

CHARTING STARS

Pathways to Success

Written by

Jody C. Isernhagen, Ed.D.

Principal Investigator, STARS Comprehensive Evaluation

Associate Professor of Educational Administration

132 Teachers College Hall

University of Nebraska-Lincoln

Lincoln, NE 68588-0360

402.472.1088

jisernhagen3@unl.edu

Shirley J. Mills, Ph.D.

Secondary Investigator

Department of Educational Leadership

Educational Complex 1.606

University of Texas-Pan American

1201 West University Drive

Edinburg, TX 78539-2999

Study Contracted by the Nebraska Department of Education

Also supported by

College of Education and Human Sciences,

University of Nebraska-Lincoln

Submitted to Douglas Christensen, Commissioner of Education

October 2007

<p>Table of Contents Charting STARS: Pathways to Success</p>

Section 1: Introduction	11
Section 2: Executive Summary	15
Section 3: Research Studies	
Study I: Nebraska-led Portfolio Peer Review Process	51
Theme 1: <i>Alignment</i>	53
Theme 2: <i>Sufficiency</i>	56
Theme 3: <i>Clarity</i>	58
Theme 4: <i>Appropriateness</i>	61
Theme 5: <i>Scoring Procedures</i>	63
Theme 6: <i>Summarizing the Review Process</i>	65
Theme 7: <i>Support for the Review Process</i>	68
Theme 8: <i>Challenges</i>	70
Theme 9: <i>New Learnings</i>	73
Study II: Leader and Teacher Perceptions of STARS	75
Theme 1: <i>District Support</i>	77
Theme 2: <i>Assessment Literacy</i>	80
Theme 3: <i>Data</i>	83
Theme 4: <i>Leadership</i>	85
Theme 5: <i>Instructional Impact</i>	88
Theme 6: <i>External Support</i>	92
Theme 7: <i>Special Populations</i>	94
Theme 8: <i>Communication</i>	98
Theme 9: <i>The STARS Process</i>	99
Study III: 2001-2006 Reading, Math and Science Achievement	101
Study IV: 2002-2006 Writing Achievement	109
Study V: 2001-2006 Achievement for Special Populations	115

Study VI: Use of Data to Inform Decisions by Elementary Building Principals.....	123
Study VII: The Role of Technology in Supporting STARS.....	133
Study VIII: Statewide Writing Assessment Scoring: Effective Professional Development for the Classroom Teacher	157
Study IX: The Effect of an Intensive Writing Instruction Assessment Preparation Program on Student Writing Proficiency	179

Section 5: Appendices:

A. IRB Approval Letter.....	187
B. Researchers for the Project.....	191
C. Validation Framework	195
D. 2006-07 Study I Nebraska-Led Portfolio Survey	203
E. 2006-07 Study I Nebraska-Led Portfolio Interview Protocol	207
F. 2006-07 Study II Leader Research Survey	215
G. 2006-07 Study II Teacher Research Survey	219
H. 2006-07 Study II Leader Interview Protocol.....	223
I. 2006-07 Study II Teacher Interview Protocol	231
J. 2006-07 Study VI Use of Data to Inform Decisions by Elementary Building Principals Survey	239
K. 2006-07 Study VII Role of Technology in Supporting STARS Survey.....	251
L. 2006-07 Study VIII 2007 Statewide Writing Assessment Scoring Rater Evaluation	259

TABLES & FIGURES

Tables

1	Percent Proficient or Higher (Change) on Criterion-Referenced Tests in Reading	104
2	Percent Proficient or Higher (Change) on Norm-Referenced Tests in Reading	104
3	Percent Proficient or Higher (Change) on Criterion-Referenced Tests in Math	105
4	Percent Proficient or Higher (Change) on Norm-Referenced Tests in Math	105
5	Reading and Math District Average Portfolio Ratings (Gain/Loss) 2001-2006	106
6	Math District Average Portfolio Ratings (Gain/Loss) 2001-2006	106
7	Statewide Writing Assessment 2002-2007: Mean District Percent of Student Scores at the Proficient Level or Higher.....	112
8	District Average Reading Criterion Referenced Assessments English Language Learners (ELL) 2001-2006	117
9	District Average Math Criterion Referenced Assessments English Language Learners (ELL) 2002-2007	118
10	District Average Writing Criterion Referenced Assessments English Language Learners (ELL) 2002-2006	119
11	District Average Reading Criterion Referenced Assessments Special Education (SPED) Students 2001-2006.....	119
12	District Average Math Criterion Referenced Assessments Special Education (SPED) Students 2002-2006.....	120
13	District Average Writing Criterion Referenced Assessments Special Education (SPED) Students 2002-2006	121
14	Degree of Entire School Faculty Involvement in Analysis of Student Achievement Data.....	126
15	Degree to Which Principals Disaggregate Student Data to Identify Goals.....	127
16	Degree to Which Teachers, Parents, Community Members are Engaged in Data-Driven Decision Making	128
17	Degree Data was Used to Determine Professional Development	128
18	Degree Data was Used to Monitor School Improvement Goals	129
19	Frequency and Percentage of Respondents' School Enrollment.....	135
20	Frequency and Percentage of Respondents' Job Titles	136
21	Frequency and Percentage of Respondents' Assessment Responsibilities	137
22	Frequency and Percentage of Respondents' Type of Participation in STARS	137
23	Frequency and Percent of ESU Roles in the Assessment Process	138
24	Ability to Invest In Assessment Technology	139
25	Frequency and Percent of ARMS Grants Investment Areas	140
26	Software Use in Developing and Documenting Assessment Procedures	141
27	Software Use in Administering Assessments to Students.....	142
28	Software Use in Aggregation, Management and Storage of Assessment Data	143
29	Software Use in Interpretation and Reporting of Assessment Data	143
30	Technology-based Assessment Systems are a Good Match for the Nebraska Assessment Model	144
31	Self-Assessment of the Quality of Technology Currently Used with the District in Various Assessment Areas	145
32	Contributions of Technology	147
33	Technology Significantly Reduces the Amount of Time Involved in STARS	147
34	Use of Technology Significantly Increased the Complexity of the STARS Process.....	148
35	Use of Technology Increases the Value of the Assessment Data Collected in the STARS Process.....	149
36	Technology Assists Teachers in Understanding the Data Collection Process	149

37	Technology Assists Teachers in Utilizing the Data Collected in the STARS Process for Instructional Decision-Making.....	150
38	Potential Barriers Related To Using Technology to Support Assessment	152
39	Quality of Currently Available Technology Used to Support the Collection, Management, and Reporting of Assessment Data	153
40	Desirability of Characteristics in a Technology-based Data Collection, Management, and Reporting System.....	154
41	Likely Impact of Potential Areas of State Support or Resources.....	155
42	Raters' Years of Teaching Experience by Grade Level and Percentage of Raters Currently Teaching.....	162
43	Raters' Previous Statewide Writing Assessment Scoring Experience by Grade Level and Year	163
44	Being a rater has increased my understanding of the qualities of good writing.....	164
45	Being a rater has increased my confidence in making accurate ratings of the quality of student writing.....	164
46	Being a rater has contributed to my ability to help my students become better writers.....	165
47	Being a rater has affirmed my current classroom writing instruction practices.....	166
48	Being a rater has given me new ideas I can use in the teaching of writing in my classroom	166
49	Being a rater has helped me develop more accurate grade-related expectations for student writing	166
50	Being a rater has helped me better understand how writing relates to thinking and reasoning	168
51	Being a rater has given me greater appreciation of writing's uses in content area learning	168
52	Being a rater has increased my general confidence as a teacher of writing	169
53	Being a rater has built my confidence in giving my students helpful feedback.....	169
54	Being a rater has given me assurance in talking to parents and others about writing and it's importance.....	170
55	Raters' Level of Agreement or Disagreement Related to the Appropriateness of the Training	171
56	Raters' Level of Agreement or Disagreement on the Effectiveness of Re-training Procedures	171
57	Raters' Levels of Agreement or Disagreement Related to the Clarity of Training/Scoring Materials.....	172
58	Raters' Levels of Agreement or Disagreement Related to the Expert Knowledge Demonstrated by Scoring Room Leaders and Table Leaders	173
59	Raters' Levels of Agreement or Disagreement Related to the Skill of Scoring Room Leaders and Table Leaders	174

Figures

1	Nebraska-led Portfolio Peer Review Process Survey Category Average Scores 2006-2007.....	17/53
2	Leader and Teacher Perceptions of STARS Leader Average Lows, Means, and Highs 2006-2007	23
3	Leader and Teacher Perceptions of STARS Teacher Average Lows, Means, and Highs 2006-2007	23
4	Leader and Teacher Perceptions of STARS Average Ratings of All Leaders and Teachers 2006-2007	77
5	STARS Technology Survey Results Distribution of Respondents by School Enrollment Categories 2006-2007	136
6	STARS Technology Survey Results-Distribution of Respondents by Job Title 2006-2007.....	136
7	STARS Technology Survey Results Respondents Reporting Various Assessment Responsibilities 2006-2007	137
8	STARS Technology Survey Results Frequency Distribution for District Methods of Participation in STARS 2006-2007.....	138

9	STARS Technology Survey Results Respondents' Selecting Various ESU Roles in the Assessment Process 2006-2007	139
10	STARS Technology Survey Results Distribution of Responses to Ability to Invest in Assessment Technology 2006-2007	139
11	STARS Technology Survey Results Investment Areas that are Targeted by ARMS Grants 2006-2007	140
12	STARS Technology Survey Results Extent to Which Software Was Used in Developing and Documenting Assessment Procedures 2006-2007	141
13	STARS Technology Survey Results Extent to Which Software Was Used in Administering Assessments to Students 2006-2007	142
14	STARS Technology Survey Results Extent to Which Software Was Used in Aggregation, Management, and Storage of Assessment Data 2006-2007	143
15	STARS Technology Survey Results Extent to Which Software was Used in Interpretation and Report Assessment Data 2006-2007	144
16	STARS Technology Survey Results Percentage of Respondents in Agreement That Technology-based Systems Are a Good Match for the Nebraska Assessment Model 2006-2007	145
17	STARS Technology Survey Self-Assessment of the Quality of Technology Currently Used with the District in Various Assessment Areas 2006-2007	146
18	STARS Technology Survey Results Contributions of Technology to Assessment 2006-2007	147
19	STARS Technology Survey Results Technology Significantly Reduces the Amount of Time Involved in STARS 2006-2007	148
20	STARS Technology Survey Results Use of Technology Significantly Increased the Complexity of the STARS Process 2006-2007	148
21	STARS Technology Survey Results Use of Technology Increases the Value of the Assessment Data Collected in the STARS Process 2006-2007	149
22	STARS Technology Survey Results Technology Assists Teachers in Understanding the Data Collection Process 2006-2007.	150
23	STARS Technology Survey Results Technology Assists Teachers in Utilizing the Data Collected in the STARS Process for Instructional Decision-making 2006-2007	151
24	STARS Technology Survey Results Potential Barriers Related to Using Technology to Support Assessment 2006-2007	152
25	STARS Technology Survey Results Quality of Currently Available Technology Used to Support the Collection, Management, and Reporting of Assessment Data 2006-2007	153
26	STARS Technology Survey Results Desirability of Characteristics in a Technology-based Data Collection, Management, and Reporting System 2006-2007	154
27	STARS Technology Survey Results Likely Impact of Potential Areas of State Support or Resources 2006-2007	155
28	2007 SWA Scoring Rater Information	163
29	2007 SWA Scoring Benefits to Teachers	167
30	2007 SWA Scoring Benefits to Teachers About Writing	168
31	2007 SWA Scoring Benefits to Teacher Confidence	170
32	2007 SWA Effectiveness of Rater Training	172
33	2007 SWA Scoring Training Effectiveness	174

In Memory of our Colleague

Dr. Leon Dappen

1943 - 2007

*Think where man's glory most begins and ends,
And say my glory was I had such friends.*

William Butler Yeats

The STARS Comprehensive Evaluation Project research team wishes to acknowledge Dr. Leon Dappen's wonderful contribution to STARS research and to the lives of each of those who worked with him. He was a kind and gentle man that loved education and contributed to many with his constant smile, support, and knowledge. Leon will be greatly missed.

At the time of his death, Leon was an Associate Professor in Educational Administration for the University of Nebraska at Omaha.. He received his doctoral degree from the University of Nebraska-Lincoln and had served as a teacher, counselor, psychologist, special education administrator, and assistant superintendent in pre-K through 12 schools. Over the past four years, Leon served as a secondary investigator for the STARS Comprehensive Evaluation Project. During the eight years preceding his untimely death, Leon collaborated on research concerning STARS and school issues. He co-authored many articles with Jody Isernhagen.

CHARTING STARS: Pathways to Success

Section 1: Introduction



The sixth annual report of the STARS Comprehensive Evaluation Project (CEP) is an independent evaluation of Nebraska's School-based Teacher-led Assessment and Reporting System (STARS). This study was approved by the Institutional Review Board (IRB) at the University of Nebraska-Lincoln (UNL) characterized by the highest level of integrity, with respect and equitable treatment for all persons involved in the study in order to maintain confidentiality and protect the privacy of participants in the study (Appendix A). The CEP was originally contracted between the Nebraska Department of Education (NDE) and the University of Nebraska-Lincoln, College of Education and Human Sciences (CEHS) in 2001. The CEP was supported jointly by the Nebraska Department of Education (NDE) and the College of Education and Human Sciences (CEHS). Dr. Jody Isernhagen, Associate Professor, served as the Principal Investigator. Dr. Leon Dappen¹, Associate Professor, and Dr. Shirley Mills, Assistant Professor, served as secondary investigators. Jadi Miller, Principal, Everett Elementary, Lincoln Public Schools, conducted Study VI. Dr. Al Steckelberg, Associate Professor, Teaching, Learning, and Teacher Education, CEHS, and Lan Li, Graduate Assistant, conducted Study VII. Sue Anderson, Director of Professional Development, Educational Service Unit #3, conducted Study VIII, and Edward J. Bennett, Assistant Principal, Omaha Central High School, Omaha Public Schools, conducted the final study (IX). All researchers and members of the research team for the Comprehensive Evaluation Project are listed in Appendix B.

¹ Posthumous

OVERVIEW

Nebraska educators have ventured down new pathways unlike educators in other states in the nation. After building their assessment literacy and data analysis skills, teachers are now honing their instructional skills to better meet the needs of Nebraska students. In 2006-07, a new Nebraska-led Portfolio Peer Review Process was introduced with great success and provided many new opportunities for learning. Throughout this process of evaluation, a growing number

of educators are finding new pathways that answer questions and provide new interventions to enable all students to excel.

SUMMARY OF THE SIXTH-YEAR STUDY

Over the past six years, the Nebraska School-based Teacher-led Assessment and Reporting System (STARS) has required Nebraska's school districts to develop a local assessment system to measure student performance on local standards that equal or exceed the state standards. Each year educators have continued to perfect the implementation of the STARS Process to better meet the needs of Nebraska's students. This process is designed to help educators build new connections between assessment, curriculum, and instruction and is intended to target appropriate interventions for students to ensure that they all meet with success. Nine studies were conducted during the sixth year of the Comprehensive Evaluation Project.

STUDY I: Nebraska-led Portfolio Peer Review Process

The first study (Study I) was an investigation of educator perceptions of the technical quality of their classroom assessments according to the requirements of the six quality criteria (Plake & Impara, 2000):

- The assessments reflect the state/local standards.
- Students have the opportunity to learn.
- The assessments are free of bias and insensitive situations.
- The assessments are at the appropriate level.
- The assessments are reliably scored.
- The assessment mastery levels are appropriately set.

The investigation was conducted in conjunction with the new Nebraska-led Portfolio Peer Review Process, with quantitative data being collected before the external review for each district. Qualitative data was collected after the actual Portfolio Peer Review in selected districts.

The Portfolio Peer Review consisted of statewide teams with two appointed state-trained assessment experts visiting each school district within the state. This visiting team read the previous district portfolio and the recommended changes for improvement from past years prior to visiting the district. The visiting team then went to their assigned districts, reviewed evidence based on the six quality indicators, and provided formative feedback to the district. A summative rating will appear on the State of the Schools Report in 2007.

STUDY II: Leader and Teacher Perceptions of STARS

The second study (Study II) was conducted to identify and explore any similarities and/or differences between leader and teacher perceptions concerning the impact of STARS on district support, assessment literacy, use of data in classroom settings, leadership, instructional impact, and external support with the use of an online survey. Qualitative interview data were also collected to extend the understanding of the survey findings.

STUDY III: 2001-2006 Reading and Math Achievement

The third study was a longitudinal achievement study (Study III) conducted in 2006-07 for the STARS Comprehensive Evaluation Project (CEP). District achievement scores for reading and math were compared on criterion-referenced and norm-referenced measures from 2001 through

2006. District portfolio ratings for reading and math from 2001 through 2006 were also compared.

STUDY IV: 2002-2006 Writing Achievement

The fourth study (Study IV) was a longitudinal study that focused on writing achievement. District writing scores on the Nebraska Statewide Writing Assessment (NSWA) from 2002 through 2006 were compared.

STUDY V: 2001-2006 Achievement for Special Populations

In study five (Study V), district reading, math, and writing scores for English Language Learners (ELL) and Special Education (SPED) students were compared on criterion-referenced measures from 2001-2006.

STUDY VI: Use of Data to Inform Decisions by Elementary Building Principals

In study six (Study VI), Nebraska Elementary Principals completed an online survey to determine their use of data to support decision making.

STUDY VII: Role of Technology in Supporting the Nebraska School-based Teacher-led Assessment System (STARS)

Study seven (Study VII) examined the perceptions of Nebraska educators on the use of technology in the STARS Process using a 2007 online survey.

STUDY VIII: Statewide Writing Assessment Scoring: Effective Professional Development for the Classroom Teacher

In study eight (Study VIII), Nebraska educators responded to a 2007 survey focused on the use of the writing assessment scoring process as professional development.

STUDY IX: The Effect of an Intensive Writing Instruction Assessment Preparation Program on Student Writing Proficiency

In study nine (Study IX), a pretest-posttest two group comparative survey was conducted to determine whether an intensive writing instruction and assessment preparation program (IWIAPP) was effective in helping eighth-grade students improve their performance on the Nebraska Statewide Writing Assessment (NSWA).

VALIDATION: Nebraska's Standards, Assessment, and Accountability System

Pat Roschewski, Director of Statewide Assessment for the NDE, designed a Validation of Nebraska's Standards, Assessment, and Accountability System for the CEP (Appendix C) to support the long-term evaluation of STARS. This validation of Nebraska's standards, assessment, and accountability framework places these studies in proper context to the overall long-term, statewide evaluation plan.

YEAR SIX COMPREHENSIVE EVALUATION FORMAT

This comprehensive report has been designed to serve multiple audiences and provide the most pertinent information available on the implementation of STARS based on the data collected during the 2006-07 school year. All of the studies connected assessment, curriculum, and instruction pathways that support increased student learning.

This report is divided into four sections beginning with an introduction of the total report (Section 1); followed by an executive summary of the complete findings of all studies conducted during the sixth-year study (Section 2); the third section (Section 3) focuses on the nine major studies (Studies I-IX) conducted during the 2006-07 school year; and, finally, the fourth section (Section 4) contains the Appendices.

ACKNOWLEDGMENTS

We offer a special thanks to the many educators in school districts across Nebraska for sharing the many pathways they have traveled to create a better learning environment for students. Their expertise and enthusiasm for tackling the complexities and connections between assessment, curriculum, and instruction helps to create opportunities and ensure learning for all Nebraska students.

The NDE staff: Douglas Christensen, Commissioner of Education; Pat Roschewski, Director of Statewide Assessment; Bob Beecham, Administrator of Education Support Services; John Moon, National Assessment of Educational Progress (NAEP) Coordinator; Jackie Naber, Office Administrator of Statewide Assessment, and Carol Bom, Office Assistant have offered great assistance for the completion of the Comprehensive Evaluation.

To our interviewers, Larry Bornschlegl and Ron Klemke, we couldn't have done this without you. Thanks for traveling across this great state to conduct interviews in school districts. A special thanks to Shirley Mills for her long hours and dedication given to this project. Without her this report surely would not have been completed. Jadi Miller, Al Steckelberg, Lan Li, Sue Anderson, and Edward Bennett have enriched the comprehensive evaluation with their studies and findings.

A special thanks to our Administrative Assistant, Susan Wilson, as she provided great attention to detail and completion of the final product. To Cindy DeRyke, Diane Gronewald, Tracy Helmink and Tammie Herrington, thank you for your support throughout the year. A special thanks to Marjorie Kostelnik, Dean of the College of Education and Human Sciences; L. James Walters, Associate Dean; and Larry Dlugosh, Chair of the Department of Educational Administration, for their continued support for the Comprehensive Evaluation Project.

CHARTING STARS: PATHWAYS TO EXCELLENCE

Section 2: Executive Summary



INTRODUCTION

Over the past six years Nebraska educators have walked a path unknown to many educators across the nation. Due to the implementation of Nebraska's unique assessment system, educators have been compelled to stretch and grow in their understanding of assessment, curriculum, and instruction. This growth has led to a common language about assessment related to the six quality criteria developed by the Buros Center during the initial stages of assessment implementation (Plake & Impara, 2000).

Building on the statewide district portfolio process used to measure the success of the assessment process, school districts have showcased their assessments, curriculum, and instruction to teams of experts in Nebraska and across the nation. These teams visited all school districts during the 2006-07 school year. The path to improved student achievement has not been without bumps in the road, but Nebraska educators have committed themselves to an improved assessment process and learning for all students.

SIXTH YEAR RESEARCH STUDIES

There were nine major studies conducted during the sixth year of the STARS Comprehensive Evaluation Project (CEP). These nine studies are summarized in this section of the report and presented as separate studies in Section III of the full report.

STUDY I: Nebraska-led Portfolio Peer Review Process

Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln
Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American

Introduction

The first study (Study I) was an investigation of educator perceptions of the technical quality of their district assessments according to the requirements of the six quality assessment criteria (Plake & Impara, 2000):

- The assessments reflect the state/local standards.
- Students have the opportunity to learn.
- The assessments are free of bias and insensitive situations.
- The assessments are at the appropriate level.
- The assessments are reliably scored.
- The assessment mastery levels are appropriately set.

The portfolio review consisted of teams of two state-trained assessment experts that visited each school district within the state. The assessment reviewers read the previous district portfolio and the recommended changes for improvement from past years. The visiting team then went to their assigned district, reviewed the evidence of assessment quality based on the six assessment quality indicators, and provided formative feedback to the district. Two assessment outside experts located in each of the regional areas assisted the peer reviewer teams when there were questions.

Methodology

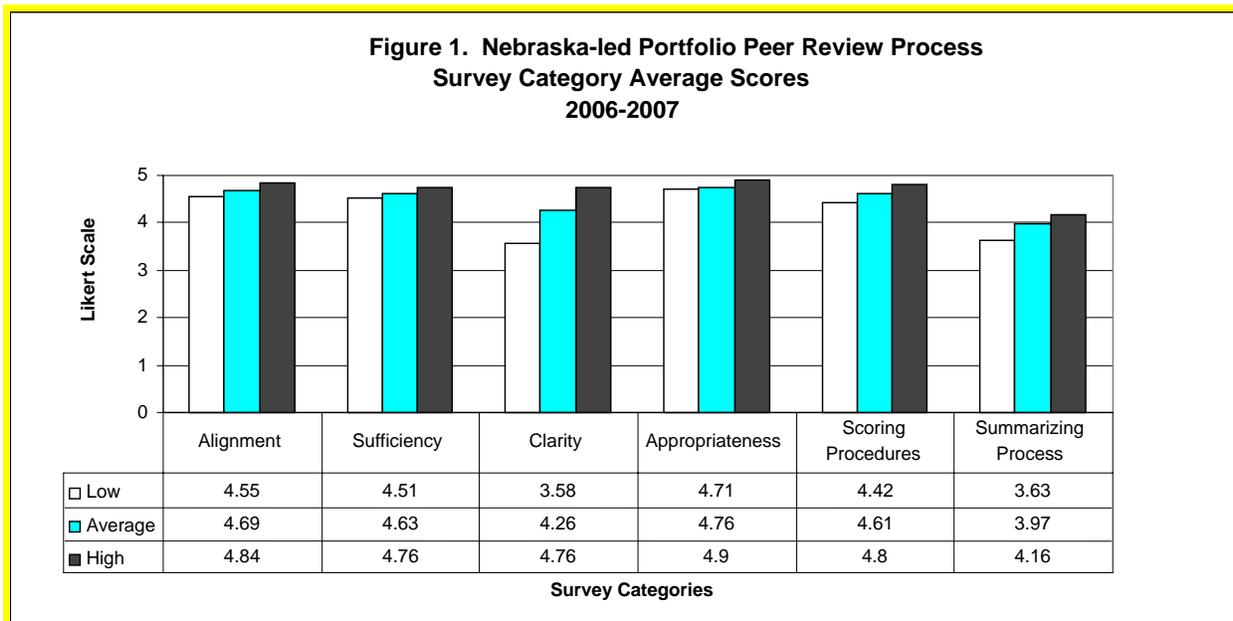
This six year mixed-methods research study focused on the Nebraska-led Assessment Portfolio Peer Review Process. The District Internal Portfolio Review Team members were surveyed prior to the Portfolio External Review. For the purpose of this research, Nebraska school districts were divided into two classifications, non-rural and rural, based on population characteristics unique to Nebraska. Non-rural districts were defined as metro-area districts in large and mid-sized cities, large towns, and the urban fringe. Each district was asked to distribute up to five surveys, as appropriate, to internal review team members. One hundred seventy-nine (179) districts (70% of 254 districts) returned surveys for this study. Of those 179 districts, 23 districts were disallowed because they did not return surveys prior to their actual Nebraska-led Portfolio Peer Review. Of the 156 districts responding within the prescribed timeframe, 5% were non-rural and 95% were rural. Participants responded to the 49-item survey using a five-point Likert scale for each item, with “1” representing “none of the time,” “2” “very little of the time,” “3” “some of the time,” “4” “most of the time,” “5” “all of the time.” The survey was structured to explore six themes: Alignment, Sufficiency, Clarity, Appropriateness, Scoring Procedures, and Summarizing the Review Process (Appendix D).

In the second phase of this study, open-ended interviews were conducted in two districts from each of four geographical areas identified by the Nebraska Department of Education (NDE) for training of reviewers. Detailed views were collected about the Nebraska-led Portfolio Peer Review Process in the sample districts. The interview protocol for educators (Appendix E) was utilized to gather qualitative data. The six themes based on the survey were also evident in the interviews. They were: Alignment, Sufficiency, Clarity, Appropriateness, Scoring Procedures,

and Summarizing the Review Process. Three additional themes emerged from the interviews: Support for the Peer Review Process, Challenges and New Learnings.

Findings

Educators’ perceptions of the Nebraska-led Portfolio Review Process were explored using a survey based on six categories. This survey was administered to members of the school based review team prior to the visit to school districts by experts. These items were rated by participants on a “1” to “5” Likert scale with “5” being the highest. Noted in the graph are the mean and the lowest and highest average scores per category. The highest average score was in “Appropriateness” while “Summarizing the Process” was the lowest average score. The lowest score was found in the category, “Clarity” while the highest was found in “Appropriateness”.



Six themes with focus on the six quality criteria were identified within the interviews based on the survey categories. Three additional themes emerged during the interview. These nine themes are addressed in this summary of Study I.

Alignment

In the area of Alignment, educators indicated that Alignment was the impetus for change in schools as teachers and administrators worked together to figure out how “curriculum and assessments matched the standards.” A rural superintendent shared, “You walk into any classroom, in any one of these sites, and you’re going to see standards and assessment curriculum books out, laying out, laying open, it’s not just because its spring and we have to assess. They really use them.” The strongest item rated by all respondents within the Alignment category was “districts involved staff in the alignment of assessments to standards (4.84)” with respondents also rating strongly that “districts’ assessment items or tasks reflected a match to the appropriate standards (4.74)” and “there is a documentation process for alignment of assessments to standards (4.74).” The lowest rated item on the survey in this category was “districts had a list of specifications mapping the assessment items to the standards in order to show which items

assessed which standards.” Evidence from the interviews supported this conclusion as only two references to a table of specifications were made.

During the interviews, participants were asked if they had recommendations to make to other school districts regarding Alignment. Recommendations included having an articulated curriculum, having a clear understanding of what skills and knowledge are essential to determine proficiency to a standard, constant monitoring, involvement of all teachers including specialist teachers (special population teachers), and training and support to understand the table of specifications.

Sufficiency

Through the surveys and interviews, it became apparent to the researchers that Sufficiency was one of the most complex criteria for schools and districts to understand. In the Sufficiency category, the strongest rated item on the survey from all respondents indicated that “districts reviewed the assessment items/tasks for sufficiency results (4.76)” and “the items or tasks were distributed across all performance levels (4.71).” The lowest rated area was “assessment items/tasks used a variety of appropriate formats (4.51)” and “items/tasks included higher order thinking skills (4.53).” This low rating was supported by the fact that large non-rural districts sometimes rely on outside agencies to develop assessment items for sufficiency.

During the interviews, participants were asked if they had recommendations to make to other school districts regarding the area of Sufficiency. Recommendations included openness to how students react to assessment items, the importance of questions by performance levels and the need for the issue of sufficiency on locally developed assessments to be addressed by people other than teachers.

Clarity

The surveys and interviews seemed to indicate that schools showed a low level of understanding in the area of Clarity. Interviewers asked for recommendations about clarity, but very little new information was collected as there was limited discussion in this area. The strongest perception from all respondents indicated that “assessment directions for students were clear (4.76)” while the lowest two items were “districts provided parents with reports that gave an appropriate explanation of assessment results (3.58)” and “individual reports are provided to parents each school year (3.74).” Although this area was rated as one of the two lowest items on the survey, elementary teachers rated both items significantly higher than all other teacher categories. Respondents indicated there were many different types of reports being generated for various audiences, the most important were those prepared for parents and students. A great deal of evidence was available to the interviewers about assessment directions used by school districts as well as information provided about how districts report.

During the interviews, participants were asked if they had recommendations to make to other school districts regarding Clarity. Recommendations included taking a look at how the process can be simplified for large populations or large school districts due to the mountain of paperwork generated when administering assessments and analyzing assessment data. In rural school districts having a site coordinator monitor and administer the tests has been found to produce consistency.

Appropriateness

In the area of Appropriateness, bias review was one of the stronger areas relative to the six quality criteria. Furthermore, through the surveys and interviews conducted, it became apparent that educators consistently monitored for bias in their assessments. The strongest perception from all respondents indicated assessments were screened for fairness, bias, and sensitivity (4.90). The lowest items on a one to five point Likert scale were “assessments demonstrate an increase of expectation from one grade level to the next (4.71),” “assessments are appropriate for the assessed grade level (4.72)” and “assessments indicated our expectations for the students (4.72).”

During the interviews, participants were asked if they had recommendations to make to other school districts regarding Appropriateness. Recommendations included taking advantage of every allowable accommodation - knowing what those things are and to use them in the best way that you can to help your kids. Additionally it was recommended that school districts need to have a clear understanding of the Nebraska Department of Education guidelines regarding which assessments you can give accommodations on and those you can't.

Scoring Procedures

In the area of Scoring Procedures, the strongest perception from all respondents indicated that “participation rates were documented (4.80)” and “assessments have established scoring guidelines and directions (4.77).” The lowest item rated within Scoring Procedures was “districts provided training for those administering assessments (4.42).” Teachers and leaders with more years of experience rated, “our subjectively scored assessments have inter-rater reliability and decision consistency methods that are within acceptable ranges,” higher than teachers and leaders with less experience.

During the interviews, participants were asked if they had recommendations to make to other school districts regarding Scoring Procedures Discussions. Recommendations included assuring tests constructs appropriately measure learning as well as maintaining high levels of reliability. A few leaders noted the challenge that as you strive to meet quality criterion five and as student performance increases, reliability calculations fall.

Summarizing the Review Process

During the interview process, it seemed apparent to the researchers that overall the participants felt that the Department of Education had shown a tremendous amount of faith in the integrity of the process. Nebraska educators across the state indicated confidence in the STARS process. The strongest perception from all respondents indicated they had “the necessary information to prepare the district assessment portfolio (4.16)” and the survey respondents “had adequate help in preparing the assessment portfolio (4.14).” The lowest item was, “I was provided compensation to prepare the district assessment portfolio when completed outside of the regular school day (3.63).” Teachers with more years of portfolio experience rated the item, “I feel prepared to present my district portfolio to my peer reviewers,” and “I have the necessary information to prepare the district assessment portfolio” higher than did those with less years of experience.

During the interviews, participants were asked if they had recommendations to make to other school districts regarding Summarizing the Review Process. As previously noted, overall the participants felt that the Department of Education had shown a tremendous amount of faith in the integrity of the process but did offer recommendations. These recommendations included: peer reviews every three years rather than every year as assessments don't change enough to warrant a review every year, develop a system that allows for portfolios to get to the reviewers a day or two ahead of time, possibly combine activities or incorporate the process through an already existing opportunity such as North Central visitation, offer more feedback on the assessments from the review process and perhaps offer a student component within the process.

Support for the Peer Review Process

During the interviews the theme, Support for the Peer Review Process, emerged. While there were challenges that faced the new process of reviewing the portfolios, educators across the state indicated strong support for the new Peer Review Process. Many educators experienced anxiety about their preparation of their staff and students but indicated confidence in the process itself. Leaders and teachers alike were positive in their comments. Many teachers and leaders were very positive about the reviewers themselves and their ability to lead the review process at the building and district level. It became evident to the researchers that many team members benefited from the discussions held during the review process. Many reviewers who conducted the reviews commented about how much *they* personally learned. Interviews revealed that both reviewers and educators believed that the Nebraska Department of Education worked diligently to support districts in preparation for the new process.

Challenges

The Nebraska-led Portfolio Peer Review Process was completed in one year with all districts presenting their reading assessment portfolios to review teams. As with any new process, there were challenges that became apparent and thus surfaced as a new theme. The challenge for Nebraska will be to create a large enough group of evaluators to continue reviewing every school district in the state. Many of Nebraska's rural districts are isolated and difficult to access. Maintaining teams of people who are able to conduct peer reviews with expertise is and will continue to be a challenge. The suggestions to use experienced reviewers with someone new to the review process would strengthen the results of the Peer Review. Although the overall peer review process was a positive one for most, as previously indicated, educators had different experiences with the evaluation process. It was a challenge to both reviewers and districts to provide a schedule of events for the day which created some problems for educators in districts who were part of the initial review visitations. Another challenge facing educators who participated in the initial reviews was the "pending" ratings that some districts received. Even though districts were informed at the onset by the NDE that those who were first to be reviewed might receive "pending" ratings, which could be rectified before the final submission, several districts still indicated frustration. This presents several challenges for future reviews due to the limited timeframe available for portfolio reviews and the number of districts within the state needing review.

The challenge to work with districts in consortia (write a common portfolio) was a problem for both reviewers and the districts being reviewed. Districts indicated their frustration with receiving a different rating from other members of the consortium that had the same portfolio

This will continue to be a challenge for the NDE. Additionally, some districts felt that their reviewers did not look favorably towards the online system and judgments were impacted by personal perceptions rather than analyzing the data gathered. As with any new process, challenges will be a part of it and educators had suggestions that they believed would strengthen the process. Their contributions offered some ideas that evolved into the final theme, New Learnings.

New Learnings

Conversations with leaders and teachers allowed time for educators to reflect upon the new knowledge and skills acquired since the beginning of STARS. Many educators felt that the Peer Review Process allowed the district to honor their teachers and the amount of work they have done to make the process strong and viable. A rural superintendent said, “It was affirmation of the high value of allowing teachers to own the process...they need to honor the teachers in the classroom.... That was just absolutely ignited again in the room with the reviewers.” While Nebraska school districts have been immersed in the accountability process for at least seven years, they continue to improve and strengthen their process, as evidenced by the new learnings gained and acknowledged within the Nebraska Peer Review interviews. Districts were encouraged to review the process that had been in place for several years and discovered work was needed to keep their assessments current.

Summary

The original six themes were combined with the three additional themes that emerged from the interview process to give a better perspective about the technical quality of district assessments according to the six quality criteria. Alignment is now viewed by many educators as the impetus for change in schools as teachers and administrators interacted with a great deal of discussion regarding how “curriculum and assessments matched the standards.” Sufficiency seemed to be one of the most complex criteria for schools and districts to understand. The surveys and interviews also seemed to indicate that schools showed a low level of understanding in the area of clarity. In the area of Appropriateness, bias review was one of the stronger areas relative to the six quality criteria. Keeping reliability levels high will continue to be a challenge as educators strive to meet quality criterion five. Nebraska educators across the state indicated confidence in the STARS process. There was a strong positive reaction to the Portfolio Review Process as indicated by many educators who experienced anxiety about their preparation of their staff and students but indicated confidence in the process itself. All members benefited by the discussions held during the review process, with many reviewers who conducted the examinations commenting about how much *they* personally learned. Interviews revealed that both reviewers and educators believed that The Nebraska Department of Education worked diligently to prepare and support the districts in the new Portfolio Peer Review Process.

