66%
Grade 7	6632	52.83%
Grade 8	6919	48.27%

*Includes all students, both above and below proficient, who indicated sufficient growth and were on target to enter into or remain proficient within three years.

Table 2.3(e): Number and percent of students tested in fall 2007 whose scale score increased by at least the targeted amount, thereby making sufficient growth, by achievement level*

READING	Number of students tested in fall 2007 with a growth score	Number of fall 2007 students who made sufficient growth
Novice	2344	8.57%
Partially Proficient	6181	24.96%
Proficient	18458	61.76%
Advanced	5888	89.33%

MATH	Number of students tested in fall 2007 with a growth score	Percentage of fall 2007 students who made sufficient growth
Novice	1861	10.64%
Partially Proficient	5457	25.20%
Proficient	17109	57.52%
Advanced	8657	83.79%

*Includes all students, both above and below proficient, who indicated sufficient growth and were on target to enter into or remain proficient within three years.

Table 2.3(f): Number and percent of students tested in fall 2007 whose scale score increased by at least the targeted amount, thereby making sufficient growth, by subgroup*

READING	Number of students tested in fall 2007 with a growth score	Percentage of fall 2007 students who made sufficient growth
Asian American	267	68.16%
Black	551	48.64%
Hispanic	502	49.80%
Native American	2662	48.95%
White	28765	56.80%
Other Ethnicity	114	52.63%
Females	16027	56.77%
Males	16824	55.27%
Non-IEP	28662	58.91%
IEP	4199	36.10%
Non-LEP	31790	56.33%
LEP	1071	46.13%
Non-NSLP	22549	58.07%
NSLP	10312	51.47%
Non-Migrant	32813	56.01%
Migrant	48	50.00%

MATH	Number of students tested in fall 2007 with a growth score	Percentage of fall 2007 students who made sufficient growth
Asian American	268	62.69%
Black	572	52.62%
Hispanic	509	47.54%
Native American	2707	49.61%
White	28918	57.29%
Other Ethnicity	110	42.73%
Females	16077	57.08%
Males	16997	51.39%
Non-IEP	28657	58.56%
IEP	4427	13.38%
Non-LEP	31983	56.69%
LEP	1101	48.86%
Non-NSLP	22642	58.05%
NSLP	10442	52.90%
Non-Migrant	33038	56.45%
Migrant	46	39.13%

***Includes all students, both above and below proficient, who indicated sufficient growth and were on target to enter into or remain proficient within three years.**

Table 2.5(a): Comparing the number of schools that did not meet a certain achievement indicator based on either the current status model or the proposed growth-to-standard model

Indicators	Primary Indicators: Achievement			
	Reading		Math	
	Current Status Model	Proposed Growth Model	Current Status Model	Proposed Growth Model
Overall	90	70	22	22
White	41	33	4	4
Native American	38	34	24	24
Black	7	7	5	5
Asian	1	1	0	0
Hispanic	5	4	0	0
Other	0	0	0	0
Low Income	105	85	29	29
Disabilities (IEP)	102	88	32	31
Limited English Proficiency (LEP)	22	19	16	13
Total	411	341	132	128

Table 2.5(b): Number of schools identified for not making AYP comparing the current status model and the proposed growth-to-standard model*

	Current Status Model	Proposed Growth Model
Number of schools that did not meet AYP:	169	133

* The 169 schools identified under the current status model were identified for 411 combined indicators in reading and 132 combined indicators in mathematics. The 133 schools identified under the proposed growth-to-standard model were identified for 341 combined indicators in reading and 128 combined indicators in mathematics.

Addendum: Data Files

The tables and figures presented throughout this proposal’s narrative and appendices have been generated from a variety of data files. In the interest of meeting the specified space limitation within this submission, the data files have not been included. The NDDPI commits itself to provide to the ED any and all files that the ED requires to make a full and complete determination of the proposal’s value. The NDDPI will accommodate any such ED requests.