However, as with any new process, there were challenges that emerged. Some of these included: Nebraska's rural districts are isolated and difficult to access; training and accessibility of annual reviewers; scheduling the events for the day enough in advance so districts can adequately plan for the review; the challenge to work with districts in a consortia (write a common portfolio) was a problem for both reviewers and the districts being reviewed as some ratings within the consortia were not consistent; those who were first to be reviewed and received "pending" ratings, even though these could be rectified before the final submission, indicated frustration;

and some districts felt that their reviewers did not look favorably upon the online system. Overall, the process allowed districts to re-evaluate the work that had been done and the work that needs to be accomplished to reinforce student learning.

STUDY II: Leader and Teacher Perceptions of STARS

Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln

Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American

Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹

¹ Posthumous

Introduction

The second study (Study II) was conducted to determine if there were any similarities and/or differences between leader and teacher perceptions concerning the impact of STARS on district support, assessment literacy, use of data in classroom settings, leadership, instructional impact, and external support. Surveys were administered for each population, teachers and leaders. Qualitative interview data were also collected to extend the understanding of the survey findings.

Methodology

For the purpose of the 2006-07 Leader and Teacher Perceptions of STARS study, Nebraska school districts were divided into two classifications, non-rural and rural, based on population characteristics unique to Nebraska. Non-rural districts were defined as metro-area districts in large and mid-sized cities, large towns, and the urban fringe. All other districts were classified as rural.

District leaders were asked to complete an online survey (Appendix F) made available to the membership of the Nebraska Council of School Administrators, an organization representing over 98% of the leaders in the state. Emails inviting survey participation were sent to 1220 leaders representing 250 districts of the 254 districts in the state. Of the 250 districts receiving the invitation to participate, 11 were non-rural districts and 239 were rural districts. Of all the leaders invited to participate, 334 leaders representing 91% of non-rural districts and 72% of rural districts completed surveys.

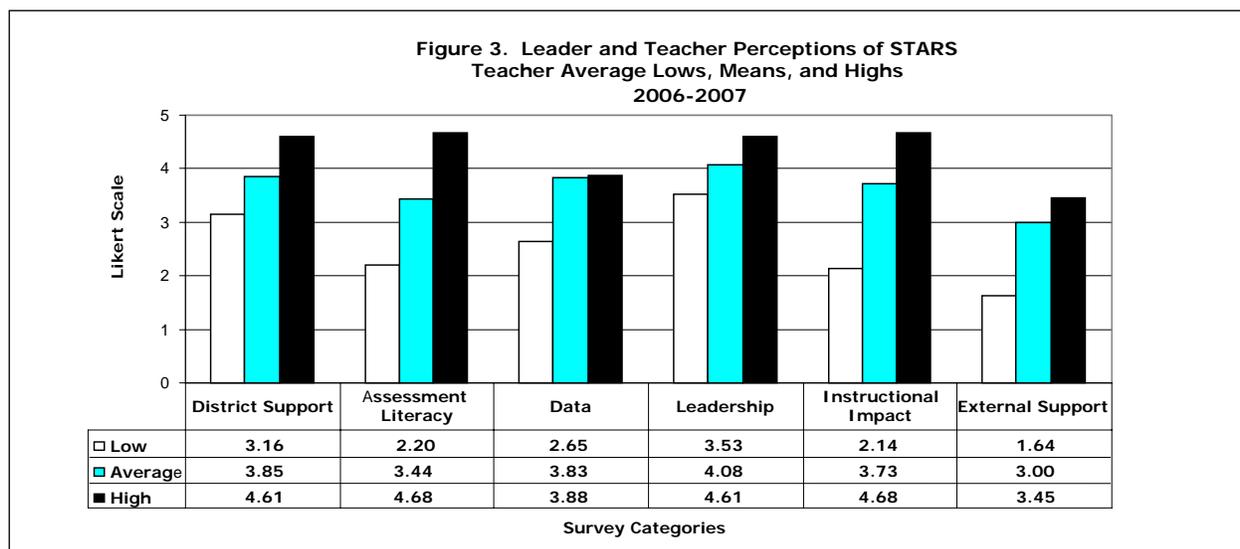
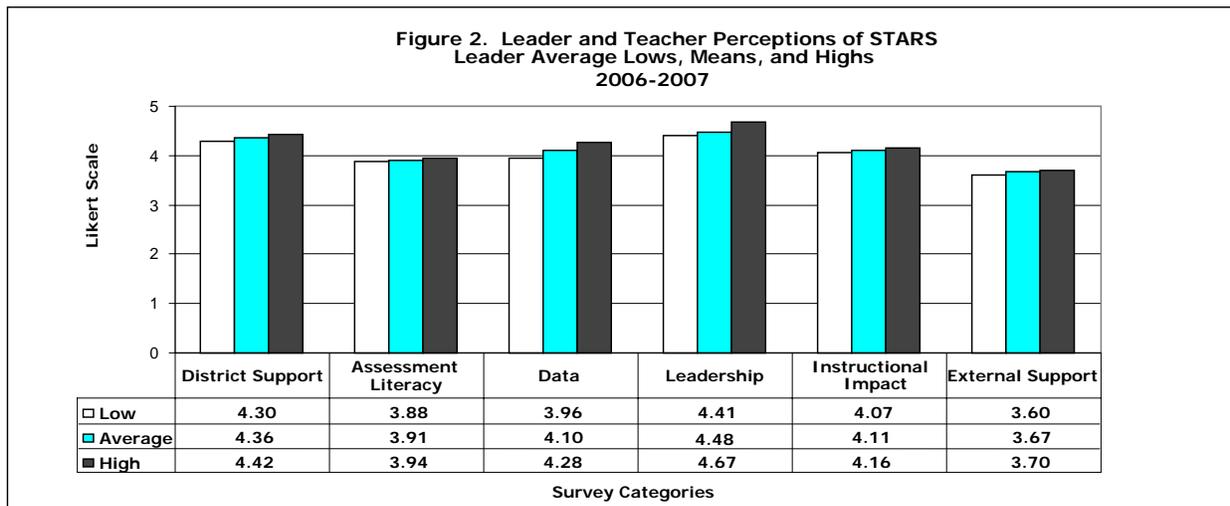
Language arts, math, science, English Language Learner (ELL), and Special Education (SPED) teachers were asked to complete an online survey made available to the membership of the Nebraska School Education Association, an organization representing about 98% of the teachers in Nebraska (Appendix G). Five hundred and sixty-two (562) teachers representing 187 (73.6%) of Nebraska's 254 districts completed the 2006-07 survey. Of the teachers in non-rural districts, nine (82%) of the eleven non-rural districts responded. Teachers in 178 (73%) of the 243 rural districts completed the surveys.

Selection of the sample school districts for the qualitative interviews was based on a stratified, purposeful sample using district class (Classes II-V), geographical areas (East, Central, and West) and free/reduced district lunch (high, middle, low based on the statewide average). Follow-up interviews were conducted with 38 leaders (Appendix H): 7 superintendents, 5 curriculum/assessment coordinators, 10 middle and secondary principals, 8 elementary principals and 8 SPED and/or ESL Directors. Teacher interviews (Appendix I) were conducted with twenty-three (23) classroom teachers: 12 elementary and 11 middle school and/or high school.

As a follow-up to the teacher and leader surveys, qualitative interviews were conducted with 23 teachers and 38 leaders. The following themes were identified from the surveys for both teachers and leaders: district support, assessment literacy, use of data in classroom settings, leadership, instructional impact, and external support. In addition to the six major themes explored by the interviews, three additional themes emerged. The emergent themes were: Special Populations, Communication, and the STARS Process.

Findings

Leader and teacher perceptions of the impact of STARS on district support, assessment literacy, the use of data in classroom settings, leadership, instructional impact and external support were explored using a survey based on these six categories. Noted in the graphs are the mean and the lowest and highest average scores per category for both leaders and teachers. The highest average score for both teachers and leaders was in “Leadership” while “External Support” was the lowest average score for both groups. However, teachers rated both areas lower than did leaders.



Through the surveys and interviews from this study it became evident to the researchers that STARS has truly begun to impact the culture of schools and school districts across the state. A superintendent shared, "...I think it (STARS) has strengthened the culture of the district and strengthened the focus of the teachers of the district. They will say that they are more focused. The ones (teachers) that we have put the most burden on are the ones that are most in favor. It's interesting, I've got to give our high school principal credit when he said, 'It's a paradox, it's like at the end of the day and you're exhausted and you think you can't do one more thing you get on the treadmill and you work really hard for half an hour and when you get off you have more energy than you had before.' He compares that to the STARS system, when he said, 'You give these teachers one more thing when they can't possibly take one more thing but they take ownership for it and they walk out with more energy than they had before.'" The themes explored by the surveys and those that emerged during the interviews are summarized below.

District Support

There were no significant differences noted between leaders and teachers within the survey category District Support. Many non-rural leaders have become more concerned about developing their own district/school based policies that yield a greater focus on monitoring and reporting of student data and, in turn, upon increasing student achievement. Rural teachers continued to emphasize the importance of resources, services, training, and record keeping as key to successful district implementation of STARS.

Assessment Literacy

The total area of assessment literacy for leaders and teachers was not significantly different during the sixth year study. Assessment literacy has become a part of the daily routine of teachers and leaders as they continue to perfect their skills. The biggest challenge identified relative to assessment literacy was the desire to get all teachers on board in large school districts. As this assessment coordinator emphasized, "We're doing our best by doing the 'trainer/trainer' model, trying to get more people understanding formative assessment so that they understand that assessment should not be something that is just tagged on at the end of something, it's something that goes on daily. It's part of that normal curriculum, instruction, assessment cycle." It was noted that, unlike rural districts, it is very difficult for large districts to make a significant change in a district without ample time and notification as illustrated by this assessment coordinator, "It becomes difficult for us when we are asked to change on a dime . . . we don't touch a subject for seven years once it is done. We continue to do training, we continue to modify, but interventions are set based on the needs of the district."

Data

The entire survey category of Data was marginally significant for leaders and was significantly different for teachers. A superintendent summarized the importance of providing leadership when using data, "I've learned a long time ago if you measure it and you watch that data and you pay attention to it, it becomes important to everybody." It is obvious to the researchers that the use of data in schools has increased over time and is impacting instruction on a daily basis. Teachers emphasized the importance of looking at gap data as noted by this elementary teacher, "I think if you see large gaps, if you see areas that perhaps you're weak in, or even areas that you're strong in, I think that dictates, then, what it is you teach. So I think that data is very important."

Leadership

The total category of Leadership was significant for both leaders and teachers. Both administrative and teacher leadership are critical to the success of the STARS process and have emerged as a strong indicator to district and school success. A rural elementary principal shared, "We have staff who are knowledgeable and who are very active in the summer workshops and can help others understand the process more. That's ... a leadership role I would say." A female rural middle school principal shared, "I count on my fourth-grade teachers, eighth-grade teachers and eleventh-grade teachers in those core areas especially reading, writing, speaking, listening and math to know the process and be able to talk about it with others. They also go out and help if a teacher needs help using the assessments or looking at the criteria."

Teachers are exhibiting ownership of the process and serving as leaders in decision making as shared by this rural superintendent, "Well, I think the staff is involved at every level in all the decisions. We just don't leave the staff out of the loop in any of the process." Another rural middle school principal shared, "Oh, its staff driven. They have to have ownership of it in order to make it effective. If it comes top down it won't mean anything." Both teachers and leaders are sharing leadership responsibility in the STARS process.

Instructional Impact

In the category, Instructional Impact, there was no significance for leaders. However, Instructional Impact was significant for teachers. A superintendent shared, "the biggest change we've had to do in education is that we're trying to teach for learning as opposed to just teaching." Focusing on student learning has driven instruction to the forefront of the STARS process.

Educators shared the importance of alignment in their conversations with researchers. Teachers emphasized that in order to provide quality instruction, curriculum must be aligned to standards. A female middle-school teacher shared, "I look at the standard. I don't look at just state standards. I look at my curriculum to make sure that, of course, it is aligned...."

Best practice research forms the basis for selecting strategies and interventions as noted by this superintendent, "We study a lot of research. I mean, we spend time looking at tried and proven research-based interventions that work." Using data to determine what your learning gaps may be and how to address those gaps through quality instruction were important to making students successful. A rural elementary principal shared, "We have conversations after we look at our STARS scores. Where are the gaps? How are we going to redirect our curriculum to fill the gaps?" An elementary principal shared, "We choose our interventions based on the data. We've done extensive in-service and gone and looked for good interventions, but all of those interventions are based on the data and the learning style of the child."

Educating all children has become critical to the success of the STARS process as illustrated by this ELL teacher, "For those kids who do not succeed, then there's always another way, another back door that possibly we can try to give them assistance." A special education director shared, "I think the interventions are more challenging and at a higher level. I think I see more varied instructional practices on the part of the general teacher and I think that they've picked up from planning and working with the special education teacher." An elementary principal shared, "So we use our funds to raise the achievement in the whole building. And we have really moved from a remediation model to a prevention model. . . . We figure out what the problem is, what

the deficit is, and we get in there and try to prevent it early." An elementary teacher shared, "I think every year I continue to grow. But I think when you have the specific objective that needs to be taught and mastered, that leads your lessons in the correct direction."

One key component of quality instruction is the grading process and its effect on student success. In order to increase achievement for students, educators are changing their thoughts about the impact of failure for students. A superintendent noted how their district's change in philosophy was supporting greater student success, "We have bought into the 'failure is not an option' concept! We've done some things in terms of our grading policy to enhance student learning. And we are working with our teachers to change a mindset about how students are graded in this district. In other words, I have told the teachers in this district that have given students zeroes for not turning in homework that this is not an option We've done a lot of good things in the district to improve student achievement and student learning and we've got a lot of support in the community to do it."

Ultimately, the goal of STARS was to increase student achievement through quality instruction. This year's interviews yielded an increase in the knowledge of teachers and leaders about improving the quality of instruction in classrooms across the state due to the implementation of the STARS Process.

External Support

External support is provided by a variety of sources including Educational Service Units (ESUs), NDE, colleges and universities, and other external experts. The category of External Support was not significant between leaders and teachers. However, there was a significant difference between rural and non-rural districts. Rural leaders and teachers were more likely to indicate that professional development, ongoing assessment training, data retreats, identification of interventions, and training in instructional strategies were provided by the ESU while non-rural schools tended to provide their own professional development and assessment literacy training. Additionally, assistance in portfolio development and review was provided more by the ESUs for rural schools than their non-rural counterparts.

External training was received from a variety of sources as indicated by this female rural assessment coordinator. She shared several sources including, ". . . training through ESU, training through the Nebraska Department of Education. I was also involved in the Leadership for Learning Assessment Cohort through UNL . . . through my involvement with that group of people, I learned more about assessment and the six quality criteria."

A rural assessment coordinator shared, "I try and facilitate communication between the Nebraska Department of Education and the district, so when we get the updates, those go out to the administrators and to the core chairs and try and keep everybody apprised of what's going on." A rural elementary principal also shared that there were web resources that supported her training, "There's tons of information on the website that is very helpful if people just take the time to read it. It's there."

Due to the size and population of rural districts, external personnel are used in many ways to support the implementation of the STARS process while non-rural schools are much more self-supportive due to their infrastructures.

Special Populations

The theme of Special Populations provided the largest number of quotes from interview participants and yielded evidence of a change in attitudes relative to special populations as a result of STARS. A high school SPED teacher shared, “I think accountability is a good thing. And as teachers we like to say, Oh, yeah, I was accountable before. But it’s a nice way to keep us all teaching the same things at a quality level so that everyone has the opportunity not just the good students.” An elementary principal noted that, “I always felt my special education kids could learn, maybe not at the same rate and maybe not always be on grade level, but I think its made us not let kids fall through the cracks.”

Finally, classroom teachers’ perceptions about accountability relative to special populations was evident as stated by this elementary SPED teacher, “The STARS data has really helped make us accountable for *all kids*, not just for a select few.”

Communication

The interviews revealed a variety of ways that educators are communicating about assessment processes and STARS results with various public groups across the state. A male rural superintendent stated, “I think the community is more aware that we are measuring student performance and are constantly looking at it.” An elementary principal shared, “I think it’s the idea that we’ve got data and we’re expected to share it and we do share it. But also, we have to kind of enlighten parents as to what this really means and we’re in that process.”

Parents are involved in many ways in school districts across the state and learn a great deal about STARS through this involvement. An elementary principal stated, “We have parents involved in our school improvement committees. We have a positive behavior support team. We have parents involved in that. We have parent meetings before school begins, talking about being school-wide Title and our goals.” A non-rural middle school teacher shared, “We have to educate the parents on the terminology like advanced, proficient, progressive; we have to let them know what those mean. I tend to send grade reports home about every 2 weeks to inform parents how their kids are doing.”

The media is used to inform parents about a variety of issues related to school. A rural elementary principal shared, “Each administrator writes an article every month that is published. I have chosen to focus my articles on assisting parents to assist their own children.” Educators are also communicating within their own districts and schools in a variety of ways as indicated by this rural assessment coordinator, “One of the things that we’ve implemented this year across the district is professional learning teams in each building. And so by grade level and by content area, we’re building goals. So teachers working on those teams are collaborating, researching, reading . . .” A rural female SPED teacher indicated that, “We work together not only with the teachers but also with parents and paraprofessionals. And I think that gives us a great deal of knowledge in the collaborative process of knowing what other people are doing.”

In summary, educators in schools and at the district level are collaborating both internally and externally in new ways about the STARS process. During the interviews this year it was more evident that there was a focus on external communication with parents and community.

STARS Process

Overall, there was a positive feeling about the STARS Process with very few educators indicating that it is not an effective process. A rural middle school principal shared, "I think it's been good for the state of Nebraska. I still think that especially with the veteran teachers they would just love to have a state test and be done with all this other stuff because they really just want to be in the classroom, and they want to teach." While an assessment coordinator shared, "I have to say that I value this process for many reasons. What we're doing in Nebraska is right for kids. It's right for teachers. I'm very grateful that we don't have a state test. It's a lot of work and a lot of time but in the long run everything that's worthwhile takes time."

Districts are becoming more effective about implementing the STARS Process in ways that are meaningful to their schools. A rural elementary principal stated, "Well, of course our policy states we will be participating and concluding with our STARS testing by the first of May. And then we will have conversations about how to plan curriculum once we see the results of those tests." A middle school teacher stated, "I think it's a good process. I think it helps level the playing field for all children so that we ensure that all children are receiving at least a baseline of education."

A non-rural assessment coordinator emphasized the impact of the STARS Process on the culture of schools and school districts across the state when she stated, "In terms of culture, I think it has strengthened the culture of the district and strengthened the focus of the teachers of the district." In summary, STARS has emerged as the way that schools do business on a daily basis.

Summary

Assessment literacy is not as prevalent in the discussion with leaders or teachers now that the initial development of most STARS assessments has been completed. Both administrators and teachers emphasized that they were committed to building their own assessment competence and supported "school-based teacher-led assessment." Science teachers are new to the process, but many have been working for several years to get their assessments ready for the reporting year. Leadership has emerged as a key to the success of STARS. A superintendent shared how leadership is critical to the use of data, "I've learned a long time ago if you measure it and you watch that data and you pay attention to it, it becomes important to everybody." It is obvious to the researchers that the use of data in schools has increased over time and is impacting instruction on a daily basis.

Using data to determine learning gaps and create interventions to address those gaps has been an important step in helping students be successful. A rural elementary principal shared, "We have conversations after we look at our STARS scores. Where are the gaps? How are we going to redirect our curriculum to fill the gaps? Where are our needs at? . . . I mean we have to look at any given test results for gaps in learning and make adjustments in curriculum." Focusing on data has helped to align curriculum and drive instruction in schools across Nebraska. In addition to aligning and modifying curriculum for learners, educators want to identify the best instructional strategies and interventions to use so that all students can be successful. In order for the use of research to be successful, it must be based on lots of communication. The biggest change observed by researchers has been the focus upon learning rather than just teaching as shared by this superintendent, "the biggest change we've had to do in education and it's probably of our teachers, is that we're trying to teach for learning as opposed to just teaching."

This has changed belief systems about student failure in particular. An elementary special education teacher shared, "We go back again and we look at it, work on it again, and try to come at it from a little bit different angle." Special populations create a unique set of needs that challenge classroom teachers daily and teachers are meeting the challenge.

STUDY III: 2001 to 2006 Reading and Math Achievement

Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln

Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American

Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹

¹ Posthumous

Introduction

Nebraska's School-based Teacher-led Assessment and Reporting System (STARS) required each district to either adopt state standards or develop local standards that are at least equal to or exceeded the state standards. Each district then developed a plan for assessing their standards. The plan was based on locally developed criterion-referenced tests (CRT's) reported at three selected grade levels. Districts were also required to administer a standardized norm-referenced test (NRT) of their choosing that provided an external common "touch point." The purpose of this study was to examine achievement data available to date, four years each of reading and math. District portfolio ratings for reading and math were also compared.

Methodology

Data were included for all Nebraska Class III, IV, and V school districts and represented just over 94% of the public school students in Nebraska. The district data for this study were available on the state website and use of the data was facilitated by the Nebraska Department of Education.

Findings

The study examined achievement data available to date including four years of achievement data for reading and math. District portfolio ratings for reading and math were also compared.

Reading Criterion-Referenced Achievement Results

The district average percent of student scores reported by districts as proficient or better in locally defined criterion-referenced reading at the fourth-grade-level increased from 74.99% in 2001 to 90.70% in 2006. The district average percent proficient at the eighth-grade level increased from 73.67% in 2001 to 87.70% in 2006. The district average percent proficient at the eleventh-grade level increased from 73.54% in 2001 to 86.10% in 2006. Proficiency on criterion-referenced measures increased at all grade levels each year; the average district gain from 2001 to 2006 was 15.71% at fourth grade, 14.03% at eighth grade, and 12.56% at eleventh grade.

Reading Norm-Referenced Achievement Results

The district average percent of students in the top two quartiles on the norm-referenced reading test used by districts at the fourth grade increased from 64.93% in 2001 to 69.42% in 2006. The eighth grade increased from 62.85% in 2001 to 63.24% 2006. The eleventh grade increased from 59.87% in 2001 to 63.59% in 2006. Proficiency, as determined by the percent of students in districts in the top two quartiles on norm-referenced measures, also increased from 2001 to 2006

with a 4.49% increase at fourth grade, a 0.39% at eighth grade, and 3.72% gain at eleventh grade.

Math Criterion-Referenced Achievement Results

The district average percent of students reported by districts as proficient or better in locally defined criterion-referenced math at the fourth-grade-level increased from 78.29% in 2002 to 90.90% in 2006. The district percent proficient at the eighth-grade level increased from 68.58% in 2002 to 82.90% in 2006. The district percent proficient at the eleventh-grade level increased from 66.22% in 2002 to 80.30% in 2006. Proficiency in criterion-referenced measures increased at all grade levels each year, the gain from 2002 to 2006 at fourth grade was 12.61%, at eighth grade was 14.32%, and at eleventh grade was 15.08%.

Math Norm-Referenced Achievement Results

The district average percent of students in the top two quartiles on the norm-referenced math test used by districts at the fourth grade increased from 68.12% in 2002 to 73.83% in 2006. The eighth grade increased slightly from 67.34% in 2002 to 67.83% in 2006. The eleventh grade increased slightly from 67.49% in 2002 to 67.62% in 2006. Proficiency on norm-referenced measures increased at all grade levels, the gain from 2002 to 2006 at fourth grade was 5.71%, at eighth grade was 0.49% and the gain for eleventh grade was 0.13%.

District Reading Assessment Portfolio Ratings

The total district average of Reading Assessment Portfolio ratings across grades four, eight, and eleven, on the “1” to “5” Likert scale, increased for grade four from 3.57 in 2001 to 4.54 in 2006. Portfolio ratings for grade eight increased from 3.48 in 2001 to 4.56 in 2006. Portfolio ratings for grade eleven increased from 3.46 in 2001 to 4.55 in 2006. The total district average of Reading Assessment Portfolio ratings across grades four, eight, and eleven increased from 3.50 in 2001 to 4.48 in 2006, with a total increase of 3.08 from 2001 to 2006.

District Math Assessment Portfolio Ratings

The total district average for Math Assessment Portfolio ratings across grades four, eight, and eleven, on the “1” to “5” Likert scale, increased for grade four from 3.98 in 2002 to 4.57 in 2006. Portfolio ratings for grade eight increased from 3.96 in 2002 to 4.65 in 2006. Grade eleven portfolio ratings increased from 3.96 in 2002 to 4.61 in 2006. The total district average of Math Assessment Portfolio rating across grades four, eight, and eleven increased from 3.97 in 2002 to 4.61 in 2006, with a total increase of 1.93 from 2002 to 2006.

Summary

District criterion-referenced measures continue to show growth over time in the areas of reading and math from 2001 to 2006, generally increasing in reading for all grades - fourth, eighth, and eleventh. District math norm-referenced measures have increased for grade four and slightly increased for grades eight, and eleven. School improvement with student academic achievement as the goal is not a short-term process. Nebraska is in its sixth year of full implementation of the program and variability still exists in achievement in some areas and for some grades.

STUDY IV: 2002-2006 Statewide Writing Achievement
Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln
Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American
Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹

¹ Posthumous

Introduction

With an emphasis to demonstrate increased writing achievement for all students, Nebraska chose to use the Six Trait writing model as a requirement for a statewide writing assessment, and it was included in the legislation establishing Nebraska's assessment system. Research supported large-scale assessment because it enhanced the ability of teachers to write themselves, to assess student work, and to rate student writing. This approach created higher expectations for student writing performance and prompted decision makers to ask important questions--when, where, and how writing was being taught. This study examined the district achievement data available for the Statewide Writing Assessment for the Nebraska STARS Program.

Methodology

Data was included for school districts in Classes III, IV, and V that represented just over 94% of the public school students in Nebraska. The process for the development of writing prompts for use in the Statewide Writing Assessment relied on the involvement of Nebraska classroom teachers who were selected by the Nebraska Department of Education (NDE) to participate in a writing development task force upon recommendation of their district superintendent or assessment contact person. From this task force, writing prompts were chosen and field-tested with students in grades four, eight, and eleven. Nebraska teachers, chosen for their expertise, then scored the writing assessments at a central location. The unit of analysis for this study was the district average percent of students rated as proficient or better in school districts in Classes III, IV, and V for the state of Nebraska in writing at grades four, eight, and eleven. Because the assessment was a common measure across districts and was an equal interval scale, inferential statistics were also used to examine statistically significant differences between pre/post test scores from inception to last scoring.

Findings

The purpose of this study was to examine the district achievement data available for the Statewide Writing Assessment for the Nebraska STARS Program. In the ratings for writing proficiency, gains were made at nearly all grades and years of comparison. At fourth grade, gains from 2002 to 2004, and 2004 to 2005, were significant ($p < .001$) with a decline indicated in 2006. At eighth grade, the gain from 2003 to 2004 was significant ($p < .001$); the gain from 2004 to 2005 was not significant. There was an increase in 2006 to 87%. The gain in eleventh grade from 2004 to 2005 was significant ($p < .05$) with a slight gain in 2006.

Writing scores increased for all grades, with grade four increasing 6.91% from 2002 to 2006, grade eight increasing 7.45%, and grade eleven increasing 1.88%.

In examining teacher perceptions and classroom practices related to the state writing assessment, results indicated that 69% of all teachers placed more emphasis on practice writing assessments,

73% placed more emphasis on sharing assessment criteria in class, and 73% placed more emphasis on explicit instruction in six trait writing.

Summary

The purpose of this study was to examine the district achievement data available for the Statewide Writing Assessment. Results indicated that fourth, eighth, and eleventh graders made gains in the pre/post comparisons on the Statewide Writing Assessment. The writing data, fourth grade comparing data from 2002 to 2004, and eighth grade comparing data from 2003 to 2004, revealed strong gains at both grades. This may reflect the involvement of many school districts with Six Trait Writing preceding the formal statewide writing assessment process. Students at the eighth and eleventh grades continued to show growth on their writing scores from 2004 to 2006. However, fourth-grade scores dropped slightly in 2005 to 2006 with an overall gain from 2002 to 2006 of 6.91%. While this writing assessment has characteristics that enable inferential statistical analysis to be used, there may be some question from the traditional measurement community as to whether this practice would be appropriate. It must be remembered that the philosophy and purpose of Nebraska STARS is to support teaching and learning, and not to focus on development of assessments of technical strengths in ranking results.

STUDY V: 2001-2006 Achievement for Special Populations

Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln

Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American

Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹

¹ Posthumous

Introduction

A key component of No Child Left Behind was to demonstrate increased academic achievement for all students. States have taken different approaches to accomplish this end. In Nebraska there was no single high-stakes test. STARS has been in place for six years and has a strong success record with total group data (Roschewski, Isernhagen, & Dappen, 2006). This is the second year that English Language Learners (ELL) and Special Education (SPED) populations were examined within the Comprehensive Evaluation for STARS.

The purpose of this study was to examine the academic change of special populations, specifically ELL and SPED students from Nebraska. This study examined criterion-referenced assessment results for reading, math, and writing.

Methodology

The criterion-referenced score (CRT) for reading and math was the district average percentage for ELL and SPED students meeting the proficiency level or better as defined by the local district for their locally developed measure for Class III, IV and V school districts for the state of Nebraska. Because the reading and math measures were not on a common scale, only descriptive data was provided.

The unit of analysis for the criterion-referenced writing score was the district average percent of ELL and SPED students rated as proficient in Class III, IV and V school districts for the state of Nebraska in writing at grades four, eight, and eleven. Because the assessment was a common measure across districts and was an equal interval scale, inferential statistics were also used to examine statistical significance between pre/post scores from inception to last scoring.

Findings

ELL Reading Achievement

The district average percent of English Language Learners (ELL) reported as proficient or higher in locally defined criterion-referenced assessments for reading at the fourth-grade level increased from 50% in 2001 to 72% in 2006. From 2001-2006 proficiency on reading criterion-referenced measures for grade four increased 22%.

The district average percent of ELL scores for eighth-grade students increased from 47% in 2001 to 60% in 2006. Proficiency on reading criterion-referenced measures for eighth-grade ELL students increased 13% from 2001 to 2006.

The district average percent of ELL scores for eleventh-grade students increased from 45% in 2001 to 53% in 2006. Proficiency on reading criterion-referenced measures for eleventh grade ELL students increased 8% from 2001 to 2006.

ELL Math Achievement

The district average percent of English Language Learners (ELL) reported as proficient or better in locally defined criterion-referenced assessments for math at the fourth-grade level increased from 53% in 2002 to 80% in 2006. Proficiency on math criterion-referenced measures increased 27% from 2002 to 2006.

The district average percent of ELL scores for eighth-grade students increased from 40% in 2002 to 61% in 2006. Proficiency on math criterion-referenced measures for eighth-grade ELL students increased 21% from 2002 to 2006.

The district average percent of ELL scores for eleventh-grade students increased from 39% in 2002 to 48% in 2006. Proficiency on math criterion-referenced measures for eleventh grade ELL students increased 9% from 2002 to 2006.

ELL Writing Achievement

The district average percent of English Language Learners (ELL) reported as proficient or better on the state criterion-referenced assessment for writing at the fourth-grade level increased from 49% in 2002 to 66% in 2006. Proficiency on the state writing criterion-referenced measure increased 17% from 2002 to 2006.

The district average percent of ELL scores for eighth-grade students increased from 37% in 2003 to 56% in 2006. Proficiency on the state writing criterion-referenced measure for eighth-grade ELL students increased 19% from 2003 to 2006.

The district average percent of ELL scores for eleventh-grade students increased from 45% in 2004 to 53% in 2006. Proficiency on the state writing criterion-referenced measure for eleventh-grade ELL students increased 8% from 2004 to 2006.

SPED Reading Achievement

The district average percent of Special Education (SPED) students reported as proficient or better on locally defined criterion-referenced assessments for reading at the fourth-grade level increased from 44% in 2001 to 74% in 2006. Proficiency of SPED fourth-grade students on reading criterion-referenced measures increased 30% from 2001 to 2006.

The district average percent of SPED eighth-grade students increased from 43% in 2001 to 66% in 2006. Proficiency on reading criterion-referenced measures for eighth-grade SPED students increased 23% from 2001 to 2006.

The district average percent of SPED scores for eleventh-grade students increased from 42% in 2001 to 61% in 2006. Proficiency on reading criterion-referenced measures for eleventh-grade SPED students increased 19% from 2001 to 2006.

SPED Math Achievement

The district average percent of Special Education (SPED) students reported as proficient or better on locally defined criterion-referenced assessments for math at the fourth-grade level increased from 51% in 2002 to 75% in 2006. Proficiency on math criterion-referenced measures increased 24% from 2002 to 2006.

The district average percent of SPED eighth-grade students increased 34% from 2002 to 56% in 2006. Proficiency on math criterion-referenced measures for eighth-grade SPED students increased 22% from 2002 to 2006.

The district average percent of SPED scores for eleventh-grade students increased from 28% in 2002 to 46% in 2006. Proficiency on math criterion-referenced measures for eleventh-grade SPED students increased 18% from 2002 to 2006.

SPED Writing Achievement

The district average percent of Special Education (SPED) students reported as proficient or better on the state criterion-referenced assessment for writing at the fourth-grade level increased from 46% in 2002 to 60% in 2006. Proficiency on the state writing criterion-referenced measure increased 14% from 2002 to 2006.

The district average percent of SPED eighth-grade students increased from 48% in 2003 to 61% in 2006. Proficiency on the state writing criterion-referenced measure for eighth-grade SPED students increased 13% from 2003 to 2006.

The district average percent of SPED scores for eleventh-grade students increased from 55% in 2004 to 64% in 2006. Proficiency on the state writing criterion-referenced measure for eleventh-grade SPED students increased 9% from 2004 to 2006.

Summary

While the ELL and SPED district average student scores were not as strong as the district average student scores for all students, they were increasing in all instances. This is consistent with most research in this area and the basis for the special supports provided for ELL and SPED students. The reading scores for SPED students reflected strong positive change in all areas.

ELL student scores at the fourth and eighth grade levels were somewhat stronger than those of eleventh-grade students. This trend could relate to the difficulty of ELL students mastering English as a second language at an older age.

STUDY VI: Use of Data to Inform Decisions by Elementary Principals

Jadi Miller, Principal, Everett Elementary, Lincoln, NE

Introduction

With its emphasis on local control, the state of Nebraska's assessment system or School-based Teacher-led Assessment and Reporting System (STARS) places a great deal of responsibility on local educational leaders, particularly building principals. The STARS system was designed to improve assessment literacy among Nebraska's educators and to provide schools with the student achievement data necessary for systemic school improvement. This study was designed to determine the extent to which Nebraska's elementary building principals use data to inform their decisions.

Methodology

Using the National Staff Development Council's standards for data-driven decision-making, the researchers designed an online survey about the practices of building principals related to data use. The 60-question survey (Appendix J) was organized around five specific outcomes for data use. These outcomes included involving the faculty in data analysis, using disaggregated data, involving faculty, parents, and the community in data analysis, using data to plan professional development, and using data to monitor school improvement efforts. The survey was sent to 521 elementary building principals in Nebraska. There was a return rate of 32% or 168 responses.

Findings

The data generated by this survey provided insight into the use of data in Nebraska's elementary schools. Overall, it appears that the majority of schools in this study are demonstrating proficiency according to the National Staff Development Council's standards for data-driven decision-making.

The elementary principals indicated that they were involving their school faculty in the analysis of data. This particular section generated the highest mean score across all of the items in the sections. The elementary principals also indicated a strong level of agreement that they were using disaggregated data in their analysis process. However, a majority of the respondents indicated that they were not involving parents and/or community members in the process.

When measuring the elementary principals' use of data to plan for professional development, the mean was lower than the previous areas of the survey. However, one of the initial questions about the alignment of the school's improvement goal and staff development plan indicated that for the vast majority of respondents, 97%, alignment was present. On other questions related to monitoring staff development and involving the whole faculty, respondents indicated much lower levels of agreement on the survey items.

The frequency of using data to monitor the school improvement process, with questions focusing on the frequency of these types of activities, yielded the lowest mean of the survey, with a score

closest to “Sometimes.” While the mean of this section was lower, no more than 4% of respondents selected the choice of “Never” on any of the items.

Summary

The results indicated that elementary principals in Nebraska most frequently used disaggregated data and involved their staff members in data analysis. There were lower scores in the areas of involving parents and community members in the process. Principals also displayed less frequent use of data for planning professional development and monitoring the school improvement process. Finally, the results of this survey indicated that principals are using data to a significant extent.

STUDY VII: Role of Technology in Supporting STARS

Dr. Allen Steckelberg, Associate Professor, University of Nebraska-Lincoln
Lan Li, University of Nebraska-Lincoln

Introduction

STARS evaluations conducted in the past suggested that further information on technology, the issues it both addresses and creates, would be helpful as schools and the Nebraska Department of Education develop policy and create support resources for technology. For the past two years a survey has been conducted to examine the perceptions of educators concerning the current uses of technology, the impact of technology, and future systems and issues.

Methodology

A 2005-06 survey was conducted to examine the perceptions of educators concerning the current uses of technology and its impact upon the STARS process. In 2006-07 minor changes were made to the prior survey tool and an additional question was added on the impact of the Assessment and Reporting Management System (ARMS) grants. The survey (Appendix K) was constructed in four sections. The first section of the survey gathered demographic information from participants. The second section of the survey gathered information on the current uses of technology including: software used by schools in various aspects of the assessment process, the match between technology-based systems and the Nebraska Assessment Model, and the quality of technology used by the district in the assessment process. The third section addressed respondents’ perceptions of the impact of technology on the assessment process including: contributions of technology to various aspects of the assessment process, impact on time and complexity of the process, impact on the value of data collected and understanding and use of that data. The final section of the survey addressed perceptions of future assessment systems and issues that might guide future policy decisions.

The survey was administered via a Web site. The survey targeted school personnel who coordinated the STARS Process in local schools and ESUs in Nebraska. From a list of STARS school contacts provided by the Nebraska Department of Education (NDE) and the STARS Comprehensive Evaluation Project, 463 names and email addresses were identified. An email solicitation to participate in the survey was sent to each name on the list. A follow-up email was sent two weeks later to those who had not yet responded. A total of 227 usable survey responses were received and used in the analysis.

Findings

Knowledge of the characteristics of typical respondents was helpful in interpreting the results. The distribution of school size collected from survey respondents closely reflected the distribution of school sizes statewide. Respondents were primarily from smaller schools (70%) but were often part of larger consortiums for the purpose of implementing STARS. Respondents were often superintendents, principals, or assessment coordinators, with teachers and curriculum coordinators also being represented. The majority indicated funds were available for technology, particularly if it reduced other administrative costs or contributed to school improvement. In a new item, introduced in the 2007 administration of the survey, 67% of respondents agreed that the ARMS grants for increasing technology dollars directly impacted their ability to invest in assessment technology. The most frequently reported investment of funds was for management purposes (42%).

The survey provided useful information on the current uses of technology. The majority of respondents indicated that basic productivity tools such as word processing and spreadsheets were the most extensively used tools in a number of aspects of the assessment process. Dedicated assessment software, Web-based software, and locally developed software or Web applications were used less extensively. When broken down by school size, the rated extent of use of dedicated and locally developed software was proportionally higher in larger schools than in smaller schools possibly because of differences in resources. Scanning, computer and Web-based systems, and electronic portfolios had low ratings for use.

Several survey items targeted general perceptions regarding current technology and the assessment process. A majority (87%) of respondents agreed or strongly agreed that the technology-based systems were a good match to the Nebraska assessment model. This is the only item in the survey that differed markedly from the 2006 survey results. In the 2006 survey, only 47% of the respondents agreed or strongly agreed with the statement. The 2007 result is substantially more positive. Respondents also scored the rating of the quality of the technology used by their district in the assessment process at a moderately to moderately high level. These results would seem to indicate that there is a potential to improve and increase the use of technology in the assessment process in many Nebraska schools.

Data from the survey clearly showed that survey participants felt technology has the potential to positively impact the assessment process. A majority of respondents indicated that technology reduced the amount of time involved, increased the value of assessment data, assisted teachers in understanding and using assessment data for instructional decision-making, and, at the same time, did not increase the complexity of the assessment process. Respondents also indicated that they saw a number of potential contributions of technology, such as consistent data collection, as having a moderate to extensive impact.

Survey data also provided information useful for guiding future policies, support, and resources. The most significant barriers were costs of technology, staff, and administrator time to learn systems and teacher training. Interoperability was also rated as an issue in larger school districts. Quality ratings of current systems regarding their comprehensiveness, quality, reporting features, and interface with state reporting were moderate.

Interestingly, ratings of classroom level reporting features and interface with state reporting, although still moderate, were higher in small schools. A number of features of technology-based

assessment systems were rated as highly desirable including: usefulness of results for instructional decisions, ease of data entry, match with reporting requirements, and ease of learning systems. Survey participants felt that the areas of state support with the greatest potential for impacting were grants for technology implementation and teacher training. These areas were closely followed by the provision of exemplars in the form of models for data collection, management, and reporting and software templates for collecting and organizing data.

Summary

Results from the survey indicated that STARS assessment leaders in Nebraska schools recognized the potential impact of technology on the assessment process, but at this point do not extensively make use of dedicated or locally developed software, although this may be occurring in larger districts or consortiums. It appeared that, among school personnel who work with the STARS assessment process, technology was viewed as having a positive impact. Schools saw grant support for implementing technology and training school staff as likely to have significant impacts on the quality, usefulness, and value of the assessment process. Survey participants felt that the areas of state support with the greatest potential for impact were grants for technology implementation and teacher training.

STUDY VIII: Statewide Writing Assessment Scoring: Effective Professional Development for the Classroom Teacher

**Sue Anderson, Director of Professional Development
Educational Service Unit #3**

Introduction

The purposes of this study were to learn raters' perceptions of the effect of the scoring experience on their classroom practices, and the effectiveness of the training they received during the scoring of the 2007 Nebraska Statewide Writing Assessment.

Methodology

The population of this study consisted of 390 raters in the scoring of the 2007 Nebraska Statewide Writing Assessment, with 131 raters at Grade 4, 129 raters at Grade 8, and 130 raters at Grade 11. A scoring session was conducted for each grade level. Raters were selected to participate in the scoring sessions based on their teaching experience, including grade level(s) taught and years of teaching, their previous writing assessment scoring experience, prior training in the Six Trait writing assessment model, and the geographic region of the state they represented with approximately 50% raters for each scoring session representing the eastern region of the state, 30% of the raters from the central region, and 20% percent from the western region. Among raters in the Grade 4 scoring session, 46 school districts were represented; for Grade 8, 47 districts were represented; and for Grade 11, 47 districts were represented.

The data for this study was collected from rater responses to a survey entitled 2007 Statewide Writing Assessment Scoring Rater Evaluation (Appendix L) created by the researcher. The survey items were intended to collect information about the raters, their perceptions of the effectiveness of the training provided during the scoring sessions, and the relationship between the raters' scoring experience and their classroom practices in the teaching of writing.

All survey responses were coded and tabulated at the Nebraska Department of Education. Analysis included the tabulation of frequency of responses for all items and the calculation of mean scores for items asking raters to indicate their levels of agreement and disagreement. The survey contained 29 questions across the following categories: rater information, scoring training, the scoring experience and its relationship to classroom practices, and comments related to reasons why raters wanted to participate in the scoring sessions.

Survey items related to the raters' scoring experience and their classroom practices asked them to indicate their levels of agreement and disagreement with how the scoring experience affirmed their current instructional practices, increased their understanding of the qualities of good writing, increased their confidence to provide meaningful feedback to students, and increased their confidence as teachers of writing. Raters indicated their levels of agreement and disagreement with items in this section of the survey on a scale of "strongly disagree" to "strongly agree."

A single open-ended item on the survey asked raters to explain their reason(s) for wanting to participate in the scoring session(s). A final item invited raters to provide additional comments on any aspect of the scoring experience they wished to share.

Survey items related to scoring training asked raters to respond to the effectiveness of the content of the training, strategies to ensure scoring accuracy, clarity of training materials, and the knowledge and skill of scoring leaders and table leaders. Raters indicated their levels of agreement and disagreement with items in this section of the survey on a scale of "strongly disagree" to "strongly agree."

Findings

This study was designed to ascertain the relationship between the 2007 Statewide Writing Assessment scoring experience of raters and their classroom practices as teachers and to determine the effectiveness of the procedures used to train them to score the assessment. Responses collected from raters pertained to broad categories: general rater information, scoring training, and scoring experience related to their teaching practices. The results across all grade level scoring sessions showed that raters associated the scoring experience with benefits to their teaching and evaluated all components of the training they received as very positive.

Participation in the scoring of the 2007 Statewide Writing Assessment was perceived by a majority of raters across all grade levels as beneficial to their classroom practices in the teaching of writing. Responses by raters were related to increasing their understanding of the qualities of good writing, building their confidence as a teacher of writing, affirming their current classroom practices, providing helpful feedback to students about their writing, and other areas related to the teaching of writing. Ninety-seven percent (97%) reported that being a rater increased their confidence to assess the quality of student writing, and 94% believed their scoring experience increased their understanding of the qualities of good writing and helped them develop more accurate grade-related expectations for student writing.

Across all grade levels 90% of raters agreed or strongly agreed that being a rater affirmed their current teaching practices, contributed to their ability to help students become better writers, and increased their general confidence as a teacher of writing. A majority of raters, 89%, agreed that being a rater gave them confidence to give students helpful feedback about their writing, and

86% reported that being a rater gave them assurance in talking to parents and others about writing and its importance. Additionally, being a rater gave 85% across all grade levels a better understanding of how writing relates to thinking and reasoning, and gave 82% of raters a greater appreciation of writing's uses in content area learning. Seventy-nine percent (79%) of raters agreed or strongly agreed that being a rater gave them new ideas they could use in the teaching of writing.

An overwhelming majority of raters across all grade levels agreed or strongly agreed that all aspects of the scoring training were effective. Ninety-nine percent (99%) of raters agreed or strongly agreed that the general content of the training (i.e. the scoring process, scoring criteria, six traits of writing, writing modes) was appropriate while 97% indicated the strategies to ensure scoring accuracy (i.e. using the rubric to assign scores, using anchor papers to demonstrate performance levels, use of re-calibration papers, and consensus building discussions) were effective. Ninety-five percent (95%) of raters agreed that the training materials were clear and understandable.

A majority of raters, 98%, agreed or strongly agreed that scoring leaders and 97% agreed or strongly agreed that table leaders demonstrated expert knowledge of the assessment and the scoring process. Ninety-seven percent (97%) of raters agreed that scoring leaders were skilled in teaching them to score the assessment, and 95% agreed that table leaders were effective in facilitating the scoring process through re-calibration procedures.

Summary

Based on the results of this study, it is possible to draw important conclusions about the raters and their perceptions of the scoring process. First, the scoring experience was perceived by a majority of raters as a valuable professional development opportunity that contributes to their knowledge and understanding of the teaching and assessment of student writing. Secondly, raters' perceptions about the effectiveness of the scoring training were very positive across all grade level scoring groups. What this means is that the training, including the content, materials, training personnel, and procedures, was appropriately designed and delivered in such a way that raters felt they were adequately prepared to score the assessment.

The findings of this study support the continuation of a state level scoring process that enlists the participation of classroom teachers as a way to provide them with valuable training that relates positively to their classroom practices.

STUDY IX: The Effect of an Intensive Writing Instruction Assessment Preparation Program on Student Writing Proficiency

**Dr. Edward J. Bennett, Assistant Principal
Omaha Central High School, Omaha, NE**

Introduction

The purpose of this pretest-posttest two group comparative survey study was to determine whether an intensive writing instruction and assessment preparation program (IWIAPP) was effective in helping eighth-grade students improve their performance on the Nebraska Statewide Writing Assessment (NSWA). To help students succeed on the NSWA, the research district developed its own Fall District Writing Assessment (FDWA) which is administered to all

students in grades four, eight, and eleven each September. District personnel score the FDWA in a fashion consistent with the NSW, and scores are reported to individual buildings to help inform the writing instruction in the weeks between FDWA and NSW. From a study design perspective, FDWA and NSW created a natural pretest posttest opportunity in which the effectiveness of a school's writing preparation program could be measured and analyzed. The students who took part in this study were eighth graders attending a 950-student middle school that served students in grades five through eight. Although all students took part, Special Education students, ELL students, and all students not attending the research school during seventh grade were not included in the unit of analysis. In addition to FDWA and NSW scores, the study also analyzed California Achievement Test language mechanics and language expression subtest Normal Curve Equivalence (NCE) scores from grades seven and eight and classroom behavior as measured by office referral and tardy to class frequencies.

Methodology

The Intensive Writing Instruction and Assessment Preparation Program (IWIAPP) was administered to all students whose scores ranged from 2 through 4 on the FDWA ($N = 74$) during regular classroom time by both of the research school's eighth-grade language arts teachers. Identified students spent approximately 90 minutes per week participating in the program. The three unique characteristics of the program were: (a) explicit instruction on the writing process; (b) explicit instruction by the teacher on one trait of the Six Traits of Writing per week; and (c) consistent use of the Six Trait based State Assessment Scoring Rubric by students and teachers.

Students whose scores ranged from 5 through 8 on the FDWA ($N = 86$) comprised the control group of the study. During regular classroom time, the control group students engaged in less structured assignments and followed a more standard language arts curriculum. Specifically, they wrote drafts of descriptive writing independently, worked on a portion of a poetry unit, and engaged in free reading of novels of their choosing.

Findings

Pretest to posttest gains made by IWIAPP students ($N = 74$) revealed significant improvement in writing performance on NSW. The pretest FDWA score ($M = 3.54$, $SD = .80$) compared to the posttest NSW writing score ($M = 5.11$, $SD = .95$) was statistically significantly different, $t(72) = 12.13$, $p = .0001$ (one-tailed), $d = 1.79$. Furthermore, 93% of IWIAPP students improved their scores from FDWA to NSW.

The control group of eighth-grade students ($N = 86$) also demonstrated significant pretest to posttest gains. The pretest FDWA writing score ($M = 5.88$, $SD = .85$) compared to the posttest NSW writing score ($M = 6.13$, $SD = .78$) was statistically significantly different, $t(84) = 2.37$, $p = .01$ (one-tailed), $d = .30$.

Despite the significant gains made from FDWA to NSW, California Achievement Test language mechanics and language expression subtest Normal Curve Equivalence (NCE) scores dipped significantly for both groups in three of the four subtests. However, the achievement qualitative description of these scores at posttest remained in either the average or above average range for all participants. Furthermore, while a statistical difference between the groups of students was found for behavior measures, the second semester levels of office referrals and tardies to class observed would be considered low for middle school students in general.

Summary

While not all IWIAPP students demonstrated proficiency at posttest, the overwhelming majority improved, which indicates the IWIAPP intervention succeeded in improving student writing. The results of the study also serve to underscore the value of the three underlying components of the IWIAPP intervention: explicit instruction in the writing process, explicit instruction in the use of a common language, and frequent references to a scoring rubric clarifying the definitions of effective writing. While it is unclear which of these components had the greatest impact on student performance, the results leave little doubt that the intervention as a whole revealed the codes of writing to the participants in a way that they internalized and applied with overwhelming success. The study's results should encourage local and state education officials to consider implementing the IWIAPP intervention to help more struggling writers in middle schools across the research district and the state.

CONCLUSIONS FOR STUDIES I-IX

STARS . . . The First Five Years

(Adapted from *A STARS Symphony (in D Major)* (Gallagher & Isernhagen, 2006))

Learning is hard, and it takes time. No one knows this better than Nebraska educators. The past five years have been spent listening to educators around the state tell their stories of learning. Their words portray the journey of the implementation of the STARS Process. Nebraska educators have learned the bumps in the road and the challenges of becoming classroom leaders of assessment. Teachers and leaders have had to listen, practice, learn, and grow—*together*. They have had to become experts in their own area *and* contributors to the new process. They have become technical experts in assessment quality and design.

Not all districts or schools have mastered the process. Some have come a long way on the journey and others have a long way to go. But when researchers analyzed the surveys and listened across the 600-plus interviews conducted, stories of both individual and collective talent emerged.

Starting out, was a difficult process with a great deal of change as illustrated by this teacher, “So where do we begin? Do we assess every single standard? Do we clump them together?” Teachers and leaders felt that they were not experts but then began to realize that others could provide guidance and training as shared by this staff developer, “Our primary role has been a facilitator of the design and refinement process . . . We really facilitate the process of both the development and the refinement of criterion referenced assessments.” With some practice, and some patience, Nebraska educators began to see the possibilities and what they truly were capable of doing when acting in the best interest of students.

As the knowledge and confidence of educators grew they began to see the rewards. Of course, you still make mistakes; old habits die hard as a teacher shared, “We’re testing too much with pen and pencil and paper.” It’s difficult to find time to learn new skills and implement them successfully. But educators continue to learn and grow. Slowly and together they began to realize they were more knowledgeable and therefore, able to develop the assessments needed to measure student achievement and use those measures to improve teaching and learning as shared by this superintendent, “I think our staff is light years ahead of where they used to be in developing assessments. I think they think more about planning backwards. In other words,

they say, ‘What are we assessing and then how do we get our student to that?’ I think they’re more concerned with what they’re assessing and how students are progressing.”

This work is not easy, and not everyone catches on. Sometimes, it seems you have too much data to use or that is readily available. Districts and schools lacked the data management systems to provide timely feedback to teachers and students. With practice, help, and lots of time, they learned how to use the information and the conversation to impact learning.

Educators began to understand the value of assessment as stated by this Assessment Coordinator, “If we’re good assessors, if we really know where students are and what their knowledge and skills are, then we can also identify the gaps that they have and, if they are not meeting the assessment criteria, provide additional instruction, reassess, and hopefully, we won’t have . . . students falling through the cracks. And that, to me, is the real goal of assessment, the real goal of the whole STARS process.”

Achieving is about making a difference. Knowing and reaching each learner is critical to achievement as explained by this teacher, “Instead of focusing on just what your objectives are, [classroom assessment] focuses on what your objectives are *for each student*. In essence, then, I think that each student almost has an individual education plan.” Educators are recognizing that “all students” includes those in special populations as illustrated by this assessment coordinator, “I think one of the strong, positive things about STARS is that we’ve disaggregated our data to see if we have any discrepancies in delivering services to special populations: SPED, Title I students, gender [and so forth]. I remember a couple years ago . . . [we disaggregated] the SPED population. The day we got the data we were already having the think tank team and saying, ‘Why do we see this large of a discrepancy at this grade level? What support services or what interventions can we provide to the students to help them increase their skill level?’”

Sometimes, it’s about challenging students—moving them outside the familiar, the comfortable as a teacher shared, “It’s kind of raised my expectations for my students, even the lower ones. You have that expectation that they should get to a certain level by the end of the year. So . . . my expectations for my students have risen since we’ve implemented STARS. I probably do a better job of teaching.”

Additionally the work has impacted instructional practices in schools as shared by a superintendent, “What we’ve seen is that we went from the beginning of this process spending all of our time on curriculum and assessment development and nothing on the instructional piece. It was the curriculum assessment process. And now it’s the curriculum, instruction and assessment process. We’ve evolved to the point that we are now focusing on the instruction piece. We’re [working in] learning teams. We’re focusing on strategies and interventions. So we have used data to be a driving force in interventions that we are providing and will continue to provide, and also a driving force in staff development.”

Leadership has truly become distributed as referenced by this teacher, “Within my building, I’ve been a leader as far as helping other teachers understand the assessment process. When they have a question as to how to administer a certain assessment, they would come to me and we would discuss how you go about doing that so that there is reliability and validity within the system, using the appropriate scoring guide, and just the approach that the teachers take and the ways in which the assessments are to be scored.”

Along the journey more educators see the lighted path to improving instruction for all students and commit to the continuous improvement as stated by this assessment coordinator, “STARS . . . has brought everybody together . . . so everybody is on the same page We are all where we need to be and the kids are all where they need to be . . . everybody is more focused.”

Instruction is more important than ever and educators have begun to understand that without the journey they could not have produced the same results. They are colleagues, partners, co-creators — all pulling in the same direction as shared by this assessment coordinator, “You see teachers having more conversations about curriculum and instruction and assessment than ever before . . . I have been a teacher for a number of years. [Before,] everybody did their own little thing—you know, you went in your room and you did it any way you chose as long as you covered basic things. And now teachers are working together, which to me is more equitable for kids.”

STARS...The Sixth Year

The path to improved student achievement has not been without bumps in the road. However, Nebraska educators have committed themselves to an improved assessment process and learning for all students. During the sixth year, educators continued their STARS journey with continued commitment and provided new stories about the STARS Process. These stories exhibited a range of perceptions that have changed over the last six years as growth within promotes growth in the system. Assessment coordinators were positive about the STARS Process and the progress that was being made. This change in perception was due, in part, to the evidence of the instructional impact and improvement of student achievement. Finally, STARS has truly begun to impact the culture of schools and school districts across the state as shared by this superintendent, “. . . I think it has strengthened the culture of the district and strengthen the focus of the teachers of the district. They will say that they are more focused. The ones [teachers] that we have put the most burdens on are the ones that are most in favor ‘You give these teachers one more thing when they can’t possibly take [on] one more thing, but they take ownership for it and they walk out with more energy than they had before.’”

Educators are acknowledging that some of the earlier issues that plagued STARS such as the amount of time required, the intimidation of the process, and the fear that tests would determine their jobs and the money their schools received are less existent now. This was particularly evident as science teachers began developing and administering their assessments for the first time . . . the fear factor has disappeared. Assessment literacy is not as prevalent in the discussion with leaders or teachers now that the initial development of most STARS assessments has been completed. Both administrators and teachers emphasized that they were committed to building their own assessment competence and supported school-based teacher-led assessment. Science teachers are new to the process, but many have been working for several years to get their assessments ready for the reporting year. Implementing assessments in science was an expectation and somewhat uneventful as was the arrival of a new statewide district assessment portfolio process introduced during the 2006-07 school year. School districts showcased their assessments, curriculum, and instruction to assessment experts in Nebraska and from across the nation. Teams of experts visited all school districts. Most school districts demonstrated exceptional understanding of the STARS Process and showcased their own process based on the six quality criteria. The Peer Review Process also revealed isolated cases of non- involvement in

the STARS process. It became evident to schools during the Peer Review Process that all schools must be involved. Hands down, people preferred the new Peer Review Process over just submitting a portfolio to the state annually. However, in some cases with consortium schools, educators did not feel that the evidence they provided to reviewers was valued due to the number of previous consortium schools visited.

Overall, participants felt that the Department of Education had shown a tremendous amount of faith in the integrity of the Portfolio Review Process and indicated confidence in the STARS process. There was a strong positive reaction to the review process as indicated by many educators who experienced anxiety about their preparation of their staff and students but indicated confidence in the process itself. As the six quality criteria were reviewed across the state, Alignment was viewed by many educators as the impetus for change in schools as teachers and administrators interacted with a great deal of discussion regarding how “curriculum and assessments matched the standards”. Sufficiency seemed to be one of the most complex criteria for districts to understand while a low level of understanding existed in the area of Clarity.

In the area of Appropriateness, bias review was one of the stronger areas relative to the six quality criteria. As with any new process, there were challenges. Nebraska's rural districts are isolated and difficult to access, keeping a large enough group of reviewers trained and accessible to conduct annual reviews, scheduling the events for the day enough in advance so districts can adequately plan, and the challenge to work with districts in a consortia (write a common portfolio) was a problem for both reviewers and the districts being reviewed as some ratings within the consortia were not consistent. Several schools who were first to be reviewed and received "pending" ratings, even though these could be rectified before the final submission, indicated frustration. Overall, the process allowed districts to re-evaluate the work that had been done and the work that needs to be accomplished to reinforce student learning. Next steps will need the support of leadership to continue to build upon the success of the process at the classroom level.

Leadership has emerged as a keystone to the success of STARS. A superintendent shared how leadership is critical to the use of data, "I've learned a long time ago if you measure it and you watch that data and you pay attention to it, it becomes important to everybody." It is obvious to the researchers that the use of data in schools has increased over time and is impacting instruction on a daily basis. The results of a study on the use of data by elementary principals conducted this year indicated that Nebraska principals most frequently involved their staff members in data analysis and that they were using disaggregated data. However, principals in this study showed less frequent use of data for planning professional development and monitoring the school improvement process. There were also lower scores in the areas of involving parents and community members in the process. Although some individual item scores were low, the overall results of the study seemed to indicate that the elementary principals were using data to a significant extent. The NDE, the ESUs, and individual school districts should be commended for their work in not only generating data for accountability but for using data effectively, according to national standards, to improve their schools.

Using data to determine learning gaps and interventions to address those gaps has been an important step in helping students become successful. A rural elementary principal shared, "We have conversations after we look at our STARS scores. Where are the gaps? How are we going

to redirect our curriculum to fill the gaps? Where are our needs at? . . . I mean we have to look at any given test results for gaps in learning and make adjustments in curriculum." Focusing on data has helped to align curriculum and drive instruction in schools across Nebraska. In addition to aligning and modifying curriculum for learners, educators want to identify the best instructional strategies and interventions to use so that all students can be successful. A superintendent noted, "We study a lot of research. I mean, we spend time looking at tried and proven research-based interventions that work." In order for the use of research to be successful it must be based on collaboration and communication which is evidenced by the use of professional learning communities as a means for formal collaboration about research and interventions. But the largest change observed by researchers has been the focus upon learning rather than just teaching as shared by this superintendent, "the biggest change we've had to do in education, and it's probably of our teachers, is that we're trying to teach for learning as opposed to just teaching."

Teaching for learning has been evidenced by the fact that district criterion-referenced measures continue to show growth over time in the areas of reading and math from 2001 to 2006. District norm-referenced measures have generally increased in reading for fourth, eighth, and eleventh grades. District math norm-referenced measures have increased for grade four and slightly increased for grades eight, and eleven.

Additionally, district achievement data results for the Statewide Writing Assessment indicated that fourth, eighth, and eleventh graders made gains in the pre/post comparisons on the Statewide Writing Assessment. The writing data, fourth grade comparing data from 2002 to 2004, and eighth grade comparing data from 2003 to 2004, revealed strong gains at both grades. This may reflect the involvement of many school districts that used Six Trait Writing preceding the formal statewide writing assessment process. Students at the eighth and eleventh grades continued to show growth on their writing scores from 2004 to 2006. However, fourth-grade scores dropped slightly in 2005 to 2006 with an overall gain from 2002 to 2006 of 6.91%. Not only are scores for writing improving, but teachers involved in the scoring of statewide writing are also reaping great professional growth benefits from the scoring experience as revealed in a new sixth-year study. First, the scoring experience was perceived by a majority of the raters as a valuable professional development opportunity that contributes to their knowledge and understanding of the teaching and assessment of student writing. Secondly, raters' perceptions about the effectiveness of the scoring training were very positive across all grade level scoring groups. The scoring training, including the content, materials, training personnel, and procedures, was appropriately designed and delivered in such a way that raters felt they were adequately prepared to score the assessment.

Teachers reported learning about writing through training and some teachers were implementing interventions to improve writing for students. A writing intervention study revealed the overwhelming majority of students improved their writing when these three valuable components were used: explicit instruction in the writing process, explicit instruction in the use of a common language, and frequent references to a scoring rubric clarifying the definitions of effective writing. While it is unclear which of these components had the greatest impact on student performance, the results leave little doubt that the intervention as a whole revealed the codes of writing to the participants in a way that they internalized and applied with overwhelming success.

While teachers were learning and implementing interventions and students were increasing their performance, a belief system change was underway. Many educators were changing their thoughts about the impact of failure for students and have embraced the concept that “failure is not an option” (Blankstein, 2004) for their students. A change in belief systems was evident in many schools as one superintendent shared, “We have bought into the failure is not an option concept! We’ve done some things in terms of our grading policy, to enhance student learning. And we’re working with our teachers to change a mindset about how students are graded in this district. In other words, I have told the teachers in this district that have given students zeroes for not turning in homework . . . that is not an option.” This change in beliefs has also yielded a change of attitudes relative to special populations and their ability to learn as a result of the STARS. An elementary, special education teacher shared, “We go back again and we look at it, work on it again and try to come at it from a little bit different angle.” While the ELL and SPED student district average scores were not as strong as the district average student scores for all students, they were increasing in all instances. The reading scores for SPED students reflected strong positive change in all areas. ELL students at the fourth and eighth grade were somewhat stronger than for the eleventh-grade students. This trend could relate to the difficulty of ELL students mastering English as a second language when they are older. There have been changes in classroom teachers’ perceptions relative to special populations and their ability to learn as stated by this elementary SPED teacher, “The STARS data has really helped make us accountable for **all kids** not just for a select few.”

Special populations do create a unique set of needs that challenge classroom teachers daily. SPED and ELL teachers emphasized some of the similarities and differences. Mobility was a major issue for ELL that impacted record keeping and increased truancy. Communicating with second language students and parents, creating cultural understanding and providing appropriate accommodations made a positive difference for students in some communities.

The sixth-year study uncovered several differences in the perceptions between rural and non-rural educators. Turn around time in implementing changes for non-rural schools can be a major obstacle because they require more lead time due to the size of the districts as opposed to rural schools that were able to make changes much more quickly. Various activities related to the assessment process were completed in very different ways based on the size of districts. Many non-rural schools contracted for the development of specific assessment items and their assessment systems were viewed as much more sophisticated while rural schools lacked the funds to do such contracting and thus were in need of training within their own ESU or district for developing assessments. Rural leaders and teachers were more likely to indicate that professional development, ongoing assessment training, data retreats, identification of interventions and training in instructional strategies were provided by the ESU while non-rural schools tended to provide their own professional development and assessment literacy training. In rural schools there was evidence that some leaders new to their position and to the STARS Process were in need of training as they lacked an understanding for the value of STARS. Finally, non-rural school districts had the expertise of curriculum and assessment experts while rural schools were pooling funds for such a leader or lacked the funds to provide such services. Assessment coordinators in rural schools were found to wear many hats. Many non-rural leaders have become more concerned about developing their own district/school-based policies that are yielding a greater focus on the monitoring and reporting of student data and, in turn, upon increasing student achievement. Rural teachers continued to emphasize the importance of

resources, services, training, and record keeping as the key to successful implementation of STARS.

Tracking STARS assessment performance and record keeping continues to be a place for growth in the STARS Process. Technology is a viable tool for use in assessment as it provides increased speed in providing end results for students, aids teachers in identifying problems with assessment questions, and for changing assessments quickly that paper/pencil does not allow. Careful selection of the way that technology is used for assessment is critical as a great deal of classroom assessment needs to be performance based. Results from a technology survey administered this year indicated that STARS assessment leaders in Nebraska schools recognized the potential impact of technology on the assessment process, but at this point do not extensively make use of dedicated or locally developed software, although this may be occurring in larger districts or consortiums. It appeared that, among school personnel who work with the STARS assessment process, technology was viewed as having a positive impact. Schools have indicated that grant support for the implementation of technology and training of school staff would likely have a significant impact on the quality, usefulness, and value of the assessment process.

Overall educators continue to grow and act in the best interests of students, even new mandates will not impact the journey as illustrated by this superintendent, “It doesn’t matter what system they throw at us. The process that we have in place is a good process and we’ll be able to make it adapt and work for whatever the case is because its child centered. And as long as it remains child-centered, we’re going to be OK. It’s not perfect. It’s going to change. We know that. We expect that. We expect to improve every year. I mean . . . you know, that’s the expectation that we have for our kids; that’s the expectation that we have for ourselves.”

Of course, many Nebraska schools are far from the pinnacle of meeting the needs of every student. However, for the schools in the sixth-year study there has been growth—at varying rates and in various ways. Districts have made new strides toward putting the pieces together and creating cultures of continuous improvement. You can hear it in their voices as they tell their stories and envision it as they lead you on their journey. Learning is hard and it takes a long time, but as the pathway becomes more visible and educators meet with greater success, you still see even in those districts making the greatest growth in student achievement, the desire to revise their processes and interventions. However far along the pathway they might be, the growth and change is never really complete.

The path to improved student achievement has not been without bumps in the road, but Nebraska educators have committed themselves to an improved assessment process and learning for all students.

Blankstein, A. M. (2004). *Failure is not an option*. Thousand Oaks, CA: Corwin Press.

RECOMMENDATIONS

These recommendations remain from previous studies:

1. Continue to educate various constituencies about the nature and purposes of STARS. Focus special attention on local media and educators.
2. Continue to sponsor a “leaders of learning” academy for principals and superintendents.
3. Assist all districts in moving toward a K-12 focused standard, assessment and reporting process to ensure ownership at all grade levels.
4. Help districts to develop quality, seamless instruction especially for students not meeting the standards.
5. Design a system of tracking students of mobility to ensure that assessment data stays with the student.

New recommendations based on 2006-07 study:

6. Work collaboratively with ESUs to provide professional development for all teachers on instructional interventions for students (i.e., not mastering the standards, special populations).
7. Work collaboratively with ESUs to provide data training for leaders for informing decision making.
8. Provide leadership to help districts connect standards to appropriate grading and reporting systems.
9. Research the assessment literacy knowledge and skills provided by higher education for teachers entering the field.
10. Research student and parent perceptions and their involvement in the STARS process.
11. Work collaboratively with ESUs to provide professional development about assessment for new teachers.
12. Designate a person(s) that coordinates assessment, instruction, curriculum, and school improvement for each school district.
13. In the portfolio review process, provide more time and standardize schedule for reviewers to assess the quality of the assessments within each district.
14. Provide training that provides consistency for members of consortiums in the portfolio review process.
15. NDE needs to find ways to train and involve new educators on an annual basis as portfolio reviewers in order to build capacity in school districts for continuous growth.
16. Continue the state writing scoring process that enlists the participation of classroom teachers as a way to provide them with valuable training that relates positively to their classroom practices.
17. Continue grant support for implementing technology and training which will provide significant impact on quality, usefulness and value of the assessment process.
18. Work collaboratively with school districts to combine portfolio review and school improvement visits, etc. to minimize disruptions for school districts.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

Study 1: Nebraska-led Portfolio Peer Review Process



“Processes are people, not just the process.”

A non-rural research coordinator

Study 1: Nebraska-led Portfolio Peer Review Process

**Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln
Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American**

INTRODUCTION

The sixth-year primary study was a mixed-methods research design using both quantitative and qualitative data. The purpose of the study was to examine the new Nebraska-led Portfolio Peer Review Process and district educator perceptions of the technical quality of their district assessments according to the six quality assessment criteria (Plake & Impara, 2000):

- The assessments reflect the state/local standards.
- Students have the opportunity to learn.
- The assessments are free of bias and insensitive situations.
- The assessments are at the appropriate level.
- The assessments are reliably scored.
- The assessment mastery levels are appropriately set.

The portfolio review consisted of teams of two state-trained assessment experts that visited each school district within the state. The assessment reviewers read the previous district portfolio and the recommended changes for improvement from past years. The visiting team then went to their assigned district, reviewed the evidence of assessment quality using on the six quality assessment quality indicators, and provided formative feedback to the district. Two external

assessment experts located in each of the regional areas assisted the peer review teams by answering questions. A summative rating will appear in the State of the Schools Report in 2007.

RESEARCH DESIGN

This mixed-methods research study focused upon the Nebraska-led Assessment Portfolio Review Process. The K-12 District Internal Portfolio Review Team members were surveyed prior to the portfolio external review. Each district was asked to distribute up to ten surveys, as appropriate, to internal review team members. Of the 254 Nebraska school districts in 2007, 179 districts (70% of districts) returned surveys for the Comprehensive Evaluation Research Study. Of the 179 districts responding, surveys from 23 districts were disallowed because they returned surveys after their portfolio review. Of the 156 districts responding within the prescribed timeframe, 6% were non-rural and 94% were rural. For the purpose of this research, Nebraska public school districts were divided into two classifications, non-rural and rural, based on population characteristics unique to Nebraska. Non-rural districts were defined as metro-area districts in large and mid-sized cities, large towns, and the urban fringe. All other districts are classified as rural. Of the 254 public school districts in 2006-07, 5% were non-rural and 95% were rural.

Participants responded to the 49-item survey (Appendix D) using a five-point Likert scale for each item, with “1” representing “none of the time,” “2” “very little of the time,” “3” “some of the time,” “4” “most of the time,” “5” “all of the time.” The survey was structured to explore six themes: Alignment, Sufficiency, Clarity, Appropriateness, Scoring Procedures, and Summarizing the Review Process.

Secondly, open-ended interviews were conducted in two districts in each of four geographical areas identified by the Nebraska Department of Education (NDE) for the training of reviewers. Detailed views were collected about the Nebraska-led Portfolio Peer Review Process in the sample districts. The interview protocol for the District Internal Portfolio Review Team (Appendix E) was used to gather qualitative data. Eight sample districts were purposefully selected based on geographical area, district class, and district free and reduced lunch rate. Thirty-nine (39) individual interviews and one focus group interview (comprised of six educators) were conducted statewide during the 2006-07 school year. Three additional themes—Support for the Peer Review Process, Challenges, and New Learnings—emerged from the qualitative interviews.

Instruments

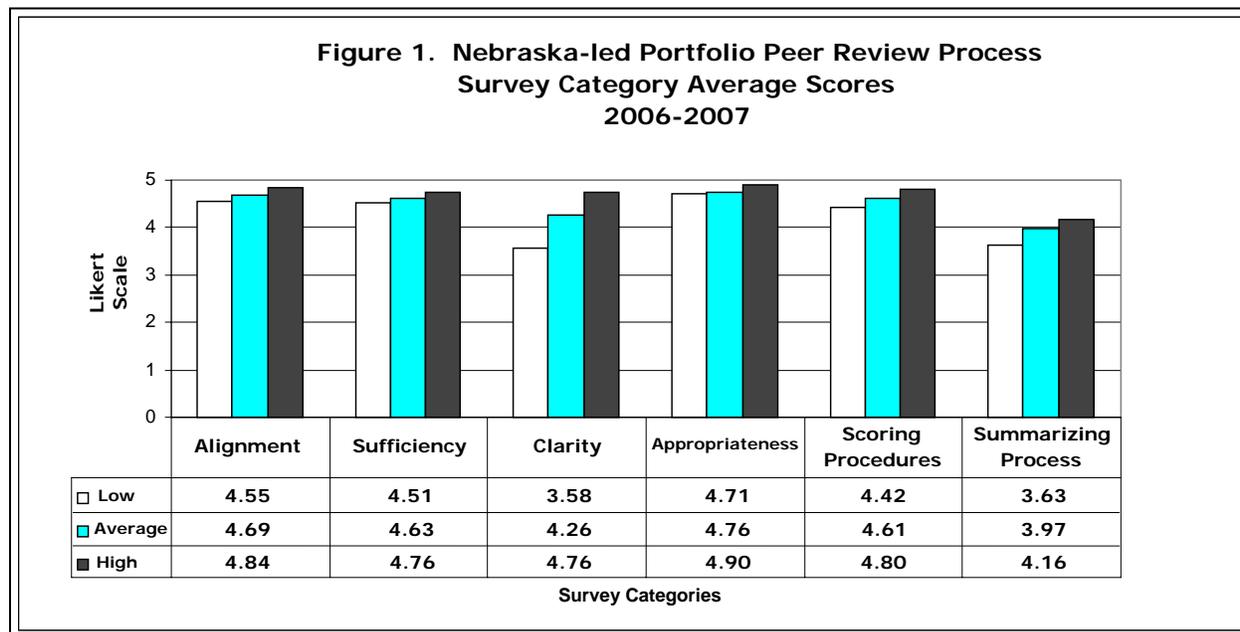
The STARS survey (Appendix D) was designed by the researchers to collect perceptions about the new Nebraska-led Portfolio Review Process and the six quality criteria. The survey examined the areas of (1) Alignment, (2) Sufficiency, (3) Clarity, (4) Appropriateness, (5) Scoring Procedures, and (6) Summarizing the Review Process. Participants responded to the 49-item survey on a five-point Likert scale for each item, with “1” representing “none of the time,” “2” “very little of the time,” “3” “some of the time,” “4” “most of the time,” and “5” “all of the time.” Analysis of variance was used to compare mean scores of the survey data.

The STARS Research Interview Protocol (Appendix E) consisted of demographic information about participants and ten questions for the selected members of the Internal Portfolio Review Team. These questions targeted the participants’ perceptions of each of the six quality criteria, their preparation and initial thoughts of the review process, and any new learnings based on the

process. Probes were identified for interviewers to use with each question. Interviewers were provided a STARS Interview Manual and received training to conduct the interviews.

RESULTS

Six categories were identified from the Nebraska-led Portfolio Review Process Survey conducted prior to the visit of the experts. They were rated by participants on a “1” to “5” Likert scale with “5” being the highest. Noted in the graph is the mean, lowest and highest score per category.



Theme 1: Alignment (Survey Questions 1-8)

In the area of Alignment for all respondents, responses ranged from 4.55 to 4.84 with an average of 4.69 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Alignment was not significant for leaders and teachers.

The strongest item rated by all respondents within the Alignment category was “districts involved staff in the alignment of assessments to standards (4.84).” Respondents also indicated that “districts’ assessment items or tasks reflected a match to the appropriate standards (4.74)” and “there is a documentation process for alignment of assessments to standards (4.74).”

In the area of Alignment for all respondents, the lowest rated area was “districts had a list of specifications mapping the assessment items to the standards in order to show which items assessed which standards (4.55).”

The survey reliability statistic (Cronbach’s Alpha) for the Alignment section of the survey was 0.776.

Alignment Discussion

The strongest perception from all respondents indicated that “district involved staff in the alignment of the assessments to standards.” Middle school teachers rated this question significantly lower than all other groups of teachers (p=.036).

- A female rural fourth grade teacher stated how staff was involved with aligning standards, “I know that they went through and unpacked the standard, looked at all the benchmarks and then tried to locate some of the teaching materials we were using already and used those to develop the assessment that we give now.”

Females tended to rate this question significantly higher than males (p=.051).

- A female rural Assessment Coordinator shared their process of matching assessments to standards, “We have a pretty thorough process in terms of matching assessments to standards. We go through the unpacking of the standards process where we actually talk about what skills and knowledge do students need to have in order to be considered a proficient student in regard to the standards. Those are long, intense conversations because once we determine what skills and knowledge are necessary for the proficient level, then we move on to the beginning, progressing, and advanced levels. ...Probably the most important piece in that alignment is getting those assessments in front of various eyes.”

The item of “district involved staff in alignment of assessments to standards” was not significant for leaders.

Another item rated high by respondents indicated that “district assessment items/tasks reflect a match to the appropriate standards.” This question was significant based on years of experience in education. Teachers with more years of experience in education tended to rate this item higher than teachers with less years experience.

- As indicated by this superintendent, alignment was the very foundation of the building of the STARS process, “I think that the alignment is number one for a reason, I mean you really have to start with that foundation, and so the process was very healthy for us in that we did make some modifications with our curriculum of what we taught and when we taught it by going through that process.”

Survey respondents were also strong in their perception that “there is a documentation process for alignment of assessments to standards.”

- This is illustrated by a female rural assessment coordinator that stated, “Years ago, when we first started this process, long before the state mandated assessment, we designed a system that identified where content was taught, matched content to standards, and determined an assessment system that allowed us to assess at the point of instruction. Through the years we have continued to review that content and alignment to make certain things are still on track. We’ve made lots of adjustments.”

The survey item indicating that “district had assessment items reviewed by external personnel” was found to be significant for leaders (p=.026) by years of experience in education. Leaders with more years of experience rated this question higher while those with fewer years of experience rated it lower.

In the area of Alignment for all respondents, the area with the lowest score on the survey was “district has a list of specifications for mapping the assessment items to the standards in order to show which items assessed which standards.” Evidence from this study supported

this conclusion as there were only two references to a table of specifications by non-rural coordinators.

- A non-rural research coordinator stated, “We begin our assessment process by developing a table of specifications. That’s a grade level table of specifications generally written by teachers”
- A non-rural assessment coordinator shared, “We send out our table of specifications to the item writers and that’s what they begin their work from. They understand building a variety of types of items.”

In the Alignment discussion during interviews, educators indicated that alignment was the impetus for change in schools as teachers and administrators interacted with a great deal of discussion regarding how “curriculum and assessments matched the standards.”

- A rural superintendent shared, “You walk into any classroom, in any one of these sites, and you’re going to see standards and assessment curriculum books out, laying out, laying open, it’s not just because its spring and we have to assess. They really use them.”
- A female rural assessment coordinator shared, “I think the one thing that districts may be overlooking is changes in curriculum. Once they establish an assessment schedule, they live with it rather than realizing that if I change the curriculum, I’ve now changed delivery of the content, possibly have changed where content is delivered. In that case, my assessments, either my assessment schedule or the way I assess have to change.”

While there has been a great deal of matching of curriculum to standards and standards to assessments, there is some fear expressed by leaders that the curriculum could be easily narrowed by this process.

- A rural superintendent shared, “Programs drive a lot of your curriculum and the thing that we’re trying at this point is to just make sure we haven’t narrowed the curriculum so much that we are just only on the standards because we know those are the minimum.”

Alignment Recommendations

During the interviews, participants were asked if they had recommendations to make to other school districts regarding each of the survey category areas.

- For Alignment, a K-12 articulated curriculum leads to success as illustrated by this female rural assessment coordinator, “Our teachers know our curriculum. I think that’s the key I guess my recommendation is that you need to have a K-12 articulated curriculum.”
- A rural assessment coordinator shared the importance of unpacking the standards so that there was a clear understanding of skills and knowledge for determining proficiency. She stated, “It’s definitely the unpacking of the standards discussion. They have to have a very clear understanding of what skills and knowledge they as a local district believe are essential in determining proficiency to a standard.”
- A fourth-grade teacher shared that constant monitoring is necessary for the assessments to be the best that they can be. She stated, “There have been several changes. Each year when the teachers notice something, then they change that to make it better.”
- Several of the interview participants emphasized that involving special population teachers is critical to the alignment process as stated by this female rural assessment coordinator, “SPED and ELL can really bring valuable things to those discussions. When

you always include them, they carry out what is the intent of things, and the alignment, and the instruction is so much better.”

- A rural superintendent emphasized the importance of involving all teachers in professional learning discussions as she shared, “I think the teachers own those groups [grade level groups] now, and I really think they’re as close to PLC’s as what we’re going to get with the professional learning community concept in that they are by and large curriculum scope and sequence, and standards and assessment based. They look at students’ work and they use student work and data.”
- An assessment coordinator emphasized that training and support to understand the table of specifications and its use is needed, “We use infinite campus as the reporting system and it’s a very sophisticated, detailed program. I spend a great deal of my time training teachers and principals on how to pull that data on individuals, on groups of kids, by teams of teachers in buildings, and then drilling down to individual strand level data on kids related to these assessments. We have a link directly to the table of specifications, so they can hold the table of specs next to the pattern of performance by their kids and can in real time adjust their instruction.”

Theme 2: Sufficiency (Survey Questions 9-13)

In the area of Sufficiency for all respondents, responses ranged from 4.51 to 4.76 with an average of 4.63 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total category of Sufficiency was not significant for leaders or teachers.

The strongest perception from all respondents was for the items “district reviewed the assessment items/tasks for sufficiency results (4.76)” and “assessment items/tasks are distributed across all performance levels (4.71).”

In the area of Sufficiency for all respondents, the areas with the lowest score were “assessment items/tasks use a variety of appropriate formats (4.51)” and “assessment items/tasks include higher order thinking skills (4.53).”

The survey reliability statistic (Cronbach’s Alpha) for the Sufficiency section of the survey was 0.815.

Sufficiency Discussion

The strongest perception from all respondents indicated that “district reviewed the assessment items/tasks for sufficiency results.” Male leaders (p=.004) and middle school principals (.007) tended to rate this item higher than did their counterparts.

- A rural assessment coordinator stated, “Sufficiency’s been an interesting piece. We all wrote assessments with no understanding of sufficiency. So that’s been a learning opportunity . . . teachers are a lot more assessment literate than they used to be.”

Another Sufficiency item on the survey that participants rated strongly was that “assessment items/tasks are distributed across all performance levels.”

- This rural superintendent shared, “The sufficiency piece is one of those statistical components that I learned more about as we went through sufficiency. I certainly did not have an adequate knowledge base about making a good assessment prior to the whole STARS piece, as we tended to have most of our questions in the very lowest level or in

the very highest level, and did not have a good distribution among beginning, progressing, proficient, and advanced.”

- A rural fourth-grade teacher added, “They [teachers] have gone through and designated which questions are at the beginning level based on what a beginning student can do. And then they wrote the rubric for those so they were able to determine what a beginning student should be able to do and did this all the way to the advanced level.”

In the area of Sufficiency for all respondents, the area with the lowest score was “assessment items/tasks use a variety of appropriate formats.” Leaders with more years of experience tended to rate the item stronger than leaders with less years of experience.

- A non-rural assessment coordinator shared, “We’ll obviously try to pay attention to the process of doing vertical articulation with our curriculum, so that is going to be the key to appropriately placing where the items are. We’re not completely at that point yet, so what we’ve done is to make sure that at least in the grades that we are accountable to report, that we’re making sure that those items are appropriate.”

Another Sufficiency survey item that scored low was “assessment items/tasks include higher order thinking skills.” This low rating was supported by the fact that large non-rural districts sometimes rely on outside agencies to develop assessment items for sufficiency.

- A non-rural assessment coordinator stated, “We send out our table of specifications to the item writers and that’s what they begin their work from. They understand building a variety of types of items. All those factors go into his or her item development but the sufficiency part, we know that there are certain items that have certain values. That’s why it’s so important for us to have BUROS involved, because they will then take us through the process of insuring that we have enough items that accurately measure students at their appropriate performance level.”
- A rural superintendent supported the need for outside assistance when looking at sufficiency by stating, “Maybe that’s why it was so difficult for us, now that I think about it [Sufficiency]. Seek outside help, unless you’re a really large district. I think if you’re small you need that.”
- Another rural superintendent indicated agreement as to the level of complexity required to develop adequate numbers of questions at each level, “I think that’s a pretty tall order to ask an outside reviewer to come in and look at the specificity of our questions throughout to see if they’re really beginning, progressing”
- On the other hand, this rural assessment coordinator indicated when developing questions, “They [teachers] have the Blooms Taxonomy list in front of them and the level definitions, and they start to look at this as the content we want. So they have that when they’re writing assessments.”

Sufficiency Recommendations

During the interviews, participants were asked if they had recommendations to make to other school districts regarding each of the survey category areas.

- For the category of Sufficiency, a rural assessment coordinator shared the value of student feedback, “I think that the openness to how students react to assessment items is an extremely important point here. It is also a fine line. Teachers need to have the best interests of their students’ learning as much as they can and not what would be a question they would get correct, because those are two different things.”

- Creating items across different performance levels is not an easy task as shared by this rural assessment coordinator, “Well, I think that you would need to make sure that there is somewhat a balance of items across all levels. In other words, I don’t want to see one beginning, three progressing, six proficient, and five advanced.”
- Teachers gained better insight about the importance of questions by performance levels as noted by this rural assessment coordinator, “One of the things that they [External Team] recommended, which I already knew, but the teachers that were part of this review had to hear from someone besides myself, was we have to have better performance level definitions.”
- Some districts were concerned about the discrepancy between norm-referenced and criterion-referenced results as indicated by this rural superintendent, “What we had was a serious mismatch between CRTs and NRTs. That was a bad thing because our kids were glowing with STARS.”
- A non-rural assessment coordinator shared, “I think that the issue of sufficiency on locally developed assessments needs to be addressed by people other than your teachers.... I’m not saying that teachers are not capable of developing those types of items, but it’s helpful to have some folks from the outside to take a look at it.”

Theme 3: Clarity (Survey Questions 14-21)

In the area of Clarity for all respondents, responses ranged from 3.58 to 4.76 with an average of 4.26 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total category of Clarity was not significant for leaders and teachers.

The strongest perception from all respondents indicated that “assessment directions for students are clear (4.76).”

In the area of Clarity for all respondents, the two items that scored the lowest were “district provides parents with reports that gave an appropriate explanation of assessment results (3.58)” and “district sends individual reports to parents each school year (3.74).”

The survey reliability statistic (Cronbach’s Alpha) for the Clarity section of the survey was 0.683.

Clarity Discussion

The strongest perception from all respondents indicated that “assessment directions for students are clear.” Leaders with more years of experience also rated assessment directions for students higher than other leaders (p=.014).

- A fourth-grade teacher shared, “We have team meetings that we discuss those [directions], usually we would find something that might not have been quite clear to the students and they may have questioned it. And we would mark that down and bring it back to the committee again. The following summer they would improve upon that, and the same way with teacher directions.”
- Another fourth-grade teacher emphasized, “I think that really comes from working with the group and giving the assessments and then each year reading them over and making sure that they still make sense to you and to the student. That’s kind of what we’ve done when we get together. We talk about that.”

Leaders recognized that directions for students and teachers were critical to the success of administering assessments.

- Those involved in online assessment administration also indicated that directions need to be clear for students as shared by this rural female assessment coordinator, “So the directions in the online system are on the screen when a teacher looks at the test. But contained in there are teacher directions so that teachers know these are the directions you are to use and you are bound by these directions.”
- In non-rural schools it is important that all persons administering assessments receive specific instructions as shared by this non-rural research coordinator shared, “We do have a set of directions that are sent out to the buildings, one per person administering the assessment including those that are giving the special needs students’ assessments related to their IEPs. . . . Each year when we print assessments for that year, we go back and re-evaluate whether we have all of the information that we need. We’re continually updating or making minor changes from things that we observe in the schools.”

The two items within Clarity with the lowest ratings were “district provides parents with reports that give an appropriate explanation of assessment results” and “district sends individual reports to parents each school year.”

- This is evidenced by this female rural assessment coordinator’s statement indicating that they really don’t provide evidence to parents when she stated, “Not really. I just think it’s important to notify parents. I know as a parent, I would have liked to have that information on my children.”
- A female rural assessment coordinator indicated that parents sometimes get reports at varying levels within the school district, “Parents get reports at the end of eighth grade. So usually at the ninth-grade parent teacher conference, there is a table set up and there are graphed reports that show student performance on all of the eighth grade standards. We do the same thing for senior parents because by then juniors have finished their testing.”
- A female rural assessment coordinator shared, “We have a letter that goes home in report cards at the end of the year that shows performance for all standards.”
- Some districts are constantly adding and refining their reporting to parents as stated by this female rural assessment coordinator, “Well this is our first year of actually reporting this. In the past we did it at the grade level, parent-teacher conferences. The teachers would share that information with them As far as actually having something to give them, some teachers made their own, but it wasn’t happening district-wide like it is now.”

Although these two lowest items were rated the lowest on the survey, elementary teachers rated both items, “district sends individual reports to parents each school year (p=.022)” and “district provides parents with reports that give an appropriate explanation of assessment results (p=.008)” significantly higher than all other teacher categories.

- A female fourth-grade teacher stated, “They’ve been reporting the writing for quite a few years, the writing scores that come back. We input our scores into the computer and now have a way of printing those off for the parents.”
- A rural superintendent shared, “Report cards are translated. Parent meetings are translated. I would say there’s more of a personal involvement prior to giving the assessments or trying to explain to the families what this was about and why we’re doing it. . . . We might make a phone call, we might handwrite a note, we might talk to the kids

more about it so that they can explain this to their parents, and they might translate the information.”

In the interviews, teachers and school leaders addressed student feedback thoroughly when examining different audiences that needed to have information about assessment performance.

- A male non-rural assessment coordinator shared, “We report assessment results. There are obviously different audiences so you don’t report the same to each of the audiences.”
- A rural middle school teacher stated, “Usually, our school district does have a report card that is really pretty cool, and they send that out to students that have taken the assessments, to let parents and students know how they’ve done on assessment scores. I actually like to talk to the kids about it after they get done.”
- A female rural assessment coordinator stated, “Teachers pass the assessments back and let students see their results on those. . . . But they pass those back and go over them which is a very important part of the process, I think.

In some non-rural districts, educators addressed individual learning plans that were developed for students not meeting the standards.

- This non-rural assessment coordinator shared, “Now any student, who is not progressing, has an individual learner plan in our district. It should not come as a surprise to a parent when we get to the next grade level of testing that this child has needed additional assistance. Every building has a pyramid of intervention and that kid is plugged into that pyramid of intervention, so the communication is pretty intense. And the parents know all the way through.”

A number of districts mentioned that STARS assessments become a part of the grades assigned to students at the end of a reporting period. Therefore, students are concerned with success on assessments.

- This female middle school teacher shared, “Well, for the students, we incorporate this as part of their grades. So they get it back. They see it on their grade. We don’t feel that children respond too well where they don’t understand the purpose.”
- In the online environment students are provided with immediate feedback as shared by this female rural assessment coordinator, “Students get assessment results the moment they hit the submit button. On the screen, they get a blue screen that will immediately identify their proficiency level for them. And we joke about telling them about reread, check your answers before you click submit. Many of them click submit and then cross their fingers and squeeze their eyes shut and go, ‘oh please, please, please.’ And when the score comes up, if it’s proficient or advanced, they jump out of their seats, ‘hurray, hurray.’ So, they get instant feedback.”

Clarity Recommendations

During the interviews, participants were asked if they had recommendations to make to other school districts regarding each of the survey category areas.

- Due to the mountain of paperwork generated when administering assessments and analyzing assessment data in a large school district, many districts have begun conversations about using an online format for their paper-pencil assessments as shared by a non-rural assessment coordinator, “We’re going to pilot the online assessment in the fall. We handle thousands and thousands of pieces of paper. I’ve got data analysts who

are spending hours scanning forms, and that's probably not the best use of their time. So for bigger districts, I would maybe take a look at how you can simplify the process for large populations."

- A rural superintendent stated, "Having a site coordinator to actually get the test certificate, bring it up and test those kids has been huge. It's like a proctor, if you will, and we have a set of protocol that they use when they administer a test. So I'm very comfortable that the administration from the teachers' point of view is pretty consistent."

The area of Clarity was defined by providing assessment directions for both students and teachers to assure assessments were administered uniformly across the district. This was evident in a number of ways. Some districts have continually refined and improved their directions for students and teachers. There are many different types of reports being generated for various audiences, the most important were those prepared for parents and students. Therefore, a great deal of evidence was available to the interviewers about the assessment directions used by school districts as well as information provided about how districts report. Although interviewers asked for recommendations about the area of Clarity, very little new information was collected as there was limited discussion in this area.

Theme 4: Appropriateness (Survey Questions 22-28)

In the area of Appropriateness for all respondents, responses ranged from 4.71 to 4.90 with an average of 4.76 on the Likert scale with "1" representing "none of the time" and "5" representing "all of the time." The total category of Appropriateness was not significant for leaders and teachers.

The strongest perception from all respondents indicated "assessments were screened for fairness, bias, and sensitivity (4.90)."

In the area of Appropriateness for all respondents, the lowest item was the "assessments demonstrate an increase of expectation from one grade level to the next (4.71);" "assessments are appropriate for the assessed grade level (4.72);" and "assessments indicated our expectations for the students (4.72)."

The survey reliability statistic (Cronbach's Alpha) for the Appropriateness section of the survey was 0.862.

Appropriateness Discussion

In the area of Appropriateness for all respondents, the lowest rated item was "assessments demonstrate an increase of expectation from one grade level to the next (4.71)."

- A rural female assessment coordinator shared, "But we did a review. I don't feel we did an adequate enough one because we didn't have adjoining grade levels. And I think that's a really important part of that"
- Another rural female assessment coordinator shared, "We lay the assessments out, particularly what we call the NCLB pieces, the third, fifth, sixth, seventh. And so we honestly lay the assessments down on a table side by side. Let's read the third-grade one because 3.1.1, 4.1.1, 5.1.1 all measure the same thing, main idea and supporting details. So . . . do we find increasing rigor as you go through the grades? Are we asking more of kiddos as they go through? That's one of the tools we use every summer in the summer workshops to look at developmental appropriateness."

In the area of Appropriateness, the item “assessments are appropriate for the assessed grade level (4.72)” was rated similar to other items in this category.

- A female rural assessment coordinator shared, “You need a mixture of teachers. The reason you need that is because there’s a wide variety of philosophies. And some people’s philosophy is, ‘Oh, but they’re just fifth graders. They probably can’t, a 20 minute test, you know.’ And then you’ve got the other end that is like, ‘They’ll be fine. Just give it to them.’ And those voices come together, that’s the great part about this process . . . they reach an agreement about appropriateness of level so that things aren’t too easy, and they’re not too hard.”
- A female rural assessment coordinator shared, “We work as cluster grade level teams. So when we develop an assessment or revise an assessment, it’s not just fourth grade teachers working to do that. It’s a second, a third, there’s representation from each grade level there. I think that ensures that the appropriate level piece is there.”

In the area of Appropriateness, the item, “assessments indicate our expectations for the students (4.72)” was rated approximately the same.

- A female rural assessment coordinator shared, “We did it with readability scores. However, we have learned that readability can fluctuate so much depending upon the scale you use to determine readability. We primarily use professional judgment. But every year teachers review passages for professional judgment determination about being developmentally appropriate. We talk about how much inference, how many scientific terms, those are all components of readability that don’t always show up on a readability scale. So professional judgment is our primary piece.”

Teachers with more years of experience with the portfolio rated these items, “our assessments are appropriate for the assessed grade level (p=.039),” and “our assessments indicate our expectations for the students (p=.050),” higher than teachers with less years of portfolio experience.

- A female rural assessment coordinator stated, “I do a bias review every year for teachers. I’ve always told the teachers if I’m not here, you guys need to know how to do this. It needs to be separate from me. And so they have a handbook, it’s called the assessment and accountability handbook that lays out all the six quality criteria.”

Elementary leaders (p=.032), leaders with more years of portfolio experience (p=.008), and female teachers (p=.022) rated the item, “our assessment plan provides for appropriate accommodations where necessary,” higher than their counterparts.”

- A female rural assessment coordinator shared, “Any SPED student that has an IEP, the SPED teacher and the regular ed teacher confer and determine if any accommodations are needed. If they are allowed for in the IEP, the expectation is the student will receive those.
- Another rural superintendent shared, “When we have a conversation, we make sure that we have someone speaking in their native language.”
- A female rural assessment coordinator shared, “First and foremost, teachers need to be informed of what accommodations are allowed. And that has been a process and that’s through our IEP writing. Our SPED Director has gotten involved with this and, we have come up with kind of a blanket statement in all Individual Education Plans (IEP) that just

allows for accommodations that need to be administered at the discretion of the classroom and special education teacher.”

In the area of Appropriateness, the strongest perception from all respondents indicated, “assessments were screened for fairness, bias, and sensitivity (4.90).”

- A rural assessment coordinator shared, “It’s almost an immersion in culture, somewhat, to understand, to be able to understand what might be bias.”
- A rural assessment coordinator shared, “A bias review. When the assessments are developed, they go through a review that’s extraneous of the development team and those people are trained in terms of what is bias. It’s a universal definition that we as a district adopted.”
- A rural assessment coordinator shared, “But it’s something I learned this year and maybe you were at one of these where they talk about elevator and what an elevator means to kids out here in western Nebraska as opposed to kids in Lincoln.”
- A non-rural research coordinator shared, “There’s a lot of things that we don’t think about that are biased. You have to look at items that a student would not be able to respond to the best of their ability. That is not just content oriented, but the way the item is stated, vocabulary or anything that might affect them negatively. You can’t say that you did this once before many years ago and you are good.”

Appropriateness Recommendations

During the interviews participants were asked if they had recommendations to make to other school districts regarding each of the survey category areas.

- A male non-rural assessment coordinator described, “We take advantage of every allowable accommodation that we can to try to help the kids and the ELL kids. Safe harbor certainly helps, confidence intervals certainly help, I guess my recommendation would be to know what those things are and to use them in the best way that you can to help your kids.”
- A female middle school teacher provided the following recommendation to the state about special populations, “Because I know that there are some that even if it says in their IEP they can have the test read to them, you can’t And that would be something I think would be helpful to hear from the Nebraska Department of Education, these assessments you can give accommodations on, these you can’t.”

Theme 5: Scoring Procedures (Survey Questions 29-42)

In the area of Scoring Procedures for all respondents, responses ranged from 4.42 to 4.80 with an average of 4.61 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total category of Scoring Procedures was not significant for leaders and teachers.

The strongest perception from all respondents indicated that “participation rates are documented (4.80)” and “assessments have established scoring guidelines and directions (4.77).”

In the area of Scoring Procedures for all respondents the lowest item was “district provides training for those administering the assessments (4.42).”

The survey reliability statistic (Cronbach's Alpha) for the Scoring Procedures section of the survey was 0.918.

Scoring Procedures Discussion

The strongest perception from all respondents indicated that “participation rates are documented.” More experienced leaders (p=.033) and female leaders (p=.050) rated this item higher than their counterparts. Middle school teachers, however, rated this item lower than all other teacher groups (p=.021).

- A rural superintendent stated, “Well I don't know about the other districts, I would like to see some recognition that some of these statistical processes, again, are holding us back from recognizing students, that all students can achieve. So the idea of the 70% reliability ranking, I mean I understand reliability but, I don't know enough about statistical analysis to know what the solution is, other than why can't 100% of our students answer the question correctly and that be ok, and not be a bad question.”

Respondents also indicated strong perceptions for “assessments have established scoring guidelines and directions.”

- A rural superintendent in a large district said, “We have gone through the development of performance level descriptors. We also, then, have a review committee of people who did not write the assessments who review those or the work of teachers and sign off, on them or basically point out any problems that we have.”

In the area of Scoring Procedures for all respondents, the lowest rated item was “district provides training for those administering assessments (p=4.42)” as evidenced by this lack of understanding about common scoring procedures.

- A rural middle school teacher said, “I must not know anything about scoring procedures. I sit down with a key and I grade them. That's what I do. I'm the person who's responsible for scoring. That's what I know about scoring procedures.”

Teachers (p=.015) and leaders (p=.027) with more years of experience rated, “our subjectively scored assessments have inter-rater reliability and decision consistency methods that are within acceptable ranges,” higher than teachers and leaders with less experience.

- A female rural assessment coordinator shared, “We have professional learning communities in place. Those are wonderful avenues for me to get in and sit down and talk with groups of teachers. One thing I absolutely insist on is that scoring takes place collaboratively, especially on anything that has subjectivity involved. So we've gone to insisting any time there's possible subjectivity doing collaborative scoring.”
- A female rural assessment coordinator shared, “We're finding out throughout this year and also yesterday in our peer review. We do have a lot of subjective tests and that we need to come up with a plan They either need to be double scored or decision consistency.”
- A rural superintendent in a large district said, “Then of course, we're running validity and reliability statistics on them at the end to see where we are as far as that coefficient. We know that over time, we have improved our statistical measure of validity and reliability considerably. We used to be kind of tickled when we met the minimum for all

assessments. We're doing the right things and I think the statistical measures are saying that we're doing the right things."

Male leaders tended to rate the item, "district has monitoring procedures in place for inclusion, standardization, and security (p=.030)," differently from female leaders as reflected by the lack of male responses.

- A female rural assessment coordinator said, "Before we joined online, we joined with another district. Their assessment coordinator and I would compile our data to try to run KR20s to determine reliability. Over a period of time, we've gotten a whole lot smarter about reliability and I think other reliability methods have been approved that were not in place when we first started. The person instructing knows the kids. So, that's where the reliability factor needs to come from. So, we've learned to shuffle the papers where they need to be, to the appropriate staff, to predict how a student will perform."
- A female non-rural research coordinator said, "Since the majority of our assessments are, objective, we use Cronbach's alpha. We have over 1500 students in each sample. With a large sample we use that reliability information as the assessment is developed. We run it on our pre-pilot as well as our pilot and if you take our pre-pilot data and our pilot data and our actual assessment data over years, you're going to see very little variance between any of those. So, you know, that directs us."
- A female non-rural research coordinator said, "When it comes to our performance assessments, in other words the speaking things, then we use inter-rater reliability."

Scoring Procedures Recommendation

During the interviews participants were asked if they had recommendations to make to other school districts regarding each of the survey category areas.

- A non-rural assessment coordinator stated, "Well I think that it's very important to make sure that your tests are defensible obviously to give students an opportunity to learn the material and then test constructs that appropriately measure whether or not the student has learned the material."
- A female rural assessment coordinator stated, "You know, this is such a huge discussion among curriculum directors because as your performance on assessments gets better, which is our goal obviously, oftentimes reliability calculations fall. So, that piece is really a struggle."
- A female rural assessment coordinator stated, "His reliability calculations are going down because his performance is going up. And we've had long discussions about it and it's like what do you do? I mean, it's a catch-22 cause you've got to meet quality criterion five. But what do you do to keep those reliability levels high?"

Theme 6: Summarizing the Review Process (Survey Questions 43-49)

In the area of Summarizing the Review Process for all respondents, responses ranged from 3.63 to 4.16 with an average of 3.97 on the Likert scale with "1" representing "none of the time" and "5" representing "all of the time." The total category of Summarizing the Review Process was not significant for leaders, but was significant for years of portfolio experience for teachers (p=.000).

The strongest perception from all respondents indicated “I have the necessary information to prepare the district assessment portfolio (p=4.16)” and the survey respondents indicated “I have had adequate help in preparing the assessment portfolio (p=4.14).”

In the area of Summarizing the Review Process for all respondents, the lowest rated item was, “Compensation is provided to prepare the district assessment portfolio when completed outside of the regular school day (p=3.63).”

The survey reliability statistic (Cronbach’s Alpha) for the summarizing the review process section of the survey was 0.852.

Summarizing the Review Process Discussion

The strongest perception from all respondents, and from leaders with more years of portfolio experience (p=.049), indicated they had “adequate help in preparing the district assessment portfolio.”

- A rural superintendent said, “The leadership component’s huge, you know, I just think any district that doesn’t, no matter how big you are it doesn’t have someone at the helm that’s knowledgeable and that respects and values the system you’re going to struggle. If you’re an administrator that doesn’t like it and doesn’t want to know about it then just make sure that you hire somebody who does and give them some resources to do it. It’s going to take some time. It’s going to take resources.”
- A female rural middle school teacher said, “We spent a lot of time working with our assessment coordinator ahead of time. You know she talked with us a little bit about what might be expected. We met with her on a few different occasions to prepare materials and make sure that we were all on the same page. Before that, we went to our ESU and got some of our materials together. And actually that was the first time I had heard anything about a portfolio review.”

In the area of Summarizing the Review Process for all respondents, the lowest rated item was “Compensation is provided to prepare the district assessment portfolio when completed outside of the regular school day.” Teachers with more years of portfolio experience rated this item higher than did teachers with less years of portfolio experience (p=.004).

Teachers with more years of portfolio experience rated the item, “I feel prepared to present my district portfolio to my peer reviewers (p=.001),” higher than did those with less years of portfolio experience.

- A female rural middle school teacher with many years of experience shared, “We spent a lot of time working with our assessment coordinator ahead of time. You know she talked with us a little bit about what might be expected. We met with her on a few different occasions to prepare materials and make sure that we were all on the same page. Before that, we went to our ESU and got some of our materials together. And actually that was the first time I had heard anything about a portfolio review.”
- A rural female assessment coordinator said, “Some people went into the day with a lot of trepidation. I had the advantage of knowing what I thought the day would look like even though we spent more time with the reviewers than I had anticipated, I at least knew the people and knew the process. So our teachers, even though we had told them a dozen times what the day would look like, still felt like it was a test going in. However, after a

short period of time, I think that everybody relaxed and people were very open, very honest.”

Teachers with more years of portfolio experience (p=.006) and leaders with more years of experience (p=.041) rated, “I was provided time within the teaching day to prepare the district assessment portfolio,” higher than did their counterparts with less experience. Experience was an important element to the preparation of the assessment portfolio and teachers and leaders with experience appeared to be in positions in their district that provided them time to work on the assessment portfolio as evidenced by the ratings.

Teachers with more years of portfolio experience rated this item, “I have the necessary information to prepare the district assessment portfolio,” higher than did teachers with less years of experience (p=.003).

- A female rural elementary teacher said, “Our curriculum director takes care of the portfolio. But we also, you know, our curriculum director has helped us each step of the way every time we come into committee meetings in the summer.”
- A female rural middle school teacher shared, “I think as much as can be expected, truthfully, our assessment coordinator bore the brunt of the portfolio review in regards to preparation. Obviously, you know, we constructed the assessments and all that other good stuff beforehand but as far as the construction of the portfolio she really took on the responsibility for that.”

Elementary teachers rated the categories Alignment (p=.025) and Clarity (p=.006) stronger than did other teacher groups.

- A female rural assessment coordinator stated, “We knew, we teach it here, we test it here. This year we were a whole lot better at our documentation. So, I think it bumped it up a notch.”

Summarizing the Review Process Recommendations

During the interviews, participants were asked if they had recommendations to make to other school districts regarding each of the survey category areas. Overall, the participants felt that the Department of Education had shown a tremendous amount of faith in the integrity of the process based on what the educators in the field provided in the way of feedback. The following recommendations come from those educators that are in "the trenches" and know first hand what may be used to strengthen the process:

- A female rural elementary teacher said, “I don’t think the assessments change often enough that you would need that (Peer Review) every year. But maybe a check every three years, because our assessments haven’t changed a lot.”
- A female rural assessment coordinator stated, “One thing that was mentioned yesterday, and I’m sure it was shared with you in your last interview, is to somehow get the portfolios to the reviewers a day or two ahead of time so that they can actually look through it before they come in.”
- A rural superintendent said, “I don’t know about strengthening, and I don’t want to say streamlining, but any time that we can combine activities whether it’s through your school improvement or North Central visitation or incorporate it to an existing opportunity would be positive.”

- A female rural assessment coordinator said, “I think they need to spend more time on assessing the assessments part. Give us more feedback. We really didn’t get a lot of feedback on our assessing the assessments.”
- A rural superintendent stated, “I have to be honest with you, I think you could have saved some time and some money if you would have brought your portfolios and a group to the service unit and met in a central location. I know it might have depersonalized it. . . .”
- A rural superintendent said, “I don’t want to be in a room looking at books. . . . I just kind of think, if there could be a student component to it, it would have been richer. I really feel like we left them out. Our kids were so excited . . . ‘The state department’s coming, the state department’s coming. Well, what did they do in that room all day, we didn’t see.’ I think the exit report should be set up to where it truly is just like a school improvement exit report, where everybody gets to celebrate”

The accountability process in Nebraska focuses on students in the classroom and their academic improvement. How it is done in Nebraska is a local issue. Nebraska's new Peer Review Process was a major undertaking and educators across the state took up the challenge to "get it done and done well." It was not perfect and everyone learned in the process but the end results seem to indicate that it is a viable means of evaluating the assessment process. Nebraska educators across the state indicated confidence in the STARS process as summed up by this rural superintendent, “I would only urge a district of any size to acknowledge that standards and assessment work regardless of what kind of test you have. It’s on-going and it’s everyone’s charge and responsibility to get it accomplished”

Theme 7: Support for the Peer Review Process

While there were challenges that faced the new process of reviewing the portfolios, educators across the state indicated strong support for the new Portfolio Peer Review Process. Many educators experienced anxiety about their preparation of their staff and students but indicated confidence in the process itself. Leaders and teachers alike were positive in their comments as indicated by a few selected responses.

- A rural superintendent said, “My initial response just like ‘Oh, this isn’t going to be painful.’ You know, it’s not going to be a bad thing. And then also, I was delighted by not only the ease and comfort level of our teachers but their eagerness to share and their excitement in talking about it. So I think it was just like ‘Whoa, they believe in this.’ So that was good.”
- A female rural elementary teacher said, “They went through every step of each one of the different criteria and related it to our portfolio and how it matched or how it didn’t and what suggestions they made.”

Some leaders were amazed by how much they did know and understand about the process when given the opportunity to discuss what their school or districts were doing to accomplish the requirements of the accountability process.

- A rural superintendent said, “I can’t believe I know what a KR21 is or a decision consistency model or the idea of sufficiency. And what’s really impressive is to sit down with a group of superintendents and have discussions like this where, you know, I think a few years ago I wouldn’t have a clue. “

Many teachers and leaders were very positive about the reviewers themselves and their ability to lead the review process at the building and district level as shared by these educators.

- A female rural middle school teacher stated, “These people don’t look nearly as scary as I thought they would. You know, of course the two women who came in to review our portfolio, you know, looked professional and everything but I don’t know, I think I had this expectation, like the men in black were going to show up with the sunglasses and what not. Very friendly, very approachable. . . .So I was a little nervous to begin with but, immediately when they began the process and put me at ease”
- A female rural assessment coordinator said, “I was very pleased. I liked that they were positive, they had good things to say, and they also had very valuable recommendations to help us improve. . . .”
- A rural superintendent in a large district said, “What I think helped with that is I believe we sent them information in advance. And they responded to some of that information in a way almost immediately that said given our overview of this information, it looks like you’re doing everything the way you’re supposed to be doing it; that you have actually some exemplary kinds of practices going on. Given some of the discussion, it was clear that they were more interested in helping us, in giving us a chance to reflect and to hear their comments about what we are doing.”

It was evident that members within all the groups benefited from the discussions held during the review process as indicated by these educators.

- A rural superintendent said, “Yeah, we talked not only about our own practices, but things that they have seen in other districts and some of the processes that they went through.”
- A non-rural assessment coordinator said, “I started this process by looking through that portfolio, the paper portfolio, and what I saw the process do is peel away some of the unnecessary formality and really was a process of good practitioners talking to you about what needed to be done in the process. And I guess I appreciated that part of it.”

Many reviewers who conducted reviews commented about how much they personally learned when they conducted reviews.

- A female rural assessment coordinator stated, “Everybody should be a portfolio reviewer because it’s like going and grading the writing tests. You really know what you’re doing then-to be on the other side of that.”
- A female rural assessment coordinator said, “I’ve grown tremendously serving on peer reviews. That’s been a wonderful experience. Everything about it has been a growth experience.”

The Nebraska Department of Education worked diligently to support districts in their accountability process and prepared districts for the new process as well.

- A female rural assessment coordinator stated, “There weren’t any surprises yesterday which I attribute to just all the correspondence that the state did with all of us to get us ready.”
- A female rural assessment coordinator stated, “We’ve told people for seven years this is what they needed to submit. But I think that they are much clearer. NDE has gotten much clearer in what they are looking for. The rubrics have been very helpful. Having people come on site just keeps people more attuned to the process.”

The Nebraska-led Portfolio Peer Review Process was completed in one year with all districts presenting their assessment portfolios to review teams. The new knowledge and skills learned by reviewers, state leadership, and district teams was extraordinary as evidenced by the results of the surveys and the interviews. As with any new process, there were challenges and new learnings.

Theme 8: Challenges

As with any evaluation, the strength of the evaluation rests with the knowledge, skills, and experience of the evaluators. The challenge for Nebraska will be to create a large enough group of evaluators to continue reviewing every school district in the state. Many Nebraska rural districts are isolated and difficult to access. Maintaining teams of people who are able to conduct peer reviews with expertise will be a challenge. Educators had different experiences with the evaluation process.

- A female rural assessment coordinator said, “I’m going to say that I think our review probably didn’t start off as I had hoped it would start off. In other words I guess, my expectations of the flow of the day differed a little bit from the reviewers, and I’m going to say specifically one of the reviewers, maybe not both of them.”
- A female rural assessment coordinator in a consortium stated, “I guess, between you and me, I was disappointed in how things started. I will take the ownership of that. . . . There was a mismatch in the expectations of what was going to take place. I think that that’s what started the day off wrong. . . . If I had to do it all over again, I wouldn’t have given them my consortium portfolio. . . . You know, I’m very proud of what we do as a district. And if that was not represented in the way I set the day up, then live and learn on my part as well.”

It was a challenge to both reviewers and districts to provide a schedule of events for the day. The vision for the day’s schedule was not always shared promptly enough between reviewers and districts to allow for substitutes to be called for teachers. This discrepancy created some problems.

- A rural assessment coordinator shared, “I had received a phone call the night before in regard to what time subs would be needed. . . . I was informed that subs wouldn’t be needed until 10 o’clock. I had established for subs to be available for the four teachers at 8:30. I think that kind of got us going on the wrong foot.”

Both reviewers and educators in districts who initially conducted the first reviews felt more inclined to feel frustration about the process. No one really knew how the actual day would evolve and there were some differences of opinions between the reviewers and the district leadership.

- A rural assessment coordinator stated, “Again, I was under the assumption that conversation was going to be the heart of determining whether or not the process was meeting the rigor of the six quality criteria. I think the reviewers thought that there would be . . . a very articulated in written form portfolio. You know my vision was that teachers would come to the table pretty much immediately and start the discussion process with each one of the quality criteria. It took us a little bit to get to that point of agreement, I guess. Once we got there, things really progressed nicely. We had a very, very, very productive, informative, affirming day. But it did take us a little bit to get

going. I'm going to say that at the onset of our day, I wouldn't call it negative, but I would say that it was . . . we had a little bit of a rocky start."

Reviewers themselves strengthened their knowledge as they conducted more than one review. The suggestions to use experienced reviewers with someone new to the review process also strengthened the results of the Peer Review.

- A research coordinator in a large non-rural school said, "So training and that ongoing psychometric support throughout the week is extremely important to that process. As a reviewer, the process is a wonderful experience. I think it was helpful on the second round to pair people that had done it before with people who were new rather than sending people out new by themselves because there's a lot of growth each time you do it. And you can share experiences."

The challenge for districts is to stay open minded about the process. This assessment coordinator described that challenge precisely.

- A female rural assessment coordinator said, "I think probably to be open-minded and receptive to other methodologies. You know when you're in a district; you get pretty tunnel vision and focused on the way you do things as a district. I think one of the healthiest pieces of the peer review, too, was that the two reviewers were able to say you might consider. . . . I think we need to realize that there are lots and lots of effective processes out there. And we need to be careful not to say this is our process, this is what we're going with. I think we always need to be open minded to new ways, more efficient and effective ways of doing things."

The challenge to work with districts in consortia (write a common portfolio) was a problem for both reviewers and the districts being reviewed. Districts indicated their frustration with receiving a different rating than a district within the consortia that had the same portfolio but received different ratings. Involvement of all staff and stakeholders is an essential piece of the accountability puzzle. Districts in consortia that received different ratings will continue to be a challenge for the NDE.

- A female rural assessment coordinator stated, "We have always tried to do the best that we know how to do and the appropriate things. However, we've learned a lot in this last year about specifically reliability. . . . We've been in the crunch where we were a consortium school and got a rating different than other schools in the consortium even though we were all doing the same work. So I think that the peer review process, even though it has caused us some long hours at times, has been really, a major improvement I think."

Some districts felt that their reviewers were biased towards the online system as evidenced by these comments.

- A female rural assessment coordinator stated, "There was some bias in the initial phases. I had districts that called and said my reviewer flat out said they don't like the online. So I know there was some bias initially. I know NDE has been made aware of those and they have tried desperately to eliminate that bias. I've had school districts who have said to me I don't want so and so or so and so, and so and so as my reviewers because I know they're biased. So I think NDE probably needs to make certain they work hard to eliminate that."

- A rural superintendent said, “The on-line system is probably one that’s fairly controversial, and because it’s so controversial I think that I really was hoping that our peer review would glean the positive pieces of what that can be and how it can be a tool in the process. And the other piece that, originally just came to my mind is there is no way in a notebook you can get what’s really happening as far as the rich discussion and terms of staff interaction and growth and collaboration. You just can’t capture that, that on paper”
- A rural superintendent said, “There are pitfalls and drawbacks to the on-line and while it’s quick and it’s good, a good way to house the data and the immediate feedback is good. Many times the first choice response is radio button clicking, you don’t always know. So even if your numbers in the end are really high, I think it’s pretty tough to know. I don’t know that I have recommendations for others in regard to that...even though secondary people have been tough to drag along, it is the secondary people who have been the best at this because they are the ones . . . the Carnegie units and the credits and we want to make sure an A is an A is an A, the grading, so I think they’ve been great help to us, they’ve pushed us a little bit.”

Revising and improving district assessments are not without challenges.

- A rural superintendent reflected on two of these challenges, “Now, there are two things here. One we don’t want to be constantly, constantly changing those. And we don’t want to make them easier every time we change them.”

Even though districts were informed at the onset by the NDE that those who were first to be reviewed might receive "pending" ratings, which could be rectified before the final submission, districts indicated frustration.

- A rural superintendent in a large district said, “We’re a little bit shocked by our rating that we received this year in the sense of we received a pending approval. The reason we received a pending approval is because we are engaged in continuous improvement. And so, despite the fact that we have good alignment, we’re trying to make them align to an even greater degree. We’ve gone in with assessments and modified those assessments. After we assess the kids, after we did all the things, we go in and modify that a little bit. We then run all the things that we have to do because we know what modifying an assessment means to us as far as work. . . . From the perspective of this portfolio review process, it would have been better for us to have said these are perfect tests and we’ll never modify them.”
- A rural superintendent in a large district said, “We walk around and say the fact that we’re pending is badge of honor because we’re trying to improve.”

As with any new process, many educators had suggestions that they believed would strengthen the process and offered some ideas that might be considered in the future.

- A rural superintendent in a large district said, “. . . but somehow, when those teams got together with their experts, the experts say, ‘well you can’t say that these people meet because they’re modifying their assessments. And until they get the statistics on their modified assessments, we have to give them a pending approval.’ That was kind of disheartening to us.”

Finally, students who required special accommodations continue to be a challenge for districts. Districts do their best to meet the intent of the law but still feel the challenge to do a better job with students with special needs.

- A rural superintendent in a large district said, “It doesn’t work 100% in my opinion. We have some times where I think we have slipped and kids that should have been accommodated were not. OK? I mean, that’s just the honest truth about it. We try to include the accommodations with the assessment so that teachers know. And we try to clearly identify who the students are who are to have those accommodations and what the accommodations are supposed to be for that student.”

Theme 9: New Learnings

While Nebraska school districts have been immersed in the accountability process for at least seven years, they continue to improve and strengthen their process, as evidenced by the new learnings for reliability and validity.

- A female rural assessment coordinator said, “One thing [I learned] was to show reliability other than the KR 20. I mean, I knew there were different ways. But, that and the double scoring was basically it. But knowing that we can use decision consistency even on a subjective test, I learned that. Then matching our standards better to our curriculum.”
- A female rural assessment coordinator said, “I heard them talking about it and every time someone talks about the process you learn little things that you didn’t know before even when you’re in it.”

Some educators indicated growth in their ability to assess whether an assessment was valid and reliable.

- A rural superintendent said, “Yeah, I think they have tremendous value. I’ve been much more critical of different assessments, published assessments that I see. I look in our textbooks and I think, “Who in the heck wrote this stuff?” and you know, that this is poor and just I guess, you know you used to take things that were in print as “Oh, somebody must have known what they were doing.”

Districts were encouraged to review the process that had been in place for several years and discovered they had some work to keep their assessments current.

- A female rural assessment coordinator said, “I’m definitely going to do some retraining on bias. They said our bias training was too old. I agree with them. I need to do it in a more formal way, so I’m definitely making that change.”
- A non-rural assessment coordinator said, “This portfolio review really made me delve into the test development, it was probably the best thing that could have happened and I learned so much about what needs to be done to have a quality assessment and a quality system, that I feel really fortunate that there’s a lot of work and, you know, that part was not always fun and easy but it was probably the best thing for me to ramp up as the new assessment coordinator.”

Many educators felt that the Peer Review Process allowed the district to honor their teachers and the amount of work they have done to make the process strong and viable.

- A rural superintendent said, “It was affirmation of the high value of allowing teachers to own the process . . . they need to honor the teachers in the classroom, and that work and discretion. That was just absolutely ignited again in the room with the reviewers.”

- A female rural assessment coordinator said, “I think it helped me, I guess, value the appreciation teachers feel in their part of the system that they’re being validated by what they do, that they see that as so important to the system and their willingness to step up and do that.”

The process allowed districts to re-evaluate the work that had been done and acknowledge the work that needs to be accomplished to reinforce student learning. The support of leadership to continue to build upon the success of the process at the classroom level is needed.

- A female rural assessment coordinator said, “There are a couple things I would say up front that I realize now as a result of that experience. First and foremost, I think it would be that it is important to keep teachers informed on a consistent basis in regard to everything that’s out there. . . . You know, you learn the process, you do the process, but you don’t really discuss the process because it just becomes such an integral part of what we do. I think that I will probably hold more periodic discussions about each of the quality criteria, so it really remains at the surface a little more readily for teachers.”
- A female rural assessment coordinator said, “Number one I think the main focus was always instruction, which it still should be. But the reality is teachers who are awesome instructors weren’t so awesome at assessment. And this has really helped look at...how curriculum, instruction, assessment are all three are intertwined and you can’t do one without the other. They’re all part of the same process.”

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

Study II: Leader and Teacher Perceptions of STARS



Classroom teachers give young people what they sometimes get nowhere else in Society—a sense that they have promise, that they have talents, that they are special.

Parker J. Palmer (Taulbert, 2006)

Study II: Leader and Teacher Perceptions of STARS

Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln

Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American

INTRODUCTION

This section will highlight the statistical or noteworthy findings of the major themes addressed by an online survey and interview data collected and analyzed during the 2006-2007 year. Online surveys were administered to leaders and teachers at the elementary, middle, and high school levels. The significant findings are supported with quotes that were collected during the interviews for those items where quotes were available.

PURPOSE OF THE STUDY

The purpose of this mixed-methods study was to determine if there were similarities and/or differences between leaders and teachers regarding district support, assessment literacy, use of data in classroom settings, leadership, instructional impact, and external support.

RESEARCH DESIGN

A mixed-method design was selected for use to strengthen the study results. The study began with a broad online survey for both teachers and leaders in order to generalize results to a large statewide population of leaders and teachers at all levels. The qualitative open-ended interviews

were conducted to collect detailed views from a purposeful sampling of leaders and teachers based on geographic location, school district class, and the percentage of students on free/reduced lunch (high, middle, low based on the statewide average). This selected group of participants was interviewed to help uncover, confirm, or qualify the basic findings from the survey and discover other pertinent understandings.

Survey Sample

For this 2006-07 mixed-method study, there were 254 school districts in the state at the time the data was collected. For the purpose of this research, Nebraska public school districts were divided into two classifications, non-rural and rural, based on population characteristics unique to Nebraska. Non-rural districts were defined as metro-area districts in large and mid-sized cities, large towns, and the urban fringe. All other districts were classified as rural. Of the 254 public school districts, 5% were classified as non-rural and 95% were classified as rural.

Leaders and teachers were asked to complete online surveys made available to the membership of the Nebraska Council of School Administrators and the Nebraska School Education Association, representing over 98% of the leaders and teachers in the state. Three hundred thirty-four (334) leaders and 562 teachers completed surveys. Of the leaders returning surveys, 83 were superintendents, 51 were curriculum/assessment coordinators, 116 were middle and secondary principals, and 84 were elementary principals. Of the 562 teachers returning surveys, 439 were classroom teachers and 90 were teacher specialists. Thirty-three (33) respondents did not classify themselves by position. Of the 529 teachers identifying themselves by position, 214 were elementary teachers, 114 were middle school teachers, and 201 were high school teachers.

Interview Sample

Selection of the sample school districts for the qualitative interviews was based on a stratified, purposeful sample using district class (Classes II-V), geographical areas (East, Central, and West) and free/reduced district lunch (high, middle, low based on the statewide average). Follow-up interviews were conducted with 38 leaders: 7 superintendents, 5 curriculum/assessment coordinators, 10 middle and secondary principals, 8 elementary principals and 8 SPED and/or ELL Directors. Teacher interviews were conducted with twenty-three (23) classroom teachers: 12 elementary and 11 middle school and/or high school.

Instruments

Leaders and teachers responded to the STARS Leader Survey (Appendix F) and the STARS Teacher Survey (Appendix G). This 67-item online survey was based on a five-point Likert scale for each item, with “1” representing “none of the time,” “2” representing “very little of the time,” “3” representing “some of the time,” “4” representing “most of the time,” and “5” representing “all of the time.” The online surveys focused on District Support, Assessment Literacy, Data, Leadership, Instructional Impact and External Support.

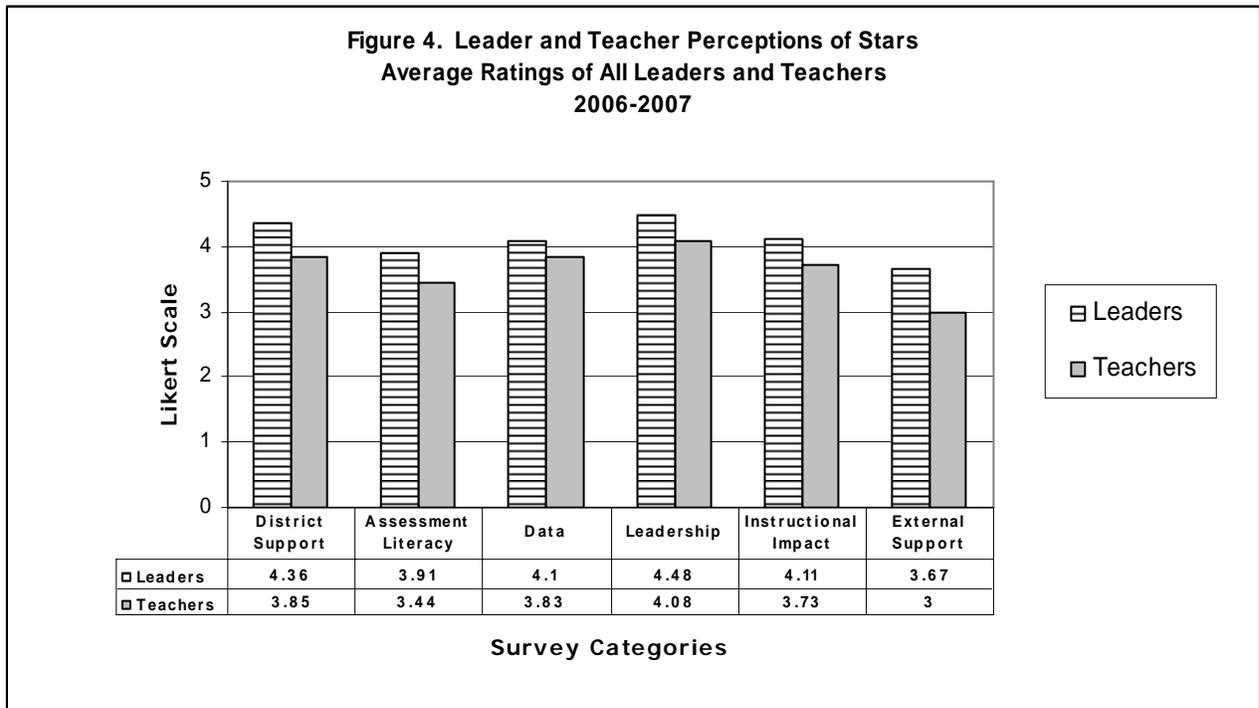
The STARS Research Interview Protocols consisted of demographic information and ten questions for both leaders (Appendix H) and teachers (Appendix I). These questions targeted the participants’ perceptions of STARS training, school leadership, student competence and academic growth, integrating assessment results into instruction, school improvement, instruction for special populations, the effects of mobility, and monitoring and reporting STARS

results. Probes were provided for use by interviewers. Interviewers were provided a STARS Interview Manual and received training in order to ensure the integrity of the qualitative data.

RESULTS

The survey and interviews supported the six themes examined on the survey: District Support, Assessment Literacy, Data, Leadership, Instructional Impact, and External Support. Three additional themes emerged: Special Populations, Communication, and the STARS Process. These themes helped to identify the many pathways that educators have taken on their journey to assessment excellence.

Noted in the graph is the mean for each survey category for both leaders and teachers. The highest average score for both teachers and leaders was in “Leadership” while “External Support” was the lowest average score for both groups. However, teachers rated both areas lower than did leaders.



Theme 1: District Support

LEADERS: Survey Questions 1-11

In the area of District Support for all leaders, responses ranged from 4.30 to 4.42 with an average of 4.36 on the Likert scale with "1" representing “none of the time” and "5" representing “all of the time.” For the total area of District Support there was no significance noted.

While the total area of District Support was not significant, there were individual survey items that indicated some significance for leaders. Non-rural leaders were more likely to score the following items as a “5” indicating “all of the time” than their rural counterparts:

- Encourages a culture of continuous school improvement (p= .006).
- Provides resources for learning teams as a means of professional development (p= .001).
- Is committed to external communication regarding student achievement (p=.022)
- Has adopted a guiding assessment philosophy, mission, and beliefs regarding STARS (p=.006).
- Provides personnel policies that reflect an expectation of assessment competence (p=.000).
- Provides policies at the district/school level that contribute to STARS assessment practices (p=.01).
- District provides ongoing assessment training for STARS (p=.015).

Female leaders were more likely than male leaders to score a “5” indicating “all of the time” for "the district has assisted teachers throughout the K-12 system for developing STARS assessments (p=.01)."

TEACHERS: Survey Questions 1-12

In the area of District Support for all teachers, responses ranged from 3.16 to 4.61 with an average of 3.85 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” For the total area of District Support, there was no significance noted.

While the total area of District Support was not significant, there were individual survey items that indicated some significance. Rural teachers were more likely to score the following items as a “5,” indicating “all of the time,” than their non rural counterparts:

- Provides services that I needed for the implementation of STARS (p=.000).
- Provides on going assessment training for STARS (p=.001).
- Provides a record keeping system for recording STARS data (p=.014).
- Has assisted teachers throughout the K-12 system for developing STARS assessments (p=.025).

Non-rural teachers were more likely to score a “5” indicating “all of the time” than their rural counterparts on the survey item, "encourages a culture of continuous school improvement.”

On the survey item, "providing resources for learning teams as a means of professional development (p=.003)," middle school teachers were more likely to score as a “5,” indicating “all of the time,” than their high school counterparts.

Discussion

While the total area of District Support was not significant for leaders and teachers, there were significances for individual questions. Leaders and teachers are supportive of a mission focused on improving student achievement through continuous school improvement.

- A rural superintendent stated, “I think the most important thing is working with the board. I want the board to understand this is how we measure student progress and . . . for the first time, its taken six years, but the board’s ultimate goal is to improve student performance and so all of our discussions are focused around that.”
- The superintendent indicated the emphasis placed on using data to monitor progress when he stated, “If the board’s number one mission is to improve student performance, obviously we’re looking at data on a regular basis. Annually we re-evaluate our data and

we are making progress. The board holds me accountable and therefore, I hold the building administrators accountable that we're moving students in the right direction or we're improving student performance."

Therefore, when board policy supports student achievement, the effects are felt throughout the system by actions taken by both leaders and teachers.

- A female rural assessment coordinator said, "Policies, policies. Plans. We have put these together in the past year and that has really helped. . . . We've had a district assessment plan, but we revised that and kind of streamlined it. We have a K-12 comprehensive plan that shows at each grade level what the various assessments are that we use throughout the district."
- A female non-rural assessment coordinator shared, "We are a district that does not require that everyone get the assessment at the same time. So we do allow for that flexibility. . . . Teachers then do long-range planning and use what we call an assessment plan card . . . when Standard I was taught, when Standard I was assessed, when Standard I was re-taught if need be, and was re-assessed."

Policy and long-range planning have guided school principals in the development of monitoring and reporting systems.

- One rural high school principal stated, "We just monitor how our kids do on the standards and then if we find an area where we're weak at, we'll go back with our curriculum mapping and we can pull that out and say, OK, standard 8.11. How many times has it been hit and what grade levels . . . and [the tech expert] will print out a report on that. And then we can get the teachers together to talk about, 'Is this accurate?'"
- A female rural assessment coordinator stated, "Probably the biggest change is . . . that ongoing monitoring of progress . . . knowing what the target is, knowing where a child is in relation to that target and knowing what steps instructionally a teacher has to take to get the child to the target. And we do that by the ongoing assessments that we have in place."

As indicated throughout the interviews, much of the training is now occurring within school buildings rather than conducted by external experts, especially when a school district has an appointed assessment coordinator and/or curriculum director.

- A female rural assessment coordinator stated, "One of the things that we've implemented this year across the district in each building is professional learning teams. . . . We have by grade level and by content area, we're building goals, so teachers working on those teams collaborating, researching, reading . . . I think that has really made a difference in raising that bar for teachers in learning the process."
- A male rural superintendent shared, "I think just through our in-service, making sure that everybody understands that we do have an assessment process in our district and we use that data."

The implementation of STARS requires that many support systems be in place that assists teachers in the use of data to support improvement of student learning.

- An assessment coordinator said, "There are two types of reports that we generate. One is, I'm just going to call it a 'class success rate'. And what it does is it gives the teacher an opportunity to see how her kids last year did on the CRTs. Although teachers are in

charge of administering and recording that information in their grade book, sometimes it's nice to have all that brought back to their attention and put into one sheet."

- Another assessment coordinator stated, "Right now we create three of those reports annually without being prompted. The other reports that we're generating are those that have been on people's wish lists. And it isn't so much that we're getting new information, it's just putting it in formats that allow the buildings what they want to see."

Because there is not a standard policy in many school districts, quality monitoring and record keeping is school based and often yields a lack of standardization for records moving from district to district. Some schools have very clear policies on record keeping that assists students that are mobile by providing for efficient transport of records from one school to another so that student learning can be maximized.

- A rural middle school principal emphasized, "And the nice thing about our records is we can quickly bring this student up and send their information to the school that they move to and say here's how they've been doing on the state assessments in all areas."

In summary, many non-rural leaders have become more concerned about developing their own district/school based policies that are yielding a greater focus on the monitoring and reporting of student data and, in turn, upon increasing student achievement. Rural teachers continue to emphasize the importance of resources, services, training, and record keeping as key to successful implementation of the STARS.

Theme 2: Assessment Literacy

LEADERS: Survey Questions 13-23

In the area of Assessment Literacy for all leaders, responses ranged from 3.88 to 3.94 with an average of 3.91 on the Likert scale with "1" representing "none of the time" and "5" representing "all of the time." For the total area of Assessment Literacy, there was no significance noted for leaders.

While the total area of Assessment Literacy was not significant, there were individual survey items that indicated some significance for leaders. Rural leaders were more likely to score a "5," indicating "all of the time," than their non-rural counterparts on the survey item "I participate in the development of the STARS assessment portfolio sent to NDE (p=.000)."

Assessment/Curriculum Coordinators were more likely to score a "5," indicating "all of the time," than superintendents (p=.000), middle/secondary principals (p=.000), or elementary principals (p=.000) for the survey item "I participate in the development of the STARS assessment portfolio." Assessment/Curriculum Coordinators were also more likely to score a "5" than either middle/secondary principals (p=.034) or elementary principals (p=.005) for the survey item, "teachers that work with students in special populations are involved with STARS."

Middle school and secondary principals were more likely to score a "5," indicating "all of the time," than assessment/curriculum coordinators for:

- Teachers are committed to improving their own assessment competence (p=.015).
- Administrators support 'school-based teacher-led' assessment (p=.049).
- Teachers support 'school-based teacher-led' assessment (p=.031).

Middle/secondary principals were more likely to score a “5” than superintendents on “I participate in the development of the STARS assessment portfolio (p=.007).”

Elementary principals were more likely to score a “5,” indicating “all of the time” than superintendents on:

- Assessments are aligned to the state/district standards (p=.015).

Elementary principals were also more likely to score a “5” than assessment/curriculum coordinators for “teachers are committed to improving their own assessment competence (p=.003).”

Female leaders were more likely than male leaders to score a “5,” indicating “all of the time,” for:

- Assessments are aligned to the state/districts standards (p=.030).
- Teachers write their own STARS assessments (p=.016).
- I participate in a learning team within my district or school (p=.033).

Leaders with more years of experience were more likely to score a “5,” indicating “all of the time,” than those with little experience for “I participate in the development of the STARS assessment portfolio sent to NDE (p=.029).”

TEACHERS: Questions 13-23

In the area of Assessment Literacy for all teachers, responses ranged from 2.20 to 4.68 with an average of 3.44 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” For the total area of Assessment Literacy, there was no significance noted for teachers.

While the total area of Assessment Literacy was not significant, there were individual survey items that indicated some significance. Rural teachers were more likely to score a “5,” indicating “all of the time,” for:

- I select assessment items from a common bank or pool (p=.004).
- Administrators support ‘school-based teacher-led’ assessment (p=.004).
- I participate in the development of the STARS assessment portfolio sent to NDE (p=.000).
- Teachers that work with students of special populations are involved in STARS(p=.001).

Non-rural teachers were more likely to score a “5,” indicating “all of the time,” for the survey item “I participate in a learning team within my school (p=.001).” Middle schools teachers were more likely to score a “5,” indicating “all of the time,” than high school teachers on the same item (p=.002).

Classroom teachers were more likely to score a “5,” indicating “all of the time,” than specialist teachers for the survey items:

- I participate in the development of the STARS assessment portfolio sent to NDE (p=.018).
- I participate in a learning team within my school (p=.018).

Discussion

Assessment literacy is not as prevalent in the discussion with leaders or teachers now that the initial development of assessments has been completed. Both administrators and teachers emphasized in the interviews that they were committed to building their own assessment competence and supported "school-based teacher-led" assessment. Science teachers are new to the process, but many have been working for several years to get their assessments ready for the reporting year.

- A middle school science teacher shared, "As I see it, it's kind of evolved. We have been working on assessments . . . basically we went through and aligned our curriculum and then we started developing assessments a couple years ago."

As teachers shared about the growth of their own assessment competence, they also spoke about their own involvement in the success of the STARS process.

Much like teachers, leaders' assessment literacy has also developed through their work with teachers in the STARS process.

- A superintendent shared, "I think just through our in-service, making sure that everybody understands that we do have an assessment process and we use data ensured that everyone became assessment literate."
- A new principal shared how they have developed their own assessment literacy skills when she stated, "My personal experience having come to this state and not familiar with the STARS initially, during the past six years I've come to understand how they develop the assessments and why they use those criteria in the process. I suppose just working with them, reading about them in the STARS updates, things like that, more than anything else has helped me to become assessment literate."

Over time there has been a noticeable increase and support of the assessment literacy of both leaders and teachers.

- An elementary principal emphasized that maintaining assessment literacy is accomplished by everyday involvement in the issues of STARS. He stated, "Well, for the staff, we meet every week and a part of the conversations that we have probably monthly include some sort of issues on where we're at with STARS testing, where we're heading to, how we're going to reach those goals, are our children well prepared?"
- Some leaders have not been satisfied with assessments developed outside of their district and now feel confident enough to develop their own assessments as illustrated by this high school principal's statement, "We're going to design our own assessments and the ESU staff developer will show us how to make sure we hit all six criteria."

Non-rural ELL teachers shared how they were involved in the STARS process and how they built their assessment literacy skills by working with their colleagues.

- An ELL teacher stated, "I think that the teachers that are very involved in writing CRTs and the textbook adoption process, are the individuals that are connecting more with the six quality standards. We absolutely do go over CRTs with all of our newly hired employees. We do it through new teacher meetings, during the course of the year."

The total area of Assessment Literacy for leaders and teachers was not significant during the sixth-year study as it has become a part of the daily routine of teachers and leaders as they perfect their assessment literacy skills through their continued work with the STARS

assessments. The biggest challenge identified relative to assessment literacy in large school districts was the desire to get all teachers on board.

- This assessment coordinator emphasized, “We’re doing our best by doing the trainer/trainer model trying to get more people understanding formative assessment so that they understand that assessment should not be something that is just tagged on at the end of something, it’s something that goes on daily. It’s part of that normal curriculum, instruction, assessment cycle.”

Additionally, unlike rural districts it was noted that it is very difficult for large districts to make a significant change without ample time and notification.

- An assessment coordinator shared, “It becomes difficult for us when we are asked to change on a dime . . . we don’t touch a subject for seven years once it is done. We continue to do training, we continue to modify, but interventions are set based on the needs of the district.”

Theme 3: Data

LEADERS: Survey Questions 24-31

In the area of Data for all leaders, responses ranged from 3.96 to 4.28 with an average of 4.10 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Data was marginally significant (p=.074) because assessment/curriculum coordinators were more likely to score a “5” than either middle/secondary principals (p=.017) or elementary principals (p=.052).

There were individual survey items for leaders that indicated some significance. Non-rural leaders were more likely to score a “5,” indicating “all of the time,” than their rural counterparts for:

- District/school provides assistance for analyzing STARS assessment data (p=.005).
- New teachers are trained to use data for improving instruction (p=.000).

Assessment/Curriculum Coordinators were more likely to score a “5,” indicating “all of the time,” than superintendents (p=.000), middle/secondary principals (p=.000), and elementary principals (p=.000) for the “district/school disaggregating STARS data.”

Female leaders were more likely than male leaders to score a “5,” indicating “all of the time,” for:

- Using STARS data results to make instructional decisions (p=.033).
- Disaggregating the STARS data (p=.022).
- Using STARS disaggregated data to make instructional decisions about students in special populations (p=.040).

TEACHERS: Survey Questions 24-31

In the area of Data for all teachers, responses ranged from 2.65 to 3.88 with an average of 3.83 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Data was significant (p=.017). The area of Data was also significant for teachers and specialists (p=.002).

There were individual survey items that indicated some significance. Rural teachers were more likely to score a “5,” indicating “all of the time,” than their non rural counterparts for “my school provides a record keeping system for STARS data (p=.000).” Non-rural teachers were more likely to score a “5,” indicating “all of the time,” than their rural counterparts for “in my school districts, new teachers are trained to use data for improving instruction (p=.003).”

Elementary teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for:

- I use the STARS data results to make instructional decisions (p=.000).
- I disaggregate STARS data for students in my classroom (p=.023).
- I use STARS disaggregated data to make instructional decisions about students in special populations (p=.004).
- My school provides assistance for analyzing STARS assessment data (p=.034).
- In my school district, new teachers are trained to use data for improving instruction (p=.034).

Middle school teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for:

- Using the STARS data results to make instructional decisions (p=.015).
- New teachers are trained to use data for improving instruction (p=.019).

Classroom teachers were more likely to score a “5,” indicating “all of the time,” than specialist teachers for the survey items:

- I disaggregate STARS data for students in my classroom (p=.015).
- I score my own STARS assessments (p=.000).
- I record my own STARS assessment results (p=.000).

Discussion

The interview question regarding the use of data yielded a wide variety of ways that educators used data to improve student learning.

- An elementary principal emphasized the use of data, “Well again, I think probably the biggest thing is just making it the norm to look at that data, assess it and think about it. That’s the process we’re in, whereas five years ago that wasn’t happening. Teachers now have data at their fingertips. They can see the trend, they can see the themes, and they can see the problem.”
- Data is being used to plan how to improve student learning though the use of objectives as stated by this elementary principal, “We’ve looked at the data and we’re establishing objectives based on that data on how to improve student outcomes.”
- A superintendent shared, “. . . I say the whole STARS school improvement process has brought data to the forefront, I think now teachers are paying a lot more attention to data than they ever have. And they better understand how students are performing. They have measures to determine if students are actually learning.”
- Teachers emphasized the importance of looking at gap data as noted by this elementary teacher, “I think if you see large gaps, if you see areas that perhaps you’re weak in, or even areas that you’re strong in, I think that dictates then what it is you teach. So I think that data is very important.”

An innovative way of building the skills of teachers in the use of data and creating ownership for data have been implemented using a train-the-trainer model where one person serves as a building data representative.

- An assessment coordinator shared, “We meet with them five times a year, and these individuals typically are leaders in the building in some form. They could be principals, they could be assistant principals, and they could be classroom teachers. Matter of fact, I’d say at least a third to half of them are probably classroom teachers. They come and obviously they learn about data and how to analyze the data and how do you convey that information to your staff.”

A superintendent summarized the importance of providing leadership when using data, “I’ve learned a long time ago if you measure it and you watch that data and you pay attention to it, it becomes important to everybody.” It is evident to the researchers that the use of data in schools has increased over time and is impacting instruction on a daily basis.

Theme 4: Leadership

LEADERS: Survey Questions 32-37

In the area of Leadership for all leaders, responses ranged from 4.41 to 4.67 with an average of 4.48 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Leadership indicated significant differences.

In the total area of Leadership there was a significant difference between non-rural and rural leaders. Non-rural leaders (M=4.58) scored significantly higher than rural leaders (M=4.38) (p=.007). In the total area of Leadership male leaders (M=4.58) responded significantly higher than female leaders (M=4.39) (p=.004).

Assessment/curriculum coordinators were more likely to score a “5,” indicating “all of the time,” than superintendents (p=.008), middle/secondary principals (p=.006), and elementary principals (p=.009).

There were individual survey items that also indicated significance. Non-rural leaders were more likely to score a “5,” indicating “all of the time,” than their rural counterparts for:

- My vision defines how STARS assessment fits into effective teaching and learning (p=.037).
- I define success as high achievement for all learners (p=.046).
- I provide a schedule or plan that coordinates testing of students (p=.014).

Assessment/curriculum coordinators were more likely to score a “5,” indicating “all of the time,” than superintendents on:

- I am assessment literate and committed to assessment literacy for all (p=.002).
- I provide a schedule or plan that coordinates testing of students (p=.000).

Assessment/curriculum coordinators were more likely to score a “5,” indicating “all of the time,” than middle/secondary principals on:

- I define success as high achievement for all learners (p=.020).
- I am assessment literate and committed to assessment literacy for all (p=.001).

- I provide a schedule or plan that coordinates testing of students (p=.003).

Assessment/curriculum coordinators were more likely to score a “5,” indicating “all of the time,” than elementary principals on:

- I am assessment literate and committed to assessment literacy for all (p=.005).
- I provide a schedule or plan that coordinates testing of students (p=.006).

Female leaders were more likely than male leaders to score a “5,” indicating “all of the time,” on:

- I am committed to academic excellence that is communicated through a clear vision (p=.011).
- My vision defines how STARS assessment fits into effective teaching and learning (p=.002).
- I define success as high achievement for all learners (p=.001).
- I am assessment literate and committed to assessment literacy for all (p=.000).
- I am committed to encouraging new teachers to become school leaders (p=.041).

TEACHERS: Survey Questions 32-37

In the area of Leadership for all teachers, responses ranged from 3.53 to 4.61 with an average of 4.08 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Leadership indicated significant differences between elementary, middle and high school teachers (p=.020).

There were individual survey items that indicated some significance. In the total area of Leadership there was a significant difference between rural and non-rural teachers. Rural teachers responded significantly higher than non-rural teachers for “my vision defines how STARS assessment fits into effective teaching and learning (p=.006).”

Elementary teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for:

- I am assessment literate and committed to assessment literacy for all (p=.026).
- I provide an assessment schedule that coordinates testing for my students (p=.001).

Middle school teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for “my vision defines how STARS assessments fits into effective teaching and learning (p=.035).”

Classroom teachers were more likely to score a “5,” indicating “all of the time,” than specialist teachers for “I provide an assessment schedule that coordinates testing for my students (p=.002).”

Discussion

A focus on student learning within the vision and mission of a district/school was the essence of the successful implementation of the STARS process.

- This rural superintendent shared, “Our mission statement is student learning, student achievement, student success. It’s very simple, it’s right to the point and that’s what we’re all about.”

- Another superintendent indicated, “If our number one mission as a district is to improve student performance, then I have to make sure that we have measurement tools in place to measure student performance that we all agree is important and STARS is a piece of that.”
- Yet another superintendent summarized the connection to vision and mission when he stated, “We’ve always had a mission statement that we want all students to achieve at the highest possible level and to be successful and all means all.”

In order for students and teachers to be successful, leaders looked at evidence within the teaching process to ensure that students are learning.

- A rural principal stated “accountability, making sure my teachers are teaching and then that my teachers are assessing.”
- A rural female elementary principal stated, “Well, of course careful monitoring through observation, my own observation, . . .so if I have a teacher who is not doing well at bringing closure to her lessons, she can go in and look at master teachers who do an absolutely phenomenal job at closure.”
- A secondary principal shared, “through our evaluation and observation process, as far as instructional management, we try to make sure that our teachers are doing as good a job as we can. It’s kind of a self-improvement model. We rely a lot on trying to get teachers to take a look at what they’re doing and kind of self-evaluate a little bit.”

Finally, leaders emphasized the importance of maximizing the time for instruction.

- A rural middle school principal shared, “We try to eliminate all interruptions in the classroom. When the teachers are in there, the students are number one.”
- Another elementary principal stated, “By protecting that time, I’d say would be the best way I have maximized time. The primary grades have their extra activities, such as PE and music in the afternoon or as late in the morning as possible to try to maximize the use of that time in the morning for reading, for the basic skills.”

Teacher leadership is essential to the success of the STARS process.

- A female rural middle school principal shared, “I count on my fourth grade teachers, eighth grade teachers and eleventh grade teachers in those core areas especially reading, writing, speaking, listening and math to know the process and be able to talk about it with others. They also go out and help if a teacher needs help using the assessments or looking at the criteria, so they’re there to help me, I use them a lot.”

Teachers are also leaders in decision making and ownership.

- This rural superintendent shared, “Well, I think the staff is involved at every level in all the decisions. We just don’t leave the staff out of the loop in any of the process.”
- Another rural middle school principal shared, “Oh, its staff driven. They have to have ownership in it in order to make it effective. If it comes top down it won’t mean anything.”

The total area of Leadership indicated significant differences for both leaders and teachers. Leadership is critical to the success of the STARS process and has emerged as a strong indicator of district and school success.

Theme 5: Instructional Impact

LEADERS: Survey Questions 38-61

In the area of Instructional Impact for all leaders, responses ranged from 4.07 to 4.16 with an average of 4.11 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Instructional Impact was not significant.

There were individual survey items that indicated significance. Non-rural leaders were more likely to score a “5,” indicating “all of the time,” than their rural counterparts for:

- A curriculum implementation plan describes roles and responsibilities and ensures standards based instruction (p=.000).
- Consistency in achievement expectation for all educators is ensured (p=.046).
- New teachers are given specialized assistance in understanding the STARS process (p=.037).
- Achievement standards are held high for all students (p=.026).
- Administrators and teachers collectively use data to develop support systems for students not meeting the standards (p=.015).
- Leaders are focused on instruction and continuous improvement (p=.001).

Rural leaders were more likely to score a “5,” indicating “all of the time,” than their non-rural counterparts for “I have evidence that standards are taught (p=.029).”

Assessment/curriculum coordinators were more likely to score a “5,” indicating “all of the time,” than superintendents on:

- Curriculum is aligned to the state standards (p=.035).
- I have evidence that lesson plans are aligned to standards (p=.040).
- I have evidence that standards are taught (p=.003).
- I have evidence that all students have an opportunity to learn (p=.056).

Middle/secondary principals were more likely to score a “5,” indicating “all of the time,” than superintendents on:

- Curriculum is aligned to the state standards (p=.017).
- I have evidence that standards are taught (p=.005).

Middle/secondary principals were more likely to score a “5,” indicating “all of the time,” than assessment/curriculum coordinators on:

- Teachers are held accountable for teaching the adopted curriculum (p=.003).
- Instructional differentiation is provided to meet individual needs of students (p=.006).

Middle/secondary principals were more likely to score a “5,” indicating “all of the time,” than elementary principals on “opportunities are provided for teachers to share STARS assessment results with parents (p=.004).”

Elementary principals were more likely to score a “5,” indicating “all of the time,” than superintendents (p=.004), assessment/curriculum coordinators (p=.000), and middle/secondary principals (p=.054) on “teachers are held accountable for teaching the adopted curriculum.”

Elementary principals were also more likely to score a “5,” indicating “all of the time,” than superintendents on:

- New teachers are given specialized assistance in understanding the STARS process (p=.041).
- Teachers understand and apply the principles of sound grading practices (p=.045).
- I have evidence that lesson plans are aligned to standards (p=.001).
- I have evidence that standards are taught (p=.041).
- Administrators and teachers collectively use data to develop support systems for students not meeting the standards (p=.018).

Elementary principals were more likely to score a “5,” indicating “all of the time,” than assessment/curriculum coordinators in:

- Teachers understand and apply the principles of sound grading practices (p=.032).
- Instructional differentiation is provided to meet individual needs of students (p=.007).
- Students that do not master the standards are re-taught (p=.022).

Elementary principals were more likely to score a “5,” indicating “all of the time,” than middle/secondary principals on:

- Achievement standards are held high for all students (p=.011).
- I have evidence that lesson plans are aligned to standards (p=.034).
- Students that do not master the standards are re-taught (p=.020).

Female leaders were more likely than male leaders to score a “5,” indicating “all of the time,” on:

- Teachers are held accountable for teaching the adopted curriculum (p=.002).
- I use STARS assessment results to question, modify and adjust district/school instructional decisions (p=.001).

TEACHERS: Survey Questions 38-61

In the area of Instructional Impact for all teachers, responses ranged from 2.14 to 4.68 with an average of 3.73 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of Instructional Impact was significant for elementary, middle, and high school teachers (p=.034).

There were individual survey items that indicated significance. Non-rural teachers were more likely to score a “5,” indicating “all of the time,” than their rural counterparts for:

- My curriculum is aligned to state standards (p=.054).
- A curriculum implementation plan described roles and responsibilities and ensured standards based instruction (p=.034).
- I have visited other school sites to examine successful instructional interventions (p=.000).
- Administrators and teachers collectively use data to develop support systems for students not meeting the standards (p=.000).
- Teachers in my school use STARS data to develop interventions (p=.045).
- My administrator and I focus upon student achievement results with standards in teacher evaluation conferences (p=.002).

Rural teachers were more likely to score a “5,” indicating “all of the time,” than their non rural counterparts for:

- I modify my instructional strategies when students do not perform well on the STARS assessments (p=.007).
- I record the standards that I teach to ensure that all students have the opportunity to learn (p=.001).

Elementary teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for:

- In my school achievement standards are held high for all students (p=.011).
- Administrators in my school assist teachers in making instructional decisions based on STARS data (p=.026).
- I share STARS assessment results with parents of students in my classroom (p=.000).
- Teachers in my school use STARS data to develop interventions (p=.001).

Middle school teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for:

- I share STARS assessment results with parents of students in my classroom (p=.008).
- New teachers are involved in curriculum review so they better understand how curriculum, assessment, and school improvement are aligned (p=.034).

Classroom teachers were more likely to score a “5,” indicating “all of the time,” than specialist teachers for:

- My curriculum is aligned to the state standards (p.001).
- I am responsible for weaving assessment into instruction (p.000).
- I align my lesson planning to the standards (p.005).
- I record the standards that I teach to ensure that all students have an opportunity to learn (p=.000).

Specialist teachers were more likely to score a “5,” indicating “all of the time,” than classroom teachers for “Using instructional differentiation to provide for the individual needs of students (p.000).”

Discussion

Focusing on student learning has driven instruction to the forefront of the STARS Process.

- A superintendent shared, “The biggest change we’ve had to do in education and it’s probably of our teachers, is that we’re trying to teach for learning as opposed to just teaching. We’ve got to make sure that students are learning and the assessment process measures student learning.”

Educators shared the importance of alignment in the conversations with researchers. In order to provide quality instruction, curriculum must be aligned to standards.

- A middle school science teacher shared, “Right now we’re in the process of aligning our curriculum K-12. We’ve been doing a lot of things, meeting with the elementary, figuring out where everybody’s at as far as what they’re teaching, when they’re teaching and just basically getting some continuity within the entire program.”

All educators want to identify the best instructional strategies and interventions to use so that all students can be successful. Best practice research forms the basis for selecting strategies and interventions.

- This superintendent noted, “We study a lot of research. I mean, we spend time looking at tried and proven research-based interventions that work. In other words we want to make sure we’re just not doing a fad. We haven’t picked a school that’s done something for one year and made a difference. We want to look at something that’s been tested over time; similar demographics to what we’re dealing with; populations that we’re addressing, numbers of students, things like that so we spend a lot of time with the research.”

In order for the use of research to be successful it must be based on lots of communication.

- An elementary principal illustrated communication, “So it’s a step-by-step process and a lot of conversations and looking at test results.”

This may also mean that a change in belief systems regarding student failure may be needed.

- As one superintendent shared, “We have bought into the failure is not an option concept! We’ve done some things in terms of our grading policy, to enhance student learning. And we’re working with our teachers to change a mindset about how students are graded in this district. In other words, I have told the teachers in this district that have given students zeroes for not turning in homework that is not an option.”

There are many ways to ensure that students are learning through the use of various instructional strategies.

- A middle school teacher shared, “I guess that depends on the kids. I really monitor the kids. . . . If I see they’re struggling then I’ll change tactics, and if I see they’re getting it I’ll just keep doing what I’m doing.”
- Another middle school teacher stated, “Just as creatively as I CAN. I work with groups. I work using a variety of teaching methods, a variety of strategies.”

Using data to determine what the learning gaps are and how to address those gaps was important to making students successful.

- A rural elementary principal shared, “We have conversations after we look at our STARS scores. Where are the gaps? How are we going to redirect our curriculum to fill the gaps? . . . We have to look at any given test results for gaps in learning and make adjustments in curriculum.”
- Another elementary principal shared, “We do this through curriculum committees. We do this through grade level meetings at the building level. We actually have a school improvement team at each building level that meets on a monthly basis and at those meetings we look at our school improvement goals, we look at the interventions we’re using, we’re looking at the data to see if we’re being successful with it.”
- Yet another elementary principal shared, “We choose our interventions based on the data. We’ve done extensive in-service and gone and looked for good interventions, but all of those interventions are based on the data and the learning style of the child.”

Special populations create a unique set of needs that challenge classroom teachers as well as offer opportunities to enrich teaching in inclusive environments.

- An ELL leader noted, “I think that the greater challenge is with the mainstream teachers that are responsible for serving those students in the math, language arts, and science, having a knowledge of where that child is and what their needs are regarding both assessment and instruction.”

- An ELL teacher added, “I mean because each child is different and each test is different, I think the regular classroom teacher looks at that more than I do, and if they have a concern, they’ll let me know what it is and ask me to do a little intervention on that.”
- An elementary principal shared, “We have really moved from a remediation model to a prevention model. It’s not the kind of thing where we send kids out of the room to finish up their homework. We figure out what the problem is, what the deficit is, and we get in there and try to prevent it early.”

New thinking in instructional strategies has begun to impact the grading practices in schools.

- An assessment coordinator shared, “It has impacted grading practices greatly, I think. In fact we’ve just rewritten our grading policy, so we’re still in transition. We are looking at good practice and we’ve rewritten our grading policies . . . so that grades reflect student achievement . . . and separate behavior and motivation from achievement.”
- A superintendent shared, “STARS assessment has caused us to generate a new grading policy in this district which has not been adopted by the board of education as of yet, but it will be at the first board meeting in April. It’s scheduled to be approved, and again, the assessment team put that grading policy together.”
- A female middle school teacher stated, “I’ve been looking a lot at grading and STARS. I look at those proficiency levels and for me personally, to me it should correspond to the A, B, C, D. For things to be seamless, we should be looking at, what does an A mean? What does a B mean? I don’t think it means the same in all classrooms and then parents get real confused. I’ve done some surveys on this. OK, so my kid’s got an A here, but are they proficient? Well, they should be. If everything is aligned and seamless, then everything should flow into that so there shouldn’t be a difference between the grade they receive on their report card and how they’re doing.”

Theme 6: External Support

LEADERS: Survey Questions 62-69

In the area of External Support for all leaders, responses ranged from 3.60 to 3.70 with an average of 3.67 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of External Support was not significant.

There were individual survey items that indicated significance. Rural leaders were more likely to score a “5,” indicating “all of the time,” than their non-rural counterparts for:

- I participate in professional development for STARS provided by external experts (p=.000).
- The ESU provides ongoing assessment training for me (p=.000).
- The ESU conducts data retreats to help me understand STARS assessment results (p=.016).
- The ESU assists my district/school in the identification of appropriate interventions (p=.051).
- The ESU supported my district/school in preparing evidence for the Nebraska-led peer review of the STARS (p=.000).
- I encourage new teachers in my district/school to increase their knowledge of STARS by participating in professional development at the ESU (p=.001).

Assessment/curriculum coordinators were more likely to score a “5,” indicating “all of the time,” than superintendents ($p=.002$), middle/secondary principals ($p=.030$), and elementary principals ($p=.003$) on “I participate in professional development for STARS provided by external experts.”

Female leaders were more likely to score a “5,” indicating “all of the time,” than male leaders on “I participate in professional development for STARS provided by external experts ($p=.012$).”

TEACHERS: Survey Questions 62-69

In the area of External Support for all teachers, responses ranged from 1.64 to 3.45 with an average of 3.00 on the Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” The total area of External Support was significant for rural and non-rural schools ($p.000$).

There were individual survey items that indicated significance. Rural teachers were more likely to score a “5,” indicating “all of the time,” than their non-rural counterparts for:

- I participate in professional development for STARS provided by external experts ($p=.000$).
- The ESU provides ongoing assessment training for me ($p=.000$).
- The ESU conducts data retreats to help me understand STARS assessment results ($p=.000$).
- The ESU assists my district/school in the identification of appropriate interventions ($p=.000$).
- The ESU provides training in instructional strategies that produce increased student achievement ($p=.000$).
- The ESU supported my district/school in preparing evidence for the Nebraska-led peer review of the STARS ($p=.000$).
- I encourage new teachers in my district/school to increase their knowledge of STARS by participating in professional development at the ESU ($p=.000$).

Elementary teachers were more likely to score a “5,” indicating “all of the time,” than their high school counterparts for:

- The ESU provides ongoing assessment training for me ($p=.029$).
- The ESU provides training in instructional strategies that produce increased student achievement ($p=.024$).
- The ESU supported my district/school in preparing evidence for the Nebraska-led peer review of the STARS ($p=.033$).
- I encourage new teachers in my district/school to increase their knowledge of STARS by participating in professional development at the ESU ($p=.042$).

Elementary teachers were more likely to score a “5,” indicating “all of the time,” than their middle school counterparts for “I participate in professional development for STARS provided by external experts ($p=.032$).”

Discussion

As indicated throughout the interviews, much of the training is now occurring within school buildings or by local ESU experts, especially in non-rural areas. Rural leaders and teachers were more likely to indicate ESUs provided professional development, ongoing assessment training, data retreats, identification of interventions and training in instructional strategies while non-

rural schools tended to provide their own professional development and assessment literacy training. ESUs provided more assistance with portfolio development and review for rural schools than for their non-rural counterparts.

Higher education is one of the many sources of external training for new teachers and teachers currently in the field.

- A female rural assessment coordinator shared, “. . . training through ESU, training through the Nebraska Department of Education. I was also involved in the Leadership for Learning Assessment Cohort through UNL. I finished that last summer so through my involvement with that group of people, I learned more about assessment and the six quality criteria.”
- A rural SPED teacher talked about her daughter’s experience during her practicum at a higher education institution by stating, “My daughter is in early childhood development and did her practicum. She was required to work with preschool kids and she said, ‘Mom, I can’t believe that we’re doing this. We’re matching standards to a curriculum at the school. And then I need to write my objectives and my lesson plans to go along with these standards.’”

Unlike these two positive experiences in higher education, some educators believe that there is a need for more communication about STARS in teacher training programs.

- A rural elementary principal stated, “None of the colleges are fully preparing our teachers-our potential teachers-to understand our testing formats.”

Educators shared other venues of external support.

- An elementary rural principal shared, “. . . through in-services and workshops . . . reading books and attending conferences.
- A rural assessment coordinator stated, “I try and facilitate communication between the Nebraska Department of Education and the district. When we get the updates, those go out to the administrators and to the core chairs. I try and keep everybody apprised of what’s going on.”
- A rural elementary principal also shared that there were web resources that supported her training, “There’s tons of information on the website that is very helpful if people just take the time to read it. It’s there.”

Theme 7: Special Populations

LEADERS and TEACHERS

In the interviews with educators, the theme “Special Populations” emerged. Educators across the state have risen to the challenge of finding ways to provide appropriate accommodations for all students.

- An assessment coordinator shared that a continuum of needs exists, “We believe the needs here are a little extraordinary in terms of the continuum, I mean we have these extremely bright students and we have students who are spending nights in shelters. They’re having a hard enough time learning the English language. You know the needs run such a wide gamut.”

In order to address the many varied needs of special populations, educators shared that they were provided professional development opportunities.

- A SPED Director stated, “We do a lot of professional development that is provided to teams of regular education and special education teachers from buildings to support access and progress in general curriculum for students with special needs.”
- A middle school principal shared, “Well we send them to workshops and in-services and we have workshops and in-services here. We just had one on custom testing. . . . This was especially for our special education students so that they [their teachers] can see that the test doesn’t have to be given fifteen questions at a time. It can be broken down into smaller increments that the student can take over a period of time rather than all in one day.”

Identifying solutions has been difficult for educators as they implemented the STARS process.

- A rural assessment coordinator stated, “This has been like the uphill battle and the mountain that I feel like I’ve trudged for six years. I finally felt, since last year, that we’ve gotten to the top of it. And what we’ve said and we’ve been strong and firm in it, is that the kids are going to be provided the opportunity to learn and I don’t care what their classification is.”
- An example of the expectation for improvement was shared by this curriculum director, “We will take the bottom performers in some different areas, or the special education teacher may take their special education case load and they will do a combination of things. We’ve created these spread sheets where it starts with a student’s name and then it’s the years across the top and for Stanford 10 it will be their NCE scores and they now know if I don’t have at least the same NCE score I don’t have a year’s growth. We’ve got to have more than a year’s growth in a year’s time, for this student. We have to have that or we’re widening the gap.”

Classroom teachers and special education teachers are working together to meet the needs of all students.

- An elementary principal shared, “When you look at test results we see students who have gaps, then we take those and we disseminate further the students that have gaps. How many of them are special education, Title, and ELL? Then we take those subgroups and we meet with the people who are the experts in the area. Those subgroup teachers have to work directly with the classroom teacher to align what they’re doing to make that strong base for that student.”
- A secondary ESL teacher also shared how they work with content area teachers by stating, “The ESL teacher acts as a resource to the classroom teacher so if a student needs to take an assessment with extended time or even bilingual support, that can be provided in the ESL classroom. . . . So that extended time and flexibility is available and that really is very helpful to students.”
- A rural assessment coordinator shared, “The special education teachers have done a great job in identifying those essential skills within the standards that they need to work on along with the classroom teachers, so . . . the instruction goes hand-in-hand. That helps align the curriculum with what we need to focus on.”

Classroom teachers expressed their support for the knowledge of specialized teachers as it relates to modifying and adjusting instruction for special needs students.

- A middle school teacher stated, “Oh the special education teachers are fantastic with modifying, adjusting assignments and having different creative ways to present information to kids.”
- Another SPED Director shared, “I’m thinking about instructional and assessment accommodations that would be identified on students’ IEPs or 504 plans. Once they’re written down on those documents there’s a heightened sense of awareness about the need to be sure those are available to students and there’s a monitoring of the accommodations that are needed and which are provided. Regular teachers are much more keyed into what accommodations a student needs in different subject areas.”
- A middle school principal also addressed another way to provide modifications for students having difficulty by stating, “We also have a pull out program. If a student is having trouble, maybe one or two days a week, they get pulled out. They get one-on-one assistance.”

Along with modifications for special populations, there are challenges with assessing student mastery of standards.

- A rural SPED director shared some of the struggles with assessing SPED students, “On the assessment side right now we’re struggling with the use of scribes. We probably have less than five students that use scribes throughout the year for assignments. I think I’ve talked to all five schools because they’re in the middle of the state writing assessment and they want to be sure that they’re using the scribe appropriately. And so they really are, I think, very conscientious in wanting to do everything that is allowable for the student but really wanting to follow the testing practices.”
- A middle school principal shared one way that they are trying to meet the challenges of testing special education students when she stated, “They’re going to find success. We just had an IEP meeting last night where the special education teacher was saying we’re going to use custom testing for this child and they’re not going to have to take all fifteen questions at one time and we’re going to break them down into the different areas that are being stressed, so I feel like we’re addressing that.”
- A rural SPED Director shared a challenge with providing the needed accommodations for students by stating, “Probably the biggest one (challenge) is . . . the role of accommodations and modifications with special education students. Some kids genuinely need to have modifications within their curriculum because of cognitive disabilities and they’re giving their all yet if a school chooses to use modifications their hands are slapped in that they say that’s a non-participating kid.”

Not only are leaders and teachers worried about an “opportunity to learn” for special education students they are also worried about it for other groups of students.

- A curriculum director shared, “Our high ability teachers were very concerned that STARS would be a detriment to our strongest learners. We helped them figure out a way where we could check that. They said ‘We think kids are dropping.’ So we now keep track of the percent of students staying up in the top ten percentile on different kinds of measures to see if the percent of our population scoring at the top of things is as high as it’s ever been.”
- Another group of students that present a challenge to districts are students of mobility as shared by this superintendent, “The biggest challenge that . . . we face is students coming in and out. We do have a pretty mobile population. I don’t know if we’re any more

mobile than a lot of Nebraska schools our size but the real challenges are students that come in.”

In order to provide for the needs of so many diverse students, districts are finding unique ways to assign their teaching staff and provide the needed accommodations for students to be successful.

- One special education director shared, “Most of our teachers are involved in some co-teaching arrangement, at least part of the day. So they are planning with that regular education teacher in providing the accommodations and special education materials and modifications that might support the special needs student in reaching or demonstrating mastery of those concepts. The incidental benefit for other at-risk learners in those classrooms is significant because they also benefit from the specialized attention and instruction that’s being provided to the special needs students.”

Although many teachers have risen to the challenge of providing for special populations, there is concern that perhaps special education teachers do not have the depth of knowledge for specific content areas.

- This SPED Director expressed concern that special education teachers are not subject area experts, “But they don’t have the depth of knowledge and the background in social studies or mathematics to be able to provide that same rigorous instruction to special needs students and there’s recognition of that across the district at the secondary level.”

Some teachers are still overwhelmed as illustrated by these statements regarding the paperwork and value of testing for ELL students.

- A teacher shared, “I’d just say it has provided a lot more paperwork and a lot of extra testing for the kids. Special population students go through a lot of testing the way it is and it has increased it a lot.”
- A SPED Director shared concern regarding the true progress of special education students that is not valued, “Well I’m going to talk out of the other side of my mouth now because I really do feel it’s been very good for special education to have a standard and to shoot for that standard. I have to think about other portions of our special education population. Teachers are very discouraged at not getting credit for how far the student has come. And I know that’s talking out of both sides of my mouth.”

It is evident to the researchers that there has been a change of attitudes relative to Special Populations as a result of STARS.

- A high school SPED teacher shared, “I think accountability is a good thing. And as teachers we like to say, oh, yeah, I was accountable before. But it’s a nice way to keep us all teaching the same things at a quality level. And everyone has that opportunity, not just the good students but everyone’s been given that opportunity.”
- A rural superintendent added, “I think it’s probably held them more accountable because their score counts just as much as anybody else. . . . I think it’s made them feel important as part of the school, they’re not treated different than anybody else. They all have to take the same test that everybody else does. And so I think it’s been good for special education for the most part.”
- An elementary principal noted that, “I always felt my special education kids could learn, maybe not at the same rate and maybe not always be on grade level, but I think it’s made us not let kids fall through the cracks.”

Theme 8: Communication

LEADERS and TEACHERS

The interviews revealed a variety of ways that educators are communicating details about the assessment process and STARS results with various public groups across the state.

- A rural male superintendent stated, “I think the community is more aware that we are measuring student performance and are constantly looking at it.”
- Another male rural superintendent shared, “We use our website. We use meetings at the building levels. They [Teachers] meet with the parents. I write a weekly article in the newspaper which covers assessments. We try to hold a press conference when we get data that we want to share. The media’s been very good to us in sharing that so we constantly have it out in front of people.”
- An elementary principal shared, “I think it’s the idea that we’ve got data and we’re expected to share it and we do share it but also we have to kind of enlighten parents as to what this really means and we’re in that process.”
- A rural middle school principal shared, “Well, our superintendent does an excellent job and writes a weekly column and shares a lot of data with the public on how our students are doing.”

Parents are involved in many ways in school districts across the state and learn a great deal about the STARS Process through this involvement.

- An elementary principal said, “We have student-led conferences at the elementary level or students bring their parents in and lead them through that.”
- A rural elementary teacher shared, “We actually have parent/teacher conferences coming up and we have put together a PowerPoint presentation and on one of those slides they will have the actual printout report of the STARS tests that have been taken. The teachers then will be able to show their parents the tests that they have taken based on those standards.”
- A non-rural middle school teacher shared, “We have to educate the parents on the terminology like advanced, proficient, progressive; we have to let them know what those mean. I tend to send grade reports home about every two weeks to inform parents how their kids are doing.”

The media is used to inform parents about a variety of issues related to school.

- A rural elementary principal, “Each one of the administrators writes an article every month that is published. I have chosen to focus my articles on assisting parents to assist their own children. The first article I did was on homework. What can you expect your child to have at the elementary for homework? What can you do as a parent to assist your student in being successful in school? For instance, last month I was focusing on reading tips. And in fact had several internet sites that when your child is a latchkey child, these internet sites would certainly help them to be a better reader. We do things like that.”

Educators are also communicating within their own districts and schools in a variety of ways.

- A rural assessment coordinator said, “One of the things that we’ve implemented this year across the district in each building is professional learning teams. I think that has really made a difference in raising the bar for teachers in the learning process.”

- A middle school principal shared, “That’s just pretty much common dialogue in our building. One of my opening lines when I see teachers is . . . , ‘What’s new and improved in your classroom? What’s going on? What did you try today that you haven’t tried before?’”
- A rural SPED teacher indicated, “We work together not only with the teachers but also with parents and paraprofessionals. I think that gives us a great deal of knowledge in the collaborative process of knowing what other people are doing.”

Educators in schools and at the district level are collaborating both internally and externally in new ways about the STARS Process. During the interviews this year it was more evident that information derived from the STARS Process was being shared with parents and community.

Theme 9: The STARS Process

LEADERS and TEACHERS

Again this year, the theme, STARS Process, emerged from our interviews with educators. Overall there was a positive feeling about the STARS Process, with only a few educators indicating that it is not an effective process.

- A rural curriculum coordinator shared how she had changed her beliefs over time about the STARS process by stating, “What we’ve allowed teachers to do is have authentic conversations about what does this mean? What is this telling us? What can we do to affect the outcome for our kids? And when I first moved here, I said, ‘Oh, just give us a state test. It would be so much easier.’ And I kind of would scratch my head as we’re going though all of this and always thinking that in the back of my head. But I truly believe I have made the leap. I’ve had the paradigm shift that what we’re doing is what we should be doing for kids.”

Districts are becoming more effective about implementing the STARS Process in ways that are meaningful to their schools.

- A rural superintendent stated, “I am very pleased with the STARS assessment because it puts the onus right back where it should be, in the classroom, on the classroom teacher to make the assessments, to determine what prescriptive type things need to be done in the classroom for the student. The rubber hits the road right there in the classroom and those people are in the best position to make those judgments about kids’ progress.”
- A middle school teacher stated, “I think it’s a good process. I think it helps level the playing field for all children so that we ensure that all children are receiving at least a baseline of education. And I think that’s been a problem in the past because teachers were not held accountable. I think as a profession we should hold ourselves accountable.”

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

Study III: 2001-2006 Reading and Math Achievement



**2001 to 2006 Reading and Math Achievement
Grades 4, 8, and 11**

**Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln
Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American
Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹**

¹ Posthumous

INTRODUCTION

When confronted with No Child Left Behind and Average Yearly Progress requirements, every state but Nebraska decided to use norm-referenced or state developed high-stakes measures. In a search for evidence of the positive effects of high-stakes tests on student achievement, Stiggins (2004) found only one study with small gains. Nebraska's School-based Teacher-led Assessment and Reporting System (STARS) is identified by the Partnership for the 21st Century Skills (2005) as "...the nation's most innovative assessment system" (p.13). STARS is being watched closely by national audiences. Most importantly, it is described by a Nebraska school leader as "one of the best things we've done in my 25 years in education."

Nebraska's STARS requires each district to either adopt state standards or develop local standards that are at least equal to or exceed the state standards. Each district then develops a plan for assessing their standards. The plan is based primarily on locally developed criterion-referenced tests (CRT's), which are, therefore, unique to that district. The STARS is reported at fourth, eighth, and eleventh grades. Districts also report Average Yearly Progress (AYP) at grades three through eight and one year in high school. Districts are also required to administer a standardized norm-referenced test (NRT) of their choosing (e.g., Terra Nova, Stanford

Achievement Test) which provides an external common “touch point,” and parts of which may also be used to assess some standards.

PURPOSE OF THE STUDY

The purpose of this study was to examine STARS data available to date, four years each in reading and math. It is a “snapshot” of the percentage of students in Nebraska school districts demonstrating proficiency in these areas. The report includes locally developed criterion-referenced data, norm-referenced data, District Assessment Portfolio data, and quotes from interviews with stakeholders. The research questions were:

1. What was the district average percent of students rated as proficient or advanced in reading on their locally developed criterion-referenced measure and the norm-referenced measure used in that district for 2001, 2003, and 2005 (the three years tested to date for reading)?
2. What changes occurred in the district average percent of students rated as proficient in the criterion and norm-referenced data in reading over these years?
3. What was the district average percent of students rated as proficient in math on their locally developed criterion-referenced measure and the norm-referenced measure used in that district for 2002, 2004, and 2005 (the three years tested to date for math)?
4. What changes occurred in the district percentage of students rated as proficient or advanced in the criterion-referenced and norm-referenced data in math over these years?
5. What was the average rating for the District Assessment Portfolios in reading and math over the three years of available data?
6. What changes occurred in District Assessment Portfolios in reading and math over these years?
7. What were the implications for the STARS program?

RESEARCH DESIGN

Sample

Data was included for Class III, IV, and V school districts. Class III school districts are represented by any school district with a territory having a population of more than 1000 but less than 150,000 inhabitants. Class IV school districts (Lincoln only) are represented by any school district with a territory having a population of 100,000 or more with a city of the primary class. Class V school districts (Omaha only) are represented by any school district within the territory having a population of 200,000 or more inhabitants with a city of the metropolitan class within the territory (Nebraska Education Directory, 2005-2006). The districts in this study represented just over 94% of the public school students in Nebraska. The district data for this study were included on the state website and cooperation for use of the data was facilitated by the Nebraska Department of Education (NDE).

Score Definitions

The criterion-referenced score used was the district average percent of students meeting the proficiency level or better defined by the local district for their locally developed measure. The norm-referenced score is the district average percent of students scoring in the top two quartiles on the nationally standardized test used by that district (e.g., California Achievement Test, Iowa Test of Basic Skills, and Terra Nova). While the norm-referenced measure used varies, the data reported (percent of students in the top two quartiles) is constant for all districts.

Since tests used to measure standards are a mix of locally developed criterion-referenced measures and may include sections of district specific norm-referenced tests, there are few common measures to all districts. It must be remembered that STARS was designed to support instruction in local classrooms, not to facilitate ranking of schools. This strong reliance on district developed criterion-referenced measures challenges traditional validity and reliability views. Therefore, the primary measure of credibility for assessments is a District Assessment Portfolio that is submitted annually to the Nebraska Department of Education.

The Portfolio includes school district ratings on six Quality Criteria that were identified by the Buros Center for Testing (Plake & Impara, 2000), the technical advisors to the STARS program. The Quality Criteria include: 1) the assessment matches the standards, 2) the students have the opportunity to learn, 3) the assessment has been reviewed for bias and sensitivity, 4) the assessment is at the appropriate level, 5) the assessment is reliably scored, and 6) the mastery levels have been appropriately set. Portfolios are rated by an independent measurement expert specifically trained in the rubrics of each of the six Quality Criteria. The Buros Institute also arranges for a panel of external reviewers comprised of professionals with an earned doctorate in educational measurement to complete a rating of district portfolios. The rubric-based ratings on each criterion provide the basis for an overall rating. The overall five-point rating scale ranges from “1,” “Unacceptable,” to “5,” “Exemplary.”

A survey and follow-up interviews were conducted as a part of the STARS evaluation. Quotes from various stakeholders in that process are included in this paper.

Data Analysis

Based on the fact that criterion referenced scores are unique for each district, norm referenced scores vary with the standardized test used by each district, and District Assessment Portfolio ratings are based largely on the criterion referenced measures; the data is described as unconventional. Traditional inferential statistics, therefore, were not appropriate. Descriptive data were reported and discussed.

RESULTS

STARS Reading Achievement

As shown in Table 1, the district average percent of student scores reported by districts as proficient or better in locally defined criterion referenced reading at the fourth-grade level increased from 74.99% in 2001 to 79.39% in 2003, increased to 87.20% in 2005, and increased to 90.70% in 2006. The district average percent proficient for the eighth-grade level increased from 73.67% in 2001 to 74.78% in 2003, increased to 84.49% in 2005, and increased to 87.70% in 2006. The district average percent proficient at the eleventh-grade level increased from 73.54% in 2001 to 74.74% in 2003, increased to 82.26% in 2005, and increased to 86.10% in 2006.

Table 1

Percent Proficient or Higher (Change) on Criterion-Referenced Tests in Reading^a

Grade	2001	2003	2005	2006	Total Change
4	74.99%	79.39% (+ 4.40)	87.20% (+ 7.81)	90.70% (+3.50)	+15.71%
8	73.67%	74.78% (+ 1.11)	84.49% (+ 9.71)	87.70% (+3.21)	+14.03%
11	73.54%	74.74% (+ 1.20)	82.26% (+ 7.52)	86.10% (+3.84)	+12.56%

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state.

Proficiency on criterion-referenced measures increased at all grade levels each year; the average district gain from 2001 to 2006 was 15.71% at fourth grade, 14.03% at eighth grade, and 12.56% at eleventh grade.

Table 2 reports the district average percent of students in the top two quartiles on the norm-referenced reading test used by districts at the fourth grade increased from 64.93% in 2001 to 66.75% in 2003, increased to 67.59% in 2005, and increased to 69.42% in 2006. The eighth grade declined from 62.85% in 2001 to 62.56% in 2003, increased to 63.01% in 2005, and increased to 63.24% in 2006. The eleventh grade increased from 59.87% in 2001 to 61.44% in 2003, increased to 63.67% in 2005, and decreased to 63.59% in 2006. Proficiency, as determined by the percent of students in districts in the top two quartiles on norm-referenced measures, also increased from 2001 to 2006 with a 4.49% increase at fourth grade, a 0.39% increase at eighth grade, and 3.72% gain at eleventh grade.

Table 2

Percent Proficient or Higher (Change) on Norm-Referenced Tests in Reading^a

Grade	2001	2003	2005	2006	Total Change
4	64.93%	66.75% (+ 1.82)	67.59% (+ 0.84)	69.42% (+1.83%)	+4.49%
8	62.85%	62.56% (- 0.29)	63.01% (+ 0.45)	63.24 (+.23%)	+0.39%
11	59.87%	61.44% (+ 1.57)	63.67% (+ 2.23)	63.59 (-.08%)	+3.72%

^a Percent of students scoring in the top two quartiles was calculated for each district and then averaged across the state.

STARS Math Achievement

As shown in Table 3, the district average percent of students reported by districts as proficient or better in locally defined criterion-referenced math at the fourth-grade level increased from 78.29% in 2002 to 85.16% in 2004, increased to 89.00% in 2005, and increased to 90.90% in 2006. The district percent proficient at the eighth-grade level increased from 68.58% in 2002 to 75.34% in 2004, increased to 80.27% in 2005, and increased to 82.90% in 2006. The district percent proficient at the eleventh-grade level increased from 66.22% in 2002 to 72.20% in 2004, increased to 76.24% in 2005, and increased to 80.30% in 2006. Proficiency in criterion-referenced measures increased at all grade levels each year, the gain from 2002 to 2006 at fourth grade was 12.61%, at eighth grade was 14.32%, and at the eleventh grade was 15.08%.

Table 3

Percent Proficient or Higher (Change) on Criterion-Referenced Tests in Math^a

Grade	2002	2004	2005	2006	Total Change
4	78.29%	85.16% (+ 6.87)	89.00% (+ 3.84)	90.90% (+1.9%)	+12.61%
8	68.58%	75.34% (+ 6.76)	80.27% (+ 4.93)	82.90% (+2.63%)	+14.32%
11	66.22%	72.20% (+ 6.98)	76.24% (+ 4.04)	80.30% (+4.06%)	+15.08%

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state.

Table 4 reports the district average percent of students in the top two quartiles on the norm-referenced math test used by districts at the fourth grade increased from 68.12% in 2002 to 71.31% in 2004, increased to 72.05% in 2005, and increased to 73.83% in 2006. The eighth grade declined from 67.34% in 2002, to 66.67% in 2004, increasing to 73.67% in 2005, and declining to 67.83% in 2006. The eleventh grade increased from 67.49% in 2002 to 67.90% in 2004, declined to 67.13% in 2005, and increased to 67.62% in 2006. Proficiency on norm-referenced measures increased at all grade levels however slightly at 8th and 11th grade. The gain from 2002 to 2006 at fourth grade was 5.71%, 0.49% at eighth grade and 0.13% at eleventh grade.

Table 4

Percent Proficient or Higher (Change) on Norm-Referenced Tests in Math^a

Grade	2002	2004	2005	2006	Total Change
4	68.12%	71.31% (+ 3.19)	72.05% (+ 0.74)	73.83% (+1.78%)	+5.71%
8	67.34%	66.67% (- 0.67)	73.67% (+ 7.0)	67.83% (-5.84%)	(+.49%)
11	67.49%	67.90% (+0.41)	67.13% (- 0.77)	67.62% (+.49%)	(+.13%)

^a Percent of students scoring in the top two quartiles was calculated for each district and then averaged across the state.

District Assessment Portfolio Ratings

As reflected in Table 5, the total district average of Reading Assessment Portfolio ratings across grades four, eight, and eleven, on the five-point Likert scale, increased for grade four from 3.57 in 2001 to 4.34 in 2003, increased to 4.55 in 2005, and decreased to 4.54 in 2006, with a total increase of 0.91 from 2001 to 2006. Portfolio ratings for grade eight increased from 3.48 in 2001 to 4.35 in 2003, increased to 4.56 in 2005, and remained the same (4.56) in 2006, with a total increase of 1.08 from 2001 to 2006. Portfolio ratings for grade eleven increased from 3.46 in 2001 to 4.35 in 2003, increased to 4.55 in 2005, and remained the same (4.55) in 2006, with a total increase of 1.09 from 2001 to 2006.

The total district average of Reading Assessment Portfolio rating across grades four, eight and eleven increased from 3.50 in 2001 to 4.35 in 2003, to 4.55 in 2005, and declined to 4.48 in 2006, with a total increase of 3.08 from 2001 to 2006.

Table 5

Reading District Average Portfolio Ratings (Gain/Loss) 2001-2006

Grade	Reading 2001	Reading 2003	Reading 2005	Reading 2006	Reading Total Change
4	3.57	4.34 (+.77)	4.55 (+.15)	4.54 (-.01)	+0.91
8	3.48	4.35 (+.87)	4.56 (+.21)	4.56 (0)	+1.08
11	3.46	4.35 (+.89)	4.55 (+.20)	4.55 (0)	+1.09
Total Dist Av Portfolio Rating (4, 8, 11)	3.50	4.35 (+2.53)	4.55 (+.56)	4.48 (-.01)	+3.08

As reflected in Table 6, the total district average for Math Assessment Portfolio ratings across grades four, eight, and eleven, on the five-point Likert scale, increased for grade four from 3.98 in 2002, increased to 4.67 in 2004, declined to 4.57 in 2005, and remained the same (4.57) in 2006, with an increase of 0.59 from 2002 to 2006. Portfolio ratings for grade eight increased from 3.96 in 2002 to 4.77 in 2004, declined to 4.66 in 2005, and declined to 4.65 in 2006, with a total increase of 0.69 from 2002 to 2006. Grade eleven portfolio ratings increased from 3.96 in 2002 to 4.77 in 2004, declined to 4.60 in 2005, and increased to 4.61 in 2006.

The total district average of Math Assessment Portfolio rating across grades four, eight and eleven increased from 3.97 in 2002, increased to 4.74 in 2004, declined to 4.61 in 2005, and remained at 4.61 in 2006, with a total increase of 1.93 from 2002 to 2006.

Table 6

Math District Average Portfolio Ratings (Gain/Loss) 2001-2006

Grade	Math 2002	Math 2004	Math 2005	Math 2006	Total Change
4	3.98	4.67 (+0.69)	4.57 (-.10)	4.57 (0)	(+0.59)
8	3.96	4.77 (+0.81)	4.66 (-.11)	4.65 (-0.01)	(+0.69)
11	3.96	4.77 (+0.81)	4.60 (-.17)	4.61 (+0.01)	(+0.65)
Total Dist Av Portfolio Rating (4, 8, & 11)	3.97	4.74 (+2.31)	4.61 (-.38)	4.61 (0)	(+1.93)

SUMMARY

District criterion-referenced measures continue to show growth over time in the areas of reading and math from 2001 to 2006. District norm-referenced measures have generally increased in reading for fourth, eighth, and eleventh grades. District math norm-referenced measures generally increased for grades four, eight, and eleven. School improvement with student academic achievement as the goal is not a short-term process. Nebraska is in its sixth year of full implementation of the program and variability still exists in achievement in some areas.

REFERENCES

Nebraska Department of Education. (2003-2004). *Nebraska Department of Education Directory*. Lincoln, NE. 106th Edition, 170.

Partnership for 21st Century Skills. (2005). *The Road to 21st century learning: A policymaker's guide to 21st century skills*. Retrieved on May 4, 2005 from <http://www.21stcenturyskills.org/route/>

Plake, B. C., & Impara, J. C. (2000). *Technical Quality Criteria for Nebraska's District Assessments*. Lincoln, NE: Buros Center for Testing.

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22-27.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

Study IV: 2002-2006 Writing Achievement



**2002-2006 Nebraska's STARS Statewide
Writing Achievement**

**Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln
Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American
Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹**

¹ Posthumous

INTRODUCTION

With the emphasis on education to demonstrate increased achievement for all students and all schools to be accountable to their constituents, states have developed or adopted state-mandated tests to assess student academic performance. Nebraska made a decision to implement a system of accountability that enhances student performance by focusing upon building assessment literacy among educators. In the Nebraska School-based Teacher-led Assessment and Reporting System (STARS), districts first adopt local or state standards. Districts then submit an assessment plan that may include locally developed criterion-referenced tests (CRTs) to assess the district's standards at the identified grade levels. Districts may also choose to use items from their locally chosen norm-referenced test (NRT) to assess some of their standards.

Based on previous involvement by a number of Nebraska school districts in the Six Trait writing model and the natural link of this criterion-referenced approach with the emerging philosophy of Nebraska STARS, a requirement for a statewide writing assessment was included in the

legislation establishing Nebraska's assessment system. The merits, or lack thereof, of statewide writing assessments have been an area of some debate. Concerns include such areas as reducing writing to a formulaic approach, narrowing content in the teaching of writing and teacher time spent on scoring rather than the instruction of writing (Hillock, 2003; Mabry, 1999).

Spandel and Stiggins (1997) looked at an important way to bridge the gap between large-scale assessment and classroom instruction in their discussion of large-scale assessments that enlisted classroom teachers as raters. They maintained that teachers who assessed students' work in the large-scale setting gave themselves an education in both how to write and how to assess. They further argued that teachers who participated as raters in large-scale assessment gained a broader perspective of student writing from throughout a state or district, which was very different from seeing only the writing of their students. But they cited that the primary contribution large-scale writing assessment had made to writing instruction, with the exception of scoring criteria that promoted the reliability of scoring, was a higher expectation for student performance. They also maintained that a state writing assessment, perhaps more than any other single stimulus, prompted decision makers to ask important questions that they may not have asked before, questions that examined when, where, and how writing was being taught.

PURPOSE OF THE STUDY

The purpose of this study was to examine the district achievement data available for the Statewide Writing Assessment for the Nebraska STARS program. The research questions for this study were:

1. What was the district average percent of students rated proficient or better in grades four, eight, and eleven for years 2002-2006 on the Nebraska Statewide Writing Assessment?
2. What were the changes over the years 2002-2006 in the district average percent of students rated as proficient or better in district writing scores at grades four, eight, and eleven?
3. What were the teacher perceptions of the writing assessment?
4. What were the implications for the Nebraska STARS program?

RESEARCH DESIGN

Districts Included

Data was included for Class III, IV, and V school districts. Class III school districts are represented by any school district with a territory having a population of more than 1000 but less than 150,000 inhabitants. Class IV school districts (Lincoln only) are represented by any school district with a territory having a population of 100,000 or more with a city of the primary class. Class V school districts (Omaha only) are represented by any school district within the territory having a population of 200,000 or more inhabitants with a city of the metropolitan class within the territory (Nebraska Education Directory, 2005-2006). The districts in this study represented just over 94% of the public school students in Nebraska. The district data for this study were included on the state website and cooperation for use of the data was facilitated by the Nebraska Department of Education (NDE).

Statewide Writing Assessment Prompt Development

The process for development of writing prompts for use in the Statewide Writing Assessment relied on the involvement of Nebraska classroom teachers. Participating teachers were recommended by their district superintendent or assessment contact person and selected by the NDE each year to take part in a writing development task force. The task force consisted of three panels, each consisting of 10-15 teachers representing grades four, eight, and eleven from a variety of school sizes and geographic regions. The task force was convened for a one-day workshop facilitated by the NDE for the purpose of:

- Reviewing the characteristics of mode-specific writing
- Learning the criteria for effective writing prompts
- Reviewing and examining areas of bias to be avoided
- Creating writing prompts for field testing

During the workshop, participants read and discussed examples of current research related to best practices in the teaching and assessment of student writing. Participants also read about and discussed criteria for effective writing prompts as well as issues related to bias that should be avoided when creating writing prompts. A number of examples of writing prompts including those that had been used in previous Nebraska statewide writing assessments were also reviewed.

Statewide Writing Assessment Prompts Field Testing Process

From information gathered at the Writing Prompt Development workshop, school districts representing various sizes and geographic locations were selected to field test the writing prompts with students in grades four, eight, and eleven before the end of the current school year. Multiple prompts were field tested across multiple school districts.

Participants in the field-testing were at the appropriate grade-level and completed assessments according to standard administration procedures. The size of the student sample selected for the field-testing was adequate to provide responses sufficient for scoring and anchoring purposes. At the conclusion of the field-testing, the NDE conducted a review to “fine-tune” the Statewide Writing Assessment scoring process.

Statewide Writing Assessment Assessors

Nebraska teachers were recruited by the NDE to score the writing assessment each year. The scorer qualifications included:

- (a) The teacher was currently teaching or had taught at or near the grade-level being assessed.
- (b) The teacher was familiar with student writing at the grade-level being assessed.
- (c) The teacher had basic knowledge of the Six Trait writing assessment model.

Statewide Writing Assessment Scoring Process

Scoring of the state assessment was held at a central location in the state and scorers came to the site for three days during which training and the scoring occurred. The scoring process of Nebraska’s Statewide Writing Assessment required each sample of student writing to be read and scored by two trained teacher raters who assigned a single holistic score within allowable ranges as prescribed by the rubric. The rubric criteria were ideas and content, organization, voice or tone, word choice, sentence fluency, and conventions as identified in the Nebraska Content

Standards (NDE). Raters assigned a score based on how the writing met these criteria overall. If there was more than a two-point difference, a third scoring was done. The scoring range was from one to four in + and – intervals resulting in a ten-point scale. The final score was the composite of the two individual scores. The NDE contracts with the Buros Center for Testing to establish the statewide cut-score.

In the first three years, scoring was done at three sites across the state. To improve reliability, scoring is now done at one site. A sample was sent out of state for scoring by an independent contracted testing company. The Nebraska Department of Education releases results for the statewide writing assessment and all Nebraska STARS assessments on their website each fall. Local district and individual school data shared includes the district average percent of students meeting proficiency or better on the Statewide Writing Assessment.

Data Analysis

The unit of analysis for this study was the district average percent of students rated as proficient in Class III, IV, and V school districts for the State of Nebraska in writing at grades four, eight, and eleven. While this statewide assessment took on some formal technical assessment characteristics that would more characterize norm-referenced tests (statewide common administration and scoring, common cut-score) than many criterion-referenced assessments, it was clearly not a comparison with a separate norm group. Descriptive data was, therefore, reported and discussed. However, because the assessment was a common measure across districts and was an equal interval scale, inferential statistics were also used to examine statistical significance between pre/post scores from inception to last scoring. All significance tests were two-tailed.

RESULTS

Ratings of Writing Proficiency

Table 7 indicates that gains were made at nearly all grades and years of comparison. At fourth grade, gains from 2002 to 2004, and 2004 to 2005, were significant ($p < .001$) with a decline indicated in 2006. At eighth grade, the gain from 2003 to 2004 was significant ($p < .001$); the gain from 2004 to 2005 was not significant. There was an increase in 2006 to 87.00%. The gain in eleventh grade from 2004 to 2005 was significant ($p < .05$) with a slight gain in 2006.

Writing scores increased for all grades, with grade four increasing 6.91% from 2002 to 2006, grade eight increasing 7.45%, and grade eleven increasing 1.88%.

Table 7

Statewide Writing Assessment 2002-2006: Mean District Percent of Student Scores at the Proficient Level or Higher^a

Grade	2002	2003	2004	2005	2006	Change 2006 vs. Baseline
4	76.50%		80.89%**	84.57%**	83.41%	+6.91%
8		79.55%	85.44%**	86.31%	87.00%	+7.45%
11			89.22%	91.02%*	91.91%	+1.88%

* $p < .05$ compared to the previous year

** $p < .001$ compared to the previous year

† $p < .05$ compared to the baseline year

‡ $p < .001$ compared to the baseline year

Note: 2002, 2003, and 2004 were baseline years for grades 4, 8, and 11 respectively

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state.

Teacher Perceptions

In “Charting STARS – Sustainability as Challenge and Opportunity,” Gallagher (2003) reported the results of a second year of a research study and comprehensive evaluation of Nebraska’s School-based Teacher-led Assessment and Reporting system (STARS). Among the major findings of the analyses of a survey administered to teachers on their perceptions and classroom practices related to the state writing assessment were: 69% of all teachers placed more emphasis on practice writing assessments, 73% placed more emphasis on sharing assessment criteria in class, and 73% placed more emphasis on explicit instruction in six trait writing.

In addition to these findings, Gallagher (2003) reported that 88% of teachers agreed or agreed strongly that the six traits scoring rubric used to score the state writing assessment was useful for instruction; 75% agreed or strongly agreed that the state writing assessment supported learning objectives they have for their students; 72% agreed or strongly agreed that the results of the state writing assessment were useful for teachers; and 65% agreed or strongly agreed that the six traits were the most important features of writing.

SUMMARY

The purpose of this study was to examine district achievement data available for the Statewide Writing Assessment. Results indicated that fourth, eighth, and eleventh graders made gains in the pre/post comparisons on the Statewide Writing Assessment. The writing data, which compared fourth grade data from 2002 through 2006, and eighth grade data from 2003 through 2006, revealed overall gains at both grades. This may reflect the involvement of many school districts with Six Trait Writing preceding the formal statewide writing assessment process. Students at the eighth and eleventh grades continued to show growth on their writing scores from 2004 to 2006. However, fourth-grade scores dropped slightly in 2005 to 2006 with an overall gain from 2002 to 2006 of 6.91%. While this writing assessment has characteristics that enable inferential statistical analysis to be used, there may be some question from the traditional measurement community concerning this practice. It must be remembered that the philosophy and purpose of Nebraska STARS is to support teaching and learning, and not to focus on development of assessments of technical strengths in ranking results.

REFERENCES

- Gallagher, C.W. (2004).). *Charting STARS: Sustainability as Challenge and Opportunity: Year-three report to the Nebraska Department of Education*. Lincoln: University of Nebraska.
- Hillocks, G. (2003). *The testing trap: How state writing tests control learning*. New York: Teachers College Press.
- Mabry, L. (1999). Writing to the rubric. *Phi Delta Kappan*. 9(9), 673-679.
- Nebraska Department of Education. (2003-2004). *Nebraska Department of Education Directory*. Lincoln, NE. 106th Edition, 170.
- Nebraska Department of Education. (NDE). Nebraska standards in reading/writing. Retrieved on March 3, 2006, from http://www.nde.state.ne.us/ndestandards/academic_standards.htm

Roschewski, P., Gallagher, C. & Isernhagen, J. (2001). Nebraskans reach for the STARS. *Phi Delta Kappan*. 78(8), 611-615.

Spandel, V. & Stiggins, R. (1997). *Creative writers: Linking writing assessment and instruction*. New York: Addison Wesley Longman.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

**Study V: 2001-2006 Achievement for Special
Populations**



2001-2006 Achievement for Special Populations

**Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln
Dr. Shirley Mills, Assistant Professor, University of Texas-Pan American
Dr. Leon Dappen, Associate Professor, University of Nebraska-Omaha¹**

¹ Posthumous

INTRODUCTION

A key component of No Child Left Behind is to demonstrate increased academic achievement for all students. States have taken different approaches to accomplish this end. Nebraska has built a School-based Teacher-led Assessment and Reporting System (STARS), which requires Nebraska School Districts to develop a local assessment system to measure student performance on local standards. There is no single high-stakes test. The STARS has been in place for six years and has a strong success record with total group data (Roschewski, Isernhagen, & Dappen, 2006).

PURPOSE OF STUDY

The purpose of this study was to examine the academic change of special populations, in this case, English Language Learners (ELL) and Special Education (SPED) students. The questions for this study were:

1. What were the changes in district average percent of students rated as proficient or higher in reading for 2001-2006 on their locally developed criterion-referenced tests for ELL and SPED students?
2. What were the changes in district average percent of students rated as proficient or higher in math for 2002-2006 on their locally developed criterion-referenced tests for ELL and SPED students?
3. What were the changes in district average percent of students rated as proficient or higher in writing on the statewide criterion-referenced writing assessment for ELL and SPED students?
4. What were the implications to the Nebraska STARS program from these findings?

RESEARCH DESIGN

Districts Included

Data was included for Class III, IV, and V school districts. Class III school districts are represented by any school district with a territory having a population of more than 1000 but less than 150,000 inhabitants. Class IV school districts (Lincoln only) are represented by any school district with a territory having a population of 100,000 or more with a city of the primary class. Class V school districts (Omaha only) are represented by any school district with a territory having a population of 200,000 or more inhabitants with a city of the metropolitan class within the territory (Nebraska Education Directory, 2005-2006). The districts in this study represented just over 94% of the public school students in Nebraska. The district data for this study were included on the state website and cooperation for the use of the data was facilitated by the Nebraska Department of Education.

Score Definitions

The criterion-referenced score (CRT) for reading and math was the district average percentage for ELL and SPED students meeting the proficiency level or better defined by the local district for their locally developed measure in classes III, IV, and V school districts for the state of Nebraska.

The criterion referenced score (CRT) for writing was the district average percent of ELL and SPED students rated as proficient in classes III, IV, and V school districts for the state of Nebraska in writing at grades four, eight, and eleven. Because the assessment was a common measure across districts and was an equal interval scale, inferential statistics were also used to examine statistical significance between pre/post scores from inception to the last scoring.

Data Analysis

Criterion referenced scores for reading and math were unique for each district, therefore not on a common scale and did not support common and inferential statistics. Descriptive data only was reported and discussed for reading and math.

The writing assessment was a common measure across districts and was an equal interval scale. Therefore, inferential statistics were used to examine statistical significance between pre/post scores in writing. All significance tests were two-tailed.

RESULTS

STARS Achievement for English Language Learners

ELL Reading Achievement

The district average percent of English Language Learners (ELL) reported as proficient or higher in locally defined criterion-referenced assessments for reading at the fourth-grade level decreased from 50% in 2001 to 49% in 2003, increased to 67% in 2005, and increased to 72% in 2006. As shown in Table 8, proficiency on reading criterion-referenced measures for grade four increased 22% from 2001 to 2006.

The district average percent of ELL scores for eighth-grade students decreased from 47% in 2001 to 42% in 2003, increased to 57% in 2005, and increased to 60% in 2006. Proficiency on reading criterion-referenced measures for eighth-grade ELL students increased 13% from 2001 to 2006 as shown in Table 8.

The district average percent of ELL scores for eleventh-grade students decreased from 45% in 2001 to 32% in 2003, increased to 47% in 2005, and increased to 53% in 2006. Proficiency on reading criterion-referenced measures for eleventh-grade ELL students increased 8% from 2001 to 2006 as shown in Table 8.

Therefore, the district average percent for ELL students declined from the first year (2001) to the second year (2003) for all grade levels, but increased from 2001-2006 overall.

Table 8

District Average Reading Criterion-Referenced Assessments English Language Learners (ELL) 2001-2006^a (District Sample Size – Number of districts reporting from population)

ELL Grade	District Mean Percent Proficient				Change in District Mean Percent Proficient			
	2001	2003	2005	2006	2001-03	2003-05	2005-06	2001-06
4	50%	49%	67%	72%	-1%	+18%	+5%	+22%
8	47%	42%	57%	60%	-5%	+15%	+3%	+13%
11	45%	32%	47%	53%	-13%	+15%	+6%	+8%

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state. All percentages were rounded to whole percents.

ELL Math Achievement

The district average percent of English Language Learners (ELL) reported as proficient or better in locally defined criterion-referenced assessments for math at the fourth-grade level increased from 53% in 2002 to 70% in 2004, increased to 72% in 2005, and increased to 80% in 2006. Proficiency on math criterion-referenced measures increased 27% from 2002 to 2006 as shown in Table 9.

Table 9

District Average Math Criterion-Referenced Assessments English Language Learners (ELL) 2002-2006^a (District Sample Size - Number of districts reporting from population)

ELL Grade	District Mean Percent Proficient				Change in District Mean Percent Proficient			
	2002	2004	2005	2006	2002-04	2004-05	2005-06	2002-06
4	53%	70%	72%	80%	+17%	+2%	+8%	+27%
8	40%	43%	59%	61%	+3%	+16%	+2%	+21%
11	39%	48%	51%	48%	+9%	+3%	-3%	+9%

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state. All percentages were rounded to whole percents.

The district average percent of ELL scores for eighth-grade students increased from 40% in 2002 to 43% in 2004, increased to 59% in 2005, and increased to 61% in 2006. Proficiency on math criterion-referenced measures for eighth-grade ELL students increased 21% from 2002 to 2006 as also shown in Table 9.

The district average percent of ELL scores for eleventh-grade students increased from 39% in 2002 to 48% in 2004, increased to 51% in 2005, and decreased to 48% in 2006. Proficiency on math criterion-referenced measures for eleventh-grade ELL students increased 9% from 2002 to 2006 as shown in Table 9.

ELL Writing Achievement

The district average percent of English Language Learners (ELL) reported as proficient or better on the state criterion-referenced assessment for writing at the fourth-grade level increased from 49% in 2002 to 52% in 2004, increased to 64% in 2005, and increased to 66% in 2006. Proficiency on the state writing criterion-referenced measure increased 17% from 2002 to 2006. These increases were not statistically significant. This is shown in Table 10.

The district average percent of ELL scores for eighth-grade students increased from 37% in 2003 to 56% in 2004, increased to 60% in 2005, and decreased to 56% in 2006. Proficiency on the state writing criterion-referenced measure for eighth-grade ELL students increased 19% from 2003 to 2006 as also shown in Table 10.

The district average percent of ELL scores for eleventh-grade students decreased from 45% in 2004 to 44% in 2005, and increased to 53% in 2006. Proficiency on the state writing criterion-referenced measure for eleventh-grade ELL students increased 8% from 2004 to 2006 as also shown in Table 10.

Table 10

District Average Writing Criterion-Referenced Assessments English Language Learners (ELL) 2002-2006^a (District Sample Size - Number of districts reporting from population)

Grade	District Mean Percent Proficient					Change in District Mean Percent Proficient				
	2002	2003	2004	2005	2006	2002-03	2003-04	2004-05	2005-06	2002-06
4	49%		52%	64%	66%	+3%		+12%	+2%*	+17%
8		37%	56%	60%	56%		+19%*	+4%	-4%	+19%
11			45%	44%	53%			-1%	+9%*	+8%

*p<.05 compared to the previous year

♦ p<.05 compared to the baseline year

** p<.01 compared to the previous year

♦♦ p<.01 compared to the baseline year

*** p<.001 compared to the previous year

♦♦♦ p<.001 compared to the baseline year

Note: 2002, 2003, and 2004 were baseline years for grades 4, 8, and 11 respectively

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state. All percentages were rounded to whole percents.

STARS Achievement for Special Education Students

SPED Reading Achievement

The district average percent of Special Education (SPED) students reported as proficient or better on locally defined criterion-referenced assessments for reading at the fourth-grade level increased from 44% in 2001 to 52% in 2003, increased to 67% in 2005, and increased to 74% in 2006. Proficiency of SPED fourth-grade students on reading criterion-referenced measures increased 30% from 2001 to 2006 as shown in Table 11.

The district average percent of SPED eighth-grade students remained the same from 2001 to 2003 (43%), increased to 59% in 2005, and increased to 66% in 2006. Proficiency on reading criterion-referenced measures for eighth-grade SPED students increased 23% from 2001 to 2006 as also shown in Table 11.

The district average percent of SPED scores for eleventh-grade students decreased from 42% in 2001 to 37% in 2003, increased to 54% in 2005, and increased to 61% in 2006. Proficiency on reading criterion-referenced measures for eleventh-grade SPED students increased 19% from 2001 to 2006 as shown in Table 11.

Therefore, the district average percent for SPED students increased for all grade levels from 2001-2006.

Table 11

District Average Reading Criterion-Referenced Assessments Special Education (SPED) Students 2001-2006^a (District Sample Size - Number of districts reporting from population)

Grade	District Mean Percent Proficient				Change in District Mean Percent Proficient			
	2001	2003	2005	2006	2001-03	2003-05	2005-06	2001-06
4	44%	52%	67%	74%	+8%	+15%	+7%	+30%
8	43%	43%	59%	66%	+0%	+16%	+7%	+23%
11	42%	37%	54%	61%	-5%	+17%	+7%	+19%

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state. All percentages were rounded to whole percents.

SPED Math Achievement

The district average percent of Special Education (SPED) students reported as proficient or better on locally defined criterion-referenced assessments for math at the fourth-grade level increased from 51% in 2002 to 62% in 2004, increased to 72% in 2005, and increased to 75% in 2006. Proficiency on math criterion-referenced measures increased 24% from 2002 to 2006 as shown in Table 12.

The district average percent of SPED eighth-grade students increased 34% from 2002 to 44% in 2004, increased to 54% in 2005, and increased to 56% in 2006. Proficiency on math criterion-referenced measures for eighth-grade SPED students increased 22% from 2002 to 2006 as also shown in Table 12.

Table 12

District Average Math Criterion-Referenced Assessments Special Education (SPED) Students 2002-2006^a (District Sample Size- Number of districts reporting from population)

SPED Grade	District Mean Percent Proficient				Change in District Mean Percent Proficient			
	2002	2004	2005	2006	2002-04	2004-05	2005-06	2002-06
4	51%	62%	72%	75%	+11%	+10%	+3%	+24%
8	34%	44%	54%	56%	+10%	+10%	+2%	+22%
11	28%	36%	43%	46%	+8%	+7%	+3%	+18%

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state. All percentages were rounded to whole percents.

The district average percent of SPED scores for eleventh-grade students increased from 28% in 2002 to 36% in 2004, increased to 43% in 2005, and increased to 46% in 2006. Proficiency on math criterion-referenced measures for eleventh-grade SPED students increased 18% from 2002 to 2006 as shown on Table 12.

Therefore, the district average percent for SPED students increased for all grade levels from 2002-2006.

SPED Writing Achievement

The district average percent of Special Education (SPED) students reported as proficient or better on the state criterion-referenced assessment for writing at the fourth-grade level increased from 46% in 2002 to 55% in 2004, increased to 65% in 2005, and decreased to 60% in 2006 as shown in Table 12. Proficiency on the state writing criterion-referenced measure increased 14% from 2002 to 2006 as shown in Table 13.

The district average percent of SPED eighth-grade students increased from 48% in 2003 to 55% in 2004, increased to 61% in 2005, and remained at 61% in 2006 as shown in Table 12. Proficiency on the state writing criterion-referenced measure for eighth-grade SPED students increased 13% from 2003 to 2006 as shown in Table 13.

The district average percent of SPED scores for eleventh-grade students increased from 55% in 2004 to 63% in 2005, and increased to 64% in 2006. Proficiency on the state writing criterion-referenced measure for eleventh-grade SPED students increased 9% from 2004 to 2006 as shown in Table 13.

In summary, the district average percent proficient for SPED fourth-grade students increased 14% from 2002-2006. The district average percent proficient for SPED eighth-grade students increased 13% from 2003-2006. The district average percent for SPED eleventh-grade students increased 9% from 2004-2006 as shown in Table 13.

Table 13

District Average Writing Criterion-Referenced Assessments Special Education (SPED) Students 2002-2006^a

(District Sample Size- Number of districts reporting from population)

SPED Grade	District Mean Percent Proficient					Change in District Mean Percent Proficient				
	2002	2003	2004	2005	2006	2002-04	2003-04	2004-05	2005-06	2002-06
4	46%		55%*	65%***	60%	+9%		+10%	-5%	+14%
8		48%	55%**	61%	61%		+7%	+6%	0%	+13%
11			55%	63%**	64%			+8%	+1%	+9%

*p<.05 compared to the previous year

** p<.01 compared to the previous year

***p<.001 compared to the previous year

♦ p<.05 compared to the baseline year

♦♦ p< .01 compared to the baseline year

♦♦♦ p<.001 compared to the baseline year

Note: 2002, 2003, and 2004 were baseline years for grades 4, 8, and 11 respectively

^a Percent of students scoring proficient or higher was calculated for each district and then averaged across the state. All percentages were rounded to whole percents.

SUMMARY

While the ELL and SPED student district average scores were not as strong as the total group district averages, they were increasing in most instances. This is consistent with most research in this area and, indeed, the basis for the special supports provided for ELL and SPED students. The reading scores for SPED students reflected strong positive change in all areas. ELL students for reading at the fourth and eighth grade were somewhat stronger than for the eleventh-grade students. This trend could relate to the difficulty of older ELL students when mastering English as a second language.

STARS has impacted the teaching of all students in Nebraska and especially those for special populations. An elementary principal shared the change in emphasis for expectations for special populations, "But by the same token, being the old special education teacher, I always felt that my special education students could learn, maybe not at the same rate, and maybe not always beyond grade level. I think it's made us not let kids fall through the cracks." An assessment coordinator was also aware of the change in focus upon all students, especially those in special populations. She stated, "They are going to find success. We just had an IEP meeting last night when the special education teacher was saying, 'We're going to use custom testing for this child and they're not going to have to take all fifteen questions at one time. We're going to break them down into the different areas being stressed,' so I feel like we're addressing that." In summary, STARS has impacted the learning of all students in Nebraska in positive ways.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

**Study VI: Use of Data to Inform Decisions by
Elementary Principals**



Use of Data to Inform Decisions by Elementary Building Principals

Jadi Miller, Principal,
Everett Elementary School, Lincoln, NE

INTRODUCTION

Unlike many of its counterparts in other states, the state of Nebraska has elected to implement an accountability system that relies heavily on the professional work of educators across the state. This has resulted in each individual school district working to identify essential learning targets, either through the state's standards or the development of their own, to develop an assessment system for measuring student achievement relevant to those learning targets, and to report the results of those assessments to the state. As a result of these requirements, Nebraska's system required high levels of participation by teachers and administrators in each school district. Local educators played a much more active role in the accountability system in Nebraska; they were not simply the distributors of the tests (Roschewski, Gallagher, and Isernhagen, 2001; Gallagher, 2004). Rather than implementing a single statewide assessment that is administered to all students, the School-based Teacher-led Assessment and Reporting System (STARS) required individual school districts to write assessments that align to the state standards (LEARNs) and to have those assessments meet rigorous quality criteria. Central to this process of developing assessments was the building principal (Roschewski, et al, 2001). Nebraska chose to maintain a focus on the classroom and at the school-level, and included the professional development of both teachers and administrators to build both the assessment literacy and the instructional leadership necessary to implement a school-based system (Roschewski, et al, 2001; Roschewski, 2003; Roschewski, Isernhagen, and Dappen, 2006).

The Nebraska system is focused on school improvement (Roschewski, et al, 2001). It was designed to encourage all educators to engage in conversations about expectations for student learning, measurement of that learning, and the use of that knowledge to impact classroom instruction (Roschewski, et al, 2001). Due to the varying sizes of school districts within the state of Nebraska, one of the only constants is the presence of a building administrator. As a result, building principals have had to take a much more active role in assessing and reporting student achievement. Despite the variety of sizes of schools and districts, the focus of the STARS continues to emphasize the importance of using information generated by the accountability system to impact what happens at the classroom level (Roschewski, et al, 2001; Roschewski, 2003; Roschewski, et al, 2006).

To help clarify for educators the nature of effective leadership and adult learning, one professional organization, the National Staff Development Council (NSDA), established a set of standards that guide the design of professional learning within schools. These standards (NSDC, 2001) identified twelve different areas of best practice which schools can use to analyze their own practices. One of these was the use of data to guide decisions. The NSDC has also published a document that assists schools with the implementation of the standards. *Moving NSDC's Staff Development Standards Into Practice: Innovation Configurations* (Roy and Hord, 2003) took those standards and defined them for each member of the school community, such as teachers, community members, central office administrators, and building administrators. The National Staff Development Council's standards were selected to provide a concise definition of data-driven decision-making through specific types of behaviors exhibited by building principals.

In an effort to gain more information about the use of data to impact student achievement, a web-based survey for elementary principals was designed. The survey was distributed through email to every elementary principal in the state of Nebraska, based on the identification of individuals for each school building and/or district provided to the Nebraska Department of Education by local school districts. The 60-question survey was organized around five research questions, based on the five desired outcomes identified by NSDC related to data-driven decisions by a building principal.

PURPOSE OF STUDY

The purpose of this quantitative study was to examine the relationship between data-driven decision-making practices of Nebraska elementary school principals and the professional standards written by the National Staff Development Council (NSDC). Data-driven decision-making was defined, according to the NSDC standards and the *Innovation Configurations*, to include data analysis, goal setting, engagement of others, design of professional development, and monitoring of progress. Elementary principals were surveyed about their activities in these areas.

Research Questions

The research questions that guided the study focused on describing the behaviors of elementary principals in relation to data-driven decision-making as defined by the NSDC Standards for Staff Development. The research question and sub-questions included:

1. To what degree do Nebraska elementary principals follow the National Staff Development Council's Staff Development Standards for data-driven decision-making?
 - a. To what degree was the entire school faculty involved in the analysis of student achievement data?
 - b. To what degree does the principal disaggregate student data to identify goals?
 - c. To what degree were teachers, parents, and community members engaged in data-driven decision-making?
 - d. To what degree was data used to determine professional development?
 - e. To what degree was data used to monitor the school improvement goals?

RESEARCH DESIGN AND METHODOLOGY

This study was designed as a descriptive quantitative study to analyze data-driven decision-making practices of elementary principals in Nebraska. Data was collected through the use of a web-based survey developed by the researcher. Web-based surveys can be designed to use a variety of tools and capabilities that are unique among self-administered questionnaires. The web-based surveys also make it possible to reach large numbers of respondents in a single contact (Dillman, 2000). The researcher used survey software entitled "Zoomerang" (<http://www.zoomerang.com/>) as the web-based survey tool.

The cross-sectional survey allowed the researcher to gain information from a population of elementary principals within the state of Nebraska. The cross-sectional survey collects data from the respondents at one point in time (Creswell, 2003). The cross-sectional design is used to conduct survey research about attitudes, beliefs, or practices (Creswell, 2005).

The survey (Appendix J) was organized into six main sections to answer the proposed research questions and sub-questions. The first section asked questions about how the principal has involved other members of the school community in the analysis of data. The second section asked questions about the types of data used in that process and the level of disaggregation of that data. The third section sought information about how the principal engages others, such as parents and community members in the process of analyzing data. The fourth section addressed how the principal used or did not use the student achievement data to create a professional development plan for the school. The fifth section addressed if and how the principal used student achievement data to monitor the school's improvement efforts. Finally, the last section asked demographic questions about the principal's experience as an administrator and the size of the school and district.

Population

The survey population consisted of all elementary principals employed in the state of Nebraska during the 2006-2007 school year. This group was approximately 521 individuals who have either sole responsibility as an elementary principal or who serve in that role in addition to other duties within a school district. Participants received an email about the nature of the survey. It contained a direct link to the survey, which began with an informed consent form. Once the participant acknowledged the informed consent by agreeing to participate in the web-based survey, the survey program connected him/her to the actual survey instrument. The participants had two weeks to complete the survey. The e-mail addresses of all elementary principals in Nebraska were obtained from a variety of sources including the Nebraska Department of

Education database, the Nebraska Association of Elementary School Principals, and school district contacts. Follow-up and reminder e-mails were sent to non-responders, generated by the web-based survey program to protect anonymity.

Elementary principals were selected for this study because of their role in both school improvement and as an instructional leader within the school setting. All elementary principals employed in Nebraska during the 2006-2007 school year were contacted to participate in the study. There are several factors that can create obstacles to valid inferences (Creswell, 2005). One of these is coverage error, or a limited sample. By obtaining lists of elementary principals from multiple sources, a more accurate contact list can be created. Sampling error occurs with a sample that is too small. As a result, this study included all elementary principals in the state of Nebraska. Another source of error is non-response to the survey. The use of a web-based survey instrument allowed the researcher to provide reminders to non-responders without jeopardizing the confidentiality of respondents (Creswell, 2005). Responses were received from 168 elementary principals (32% return rate). This response rate is similar to those in other web-based surveys (Dillman, 2000).

RESEARCH RESULTS

Research Question 1a: To what degree was the entire school faculty involved in the analysis of student achievement data?

The first section in the survey included questions about the extent to which the school faculty was involved in the analysis of student achievement data to identify school improvement goals. The section included five questions on a five-point Likert scale: “Strongly agree” (5), “Agree” (4), “Disagree” (3), “Strongly disagree” (2), and “No opinion” (1). The first two questions featured a reverse scale, as strong agreement to the statements indicated that the principal did not involve staff members in the data analysis, and described a less-desirable behavior on the part of the principal. The total scores could range from five to 25, with a higher score indicating a stronger alignment with the NSDC’s standards for a data-driven principal. The following table shows the mean, mode, standard deviation, and variance for these items.

Table 14

Degree of Entire School Faculty Involvement in Analysis of Student Achievement Data

Section	Mean	Mode	Standard Deviation	Variance
1	21.27	21	2.09	4.36

The scores for the principals ranged from a low of 14 to a high of 25. The mean score for the first five questions was 4.27, which would place the group between “Agree” and “Strongly agree.” A question in this section asked respondents about the level of support they received for data analysis. The choices included from “was facilitated entirely by me,” “was facilitated with assistance from others in the building,” “was facilitated with assistance from others in the district,” “was facilitated by someone in the district other than myself and/or someone from the ESU,” or “did not take place in my building.” The respondents indicated most often that the process was facilitated with assistance from others in the building and with assistance from others in the district, each were selected 37% of the time. When asked about the support received from the Educational Service Unit (ESU), 57% of the respondents indicated that the ESU had provided training in data analysis, 41% indicated that the ESU had provided

consultation, 40% of the respondents said that the ESU had hosted a data retreat for schools in the area, and 39% indicated the ESU had provided on-site consultation with the school’s staff members.

Research Question 1b: To what degree does the principal disaggregate student data to identify goals?

Another section of the survey specifically addressed the level of disaggregation that was present in the student data used to identify school improvement goals. The NSDC standards provided descriptions of the various levels that may be present for a building principal. The section included five questions on a five-point Likert scale: “Strongly agree” (5), “Agree” (4), “Disagree” (3), “Strongly disagree” (2), and “No opinion” (1). The total score for the section could range between 5 and 25, with a higher score indicating a greater use of disaggregated data by the principal. The following table shows the mean, mode, standard deviation, and variance statistics for this section.

Table 15
The Degree to which Principals Disaggregate Student Data to Identify Goals

Section	Mean	Mode	Standard Deviation	Variance
2	20	20	2.54	6.42

The principals’ mean score for each item was 4, which corresponded to “Agree” on the items related to disaggregation. Ninety-six percent (96%) of all respondents indicated that they felt capable of working with disaggregated data. The response choices indicated whether the data “was compiled, disaggregated, and organized entirely by me,” “with assistance from others in the building,” “with assistance from others in the district,” “by someone in the district other than myself,” or “was not compiled, disaggregated, and organized for my building.” The largest group of respondents, 46% of the principals indicated that the data had been compiled, disaggregated, and organized with assistance from others in the district. When asked about support from the Educational Service Unit, 45% of the respondents indicated that the ESU had provided consultation about their disaggregated data, 42% of the respondents indicated that the ESU had provided training about disaggregation, and 35% of the respondents indicated that the ESU had provided no support for data disaggregation. The principals were also asked about the assistance and/or support they received related to the disaggregation of data.

Research Question 1c: To what degree were teachers, parents, and community members engaged in data-driven decision-making?

The elementary principals were asked the extent to which teachers, parents, and community members were engaged in data-driven decision-making. The section included five questions on a five-point Likert scale: “Strongly agree” (5), “Agree” (4), “Disagree” (3), “Strongly disagree” (2), and “No opinion” (1). The first two questions featured a reverse scale, as strong agreement to the statements indicated that the principal did not involve others in the data analysis and decision-making processes. The total score for the section could range between 5 and 25, with a higher score indicating a greater level of involvement of constituent groups. The following table indicates the mean, mode, standard deviation, and variance for this section.

Table 16

Degree to which Teachers, Parents, Community Members are Engaged in Data-driven Decision Making

Section	Mean	Mode	Standard Deviation	Variance
3	18.68	19	2.57	6.59

The principals’ mean score on the items in this section was 3.72, which indicates a score nearest “Agree.” When asked about the level of support for involving others, 42% of the respondents selected that the process “was facilitated with assistance from others in the building,” and another 42% of the respondents indicated that the process “was facilitated with assistance from others in the district.” Only 3% of the respondents indicated that the process was facilitated entirely by them. The principals were also asked about the support provided by the Educational Service Unit and could respond to multiple options related to possible supports. The respondents indicated that 45% had been offered training about writing goals, and 41% had been provided consultation with ESU staff about their school’s goals. Twenty-five percent (25%) had not received any support for the writing of goals.

Research Question 1d: To what degree was data used to determine professional development?

The NSDC Standards also indicated that the use of data to plan and implement professional development programs was an important factor of data-driven decision-making for principals. Therefore, this section included eight questions on a five-point Likert scale: “Strongly agree” (5), “Agree” (4), “Disagree” (3), “Strongly disagree” (2), and “No opinion” (1). The scores in this section could range from 8 to 40, with a higher score indicating a greater use of data to determine professional development. The following table indicates the mean, mode, standard deviation, and variance for this section.

Table 17

Degree Data was Used to Determine Professional Development

Section	Mean	Mode	Standard Deviation	Variance
4	28	32	3.65	13.31

The principals mean score on the items in this section was 3.5, which is between “Disagree” and “Agree.” When asked about support for using data to identify professional development activities for his/her building, 50% of the respondents indicated that the professional development was planned with assistance from others in the district and 41% indicated that it was planned with assistance from others in the building. When asked about the support provided by the Educational Service Unit related to designing professional development, 56% of the respondents indicated that the ESU held trainings at the ESU for all schools in the area and 45% indicated that the ESU had provided consultation about his/her professional development. When asked if they felt capable of planning professional development, 96% indicated that they did feel capable.

Research Question 1e: To what degree was data used to monitor the school improvement goals?

The NSDC also stated that the on-going use of data to monitor the school improvement goals was an important behavior for building principals to exhibit. As a result, the questions in this section related to the frequency with which data was collected and reported. This section

included seven questions organized around a five-point Likert scale: “Frequently (Once or more per month)” (5), “Occasionally (Once or more per quarter)” (4), “Sometimes (Once or more per semester)” (3), “Infrequently (Once or more per year)” (2), or “Never (Not at all)” (1). The scores for this section could range from seven to 35. The following table indicates the mean, mode, standard deviation, and variance for this section.

Table 18
Degree Data Was Used to Monitor School Improvement Goals

Section	Mean	Mode	Standard Deviation	Variance
5	23.5	24	4.32	18.66

The principals mean score on the items in this section was 3.4, which is between “Sometimes” and “Occasionally.” When asked about support for collecting and reporting student achievement data for his/her building, 46% of the respondents indicated that they received assistance from others in the district and 44% of the respondents received assistance from others in the building. Only 4% of the respondents indicated that the process was completed entirely on their own and 6% indicated that the process was completed by someone other than them. When presented with possible supports provided by the Educational Service Unit related to collecting and reporting student achievement data, 52% indicated that the ESU had provided consultation and 49% had been provided with training related to collecting and reporting data. Only 29% of the respondents indicated that the ESU had not provided any support related to collecting and reporting results. When asked if they felt capable of collecting and reporting student achievement results, 92% of the respondents indicated that they did feel capable.

DISCUSSION AND IMPLICATIONS

The data generated by this survey provided insight into the use of data in Nebraska’s elementary schools. The Nebraska STARS accountability system relies heavily on the leadership provided by building principals and encourages instructional leaders to use the data generated by the system to improve the school. Overall, it appears that the majority of schools are demonstrating proficiency according to the National Staff Development Council’s standards for data-driven decision-making.

The elementary principals indicated that they were involving their school faculty in the analysis of data. This particular section generated the highest mean score across all of the items in the sections and indicated that many principals were engaging their staff members in the process of analyzing data.

The elementary principals also indicated a strong level of agreement that they were using disaggregated data in their analysis process. The questions in this section identified the types of data that were used to identify school improvement goals ranging from criterion and norm-referenced test results to discipline referrals and grade retention information and also the disaggregation by race, gender, free and reduced lunch status, and special needs. The results from this section indicated that the principals were using a variety of data and were disaggregating that data in some way.

The next section of the survey measured the engagement of teachers, parents, and the community in data-driven decision-making. While the mean for this section was lower than other sections of the survey, when looking at the individual questions along with the first section of the survey, it

appeared that teachers are being involved in the process of analyzing data. However, a majority of the respondents indicated that they were not involving parents and/or community members in the process of analyzing data.

The fourth section of the survey measured the elementary principals' use of data to plan for professional development. The mean of this section was lower than the previous sections, which seemed to indicate that principals had not used data to plan or evaluate professional development. However, one of the initial questions about the alignment of the school's improvement goal and staff development plan indicated that for the vast majority of respondents, 97%, alignment was present. On other questions related to monitoring staff development and involving the whole faculty, there were much lower levels of agreement.

The final section of the survey measured the elementary principals' use of data to monitor the school improvement process. The questions focused on the frequency of these types of activities. This section yielded the lowest mean of the survey, with a score closest to "Sometimes." While the mean of this section was lower, no more than 4% of respondents selected the choice of "Never" on any of the items.

The data generated from this survey is useful to both the Nebraska Department of Education (NDE) and to individual school districts as they move forward in the implementation of the STARS system. It would be beneficial to conduct the survey with secondary-level principals and superintendents as well to determine the levels of data use in those positions.

For professional development considerations across the state, it would be beneficial for the NDE, Educational Service Units (ESU), and individual school districts to consider both the results related to the data use skills of building principals and the results related to the levels of support that the principals perceived, both within their schools and districts and from the ESU. While many principals indicated strong levels of support, there were still many who felt they did not receive the support that they needed to implement the NSDC standards. The NDE and other professional organizations may want to consider offering data training for principals. Such training could emphasize the use of data generated from STARS and other types of data that may be present within the school setting in order for principals to gain additional knowledge and skills related to data-driven decision-making.

Finally, the results of this survey seemed to indicate that the principals are using data to a significant extent. It would be interesting to see if principals in other states that have accountability systems that do not demand as much participation on the part of the building principal would yield the same types of results. The NDE, the ESUs, and individual school districts should be commended for their work in not only generating data for accountability but for using data effectively, according to national standards, to improve their schools.

REFERENCES

- Creswell, J.W. (2005). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Columbus, OH: Pearson-Merrill-Prentice Hall.
- Creswell, J.W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage.

- Dillman, D.A. (2000). *Mail and internet surveys: the tailored design method*. New York: Wiley & Sons, Inc.
- National Staff Development Council. (2001). *National Staff Development Council's Standards for Staff Development (Revised)*. Oxford, OH: National Staff Development Council.
- Roschewski, P., Gallagher, C. and Isernhagen, J. (2001). Nebraskans reach for the STARS. *Phi Delta Kappan* 82:8, 611-615.
- Roschewski, P. (2003). Nebraska STARS line up. *Phi Delta Kappan* 84:7, 517-520.
- Roschewski, P., Isernhagen, J., and Dappen, L. (2006). Nebraska STARS: achieving results. *Phi Delta Kappan* 87:6, 433-437.
- Roy, P. & Hord, S. (2003). *Moving NSDC's staff development standards into practice: Innovation configurations*. Oxford, OH: National Staff Development Council.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

**Study VII: The Role of Technology in Supporting
STARS—2006-07**



**Role of Technology in Supporting the Nebraska School-based Teacher-led
Assessment and Reporting System**

Allen L. Steckelberg Associate Professor
Lan Li, Research Assistant
University of Nebraska-Lincoln

INTRODUCTION

Nebraska has implemented an accountability system that emphasizes both local school involvement and attempts to provide meaningful knowledge for school improvement and instructional decisions. *Charting Stars: Voices from the Field Year-Four Report*, an evaluation of the Nebraska School-based Teacher-led Assessment and Reporting System (STARS) summarized numerous challenges and obstacles schools faced in successfully implementing the STARS. It was clear from responses across teachers, principals, superintendents, and ESU personnel that the time the process took, the difficulties in the collection, reporting and management of data, and assessment literacy were key issues impacting the implementation of the STARS.

Technology-based systems are often used to address issues of time, accuracy in handling large amounts of data, and clarity in reporting and sharing information. Technology can also be designed to support participants in learning and carrying out the assessment process. The role

technology plays in the STARS process is evolving as schools develop assessment practices and acquire technology. Previous STARS evaluations suggested that further information on the role technology plays and the issues it both addresses and creates would be helpful as schools and the Nebraska Department of Education (NDE) develop policy and support resources.

This report represents the results of the second administration of a survey of school assessment coordinators in Nebraska. The survey was designed to assess the extent of technology use by schools as they participated in the STARS process; identify critical issues in using technology-supported planning, assessment and reporting systems; and provide direction for the development of policies and resources that will support more efficient and effective implementation of the STARS. The results of the second survey are highly consistent with the results of the 2006 administration of the survey.

METHOD

Survey Development

Survey items were developed based on a review of open-ended responses from the *Charting Stars: Voices from the Field Year-Four Report* and discussions of the Nebraska Assessment Reporting Management Committee, Nebraska Department of Education, and the STARS Comprehensive Evaluation Project. Minor changes were made for the 2007 version of the survey including an additional question on the impact of the “Assessment and Reporting Management System”(ARMS) grants. The survey was constructed in four sections.

The first section of the survey gathered demographic information from participants. This information provided a picture of the characteristics of respondents including: school size, job title, roles in the assessment process, relationship with ESUs, and the district investment in technology. The second section of the survey gathered information on the current uses of technology including: software used by schools in various aspects of the assessment process, the match between technology-based systems and the Nebraska Assessment Model, and the quality of technology used by the district in the assessment process. The third section addressed respondents’ perceptions of the impact of technology on the assessment process including: contributions of technology to various aspects of the assessment process, impact on time and complexity of the process, impact on the value of data collected and understanding and use of that data. The final section of the survey addressed perceptions of future assessment systems and issues that might guide future policy decisions. In this section, the survey gathered data on potential barriers to using technology, quality of current available technology, desirable characteristics in technology-based assessment and reporting systems, and the perceptions of the likely impact of potential areas for state support or resources.

The survey was administered via a Web site. Email solicitations containing coded links allowed participants to access and complete the survey. Access to the survey was recorded to allow for follow-up requests to participants who did not respond to the initial solicitation. Survey responses were recorded separately to maintain anonymity. A survey sample can be found in Appendix K.

Survey Sampling

The survey targeted school personnel who coordinated the STARS process in local schools and ESUs in Nebraska. From a list of STARS school contacts provided by the Nebraska Department

of Education (NDE) and the STARS Comprehensive Evaluation Project, 463 names and email addresses were identified. An email solicitation to participate in the survey was sent to each name on the list. A follow-up email was sent two weeks later to those who had not yet responded. A total of 227 usable survey responses were received and used in the analysis.

Analysis

Data was exported from the database and imported into Statistical Package for the Social Sciences (SPSS) for analysis. Descriptive statistics were obtained as either frequency distributions or means depending on the nature of the data. Additional analysis comparing the differences between responses for schools smaller than 600 students and those in schools with 600 or more students used t-tests for equality of means to check for disparities based on school size. Levene’s test for equality of variances was used to confirm variance assumptions and the t-test was adjusted accordingly. Where significant differences existed for items, it is noted in the results.

RESULTS

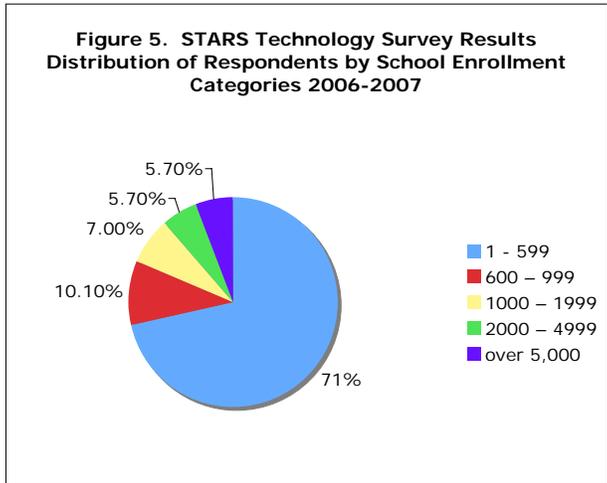
Respondent Characteristics

School Size

Respondents were asked to select one of five ranges of student enrollment that best represented their district size. Table 19 shows the distribution of responses. Results are shown in chart form in Figure 5. The distribution of school size closely reflected the distribution of school sizes in the original listing of contacts received from the Nebraska Department of Education (NDE) and the STARS Comprehensive Evaluation Project.

Table 19
Frequency and Percentage of Respondents’ School Enrollment (n=227)

	Frequency	Percent
1-599	161	70.9
600-999	23	10.1
1000-1999	16	7.0
2000-4999	13	5.7
Over 5000	13	5.7



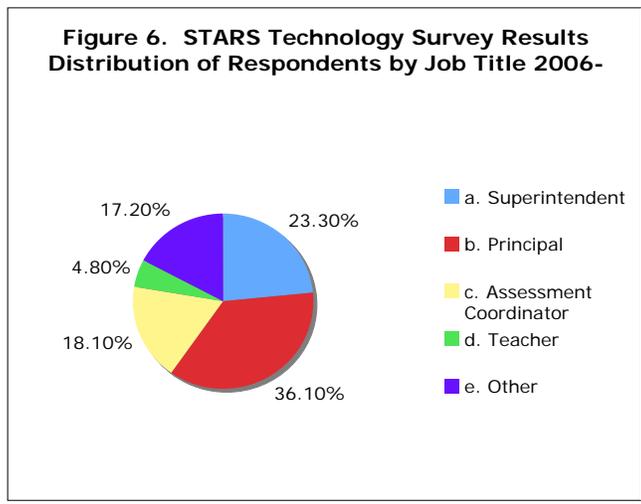
Job Title

Table 20 and Figure 6 show the distribution of the respondent job titles in five categories as superintendent, principal, assessment coordinator, teacher, or other. The other category often included variations or combinations of these titles, with the most frequent other title being a variation of curriculum coordinator/director.

Table 20

Frequency and Percentage of Respondents' Job Titles (n=227)

	Frequency	Percent
Superintendent	53	23.3
Principal	82	36.1
Assessment Coordinator	41	18.1
Teacher	11	4.8
Other	39	17.2



Responsibilities

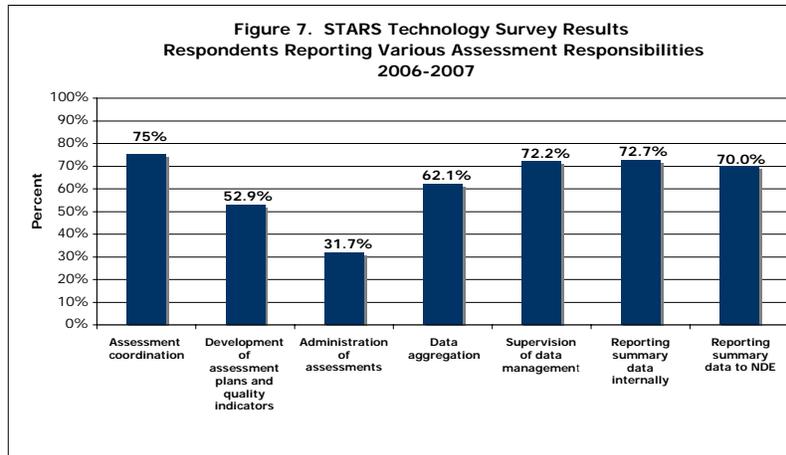
Respondents were asked to identify their responsibilities in the assessment process on the survey. They were allowed to select “all that applied” from a list of responsibilities including assessment coordination, plan development, assessment administration, data aggregation, supervision of data

management, internal reporting, and/or reporting to NDE. Table 21 provides the frequency and percent of respondents' assessment responsibilities. Figure 7 provides a comparison of percentages for each responsibility.

Table 21

Frequency and Percentage of Respondents' Assessment Responsibilities (n=227)

	Frequency	Percent
Assessment Coordinator	171	75.3
Development of Assessment Plans and Quality Indicators	120	52.9
Administration of Assessments	72	31.7
Data Aggregation	141	62.1
Supervision of Data Management	164	72.2
Reporting Summary Data Internally	165	72.7
Reporting Summary Data to NDE	159	70.0



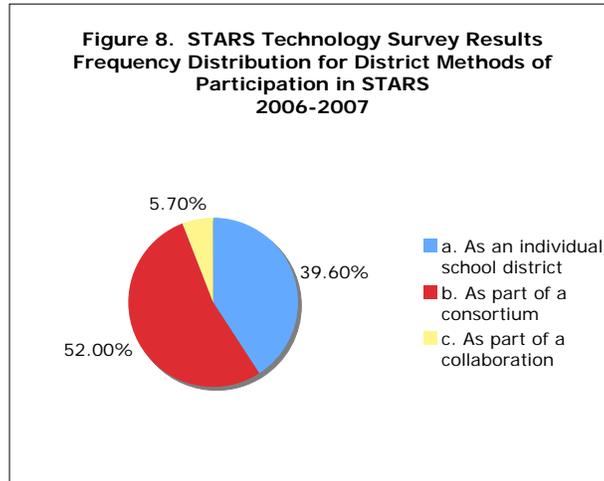
Participation

Respondents were asked to identify whether their school district participated in the STARS process as an individual school district, as part of a consortium, or as part of a collaboration. Table 22 and Figure 8 show the frequency distribution of responses.

Table 22

Frequency and Percentage of Respondents' Type of Participation in STARS (n=227)

	Frequency	Percent
As an Individual School District	90	39.6
As Part of a Consortium	118	52.0
As Part of a Collaboration	13	5.7



ESU Roles

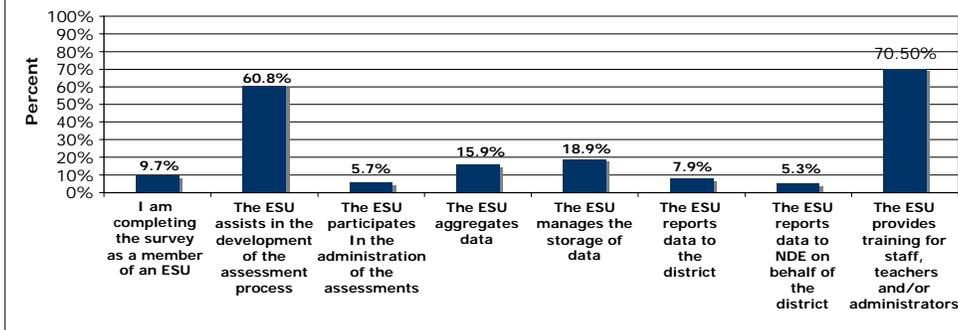
Respondents were asked to first identify whether the ESU played a role in the assessment process in their district. Eighty percent (80%) of respondents said that the ESU did play a role. Respondents who indicated that the ESU did play a role were asked to identify which of eight roles the ESU played. They were asked to indicate all the roles that applied. The frequency and percent of their responses are shown in Table 23. Figure 9 shows a profile of the percentage of respondents that indicated each of the possible ESU roles. It is clear that for schools that do use the ESU in the STARS, development of the assessment process and training were primary roles. The ESUs do play other roles, but do so less frequently.

Table 23

Frequency and Percent of ESU Roles in the Assessment Process (n=227)

	Frequency	Percent
I am completing the survey as a member of an ESU	22	9.7
The ESU assists in the development of the assessment process	138	60.8
The ESU participates in the administration of assessments	13	5.7
The ESU aggregates data	36	15.9
The ESU manages the storage of data	43	18.9
The ESU reports data to the district	18	7.9
The ESU reports data to the NDE on behalf of the district	12	5.3
The ESU provides training for staff, teachers, administrators	160	70.5

**Figure 9. STARS Technology Survey Results
Respondents' Selecting Various ESU Roles in the Assessment Process
2006-2007**



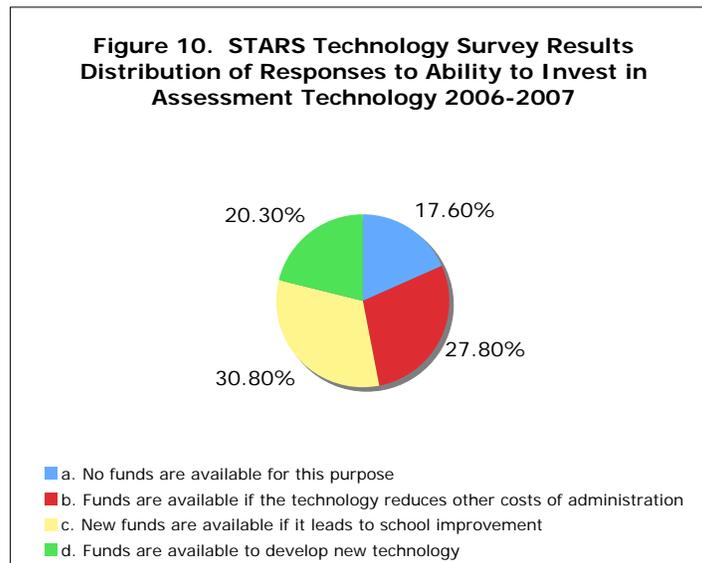
Ability to Invest in Assessment Technology

Respondents were asked to choose which of four statements best described their districts' ability to invest in technology specifically to support the assessment process. The distribution of responses is shown in Table 24 and Figure 10.

Table 24
Ability to Invest in Assessment Technology (n=227)

	Frequency	Percent
No funds are available for this purpose	40	17.6
Funds are available if the technology reduces other costs of administration	63	27.8
New funds are available if it leads to school improvement	70	30.8
Funds are available to develop new technology	46	20.3

**Figure 10. STARS Technology Survey Results
Distribution of Responses to Ability to Invest in
Assessment Technology 2006-2007**



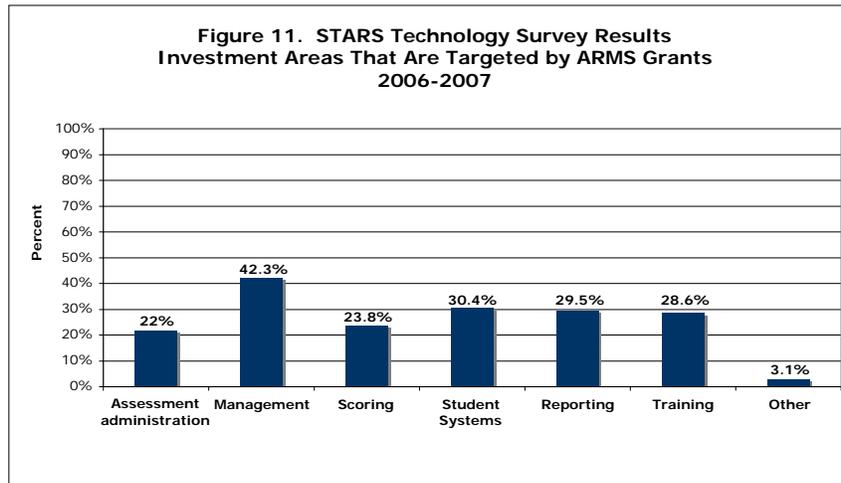
Impact of the Assessment and Reporting Management System (ARMS) Grants

Sixty-seven percent (67%) of the respondents agreed that ARMS Grants issued through the Nebraska Department of Education directly impact the district's ability to invest in assessment technology. Table 25 and Figure 11 provide a breakdown of the frequency and percent of potential investment areas that were targeted by the ARMS grants.

Table 25

Frequency and Percent of ARMS Grants Investment Areas (n=227)

	Frequency	Percent
Assessment Administration	50	22
Management	96	42.3
Scoring	54	23.8
Student Systems	69	30.4
Reporting	67	29.5
Training	65	28.6
Other	7	3.1



Current Use of Technology

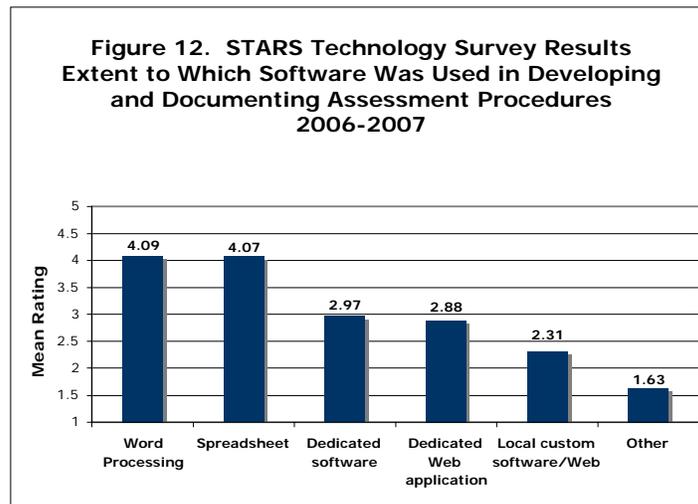
A series of six items were used to gather information on the current uses of technology in the STARS assessment process. Four questions asked participants to report the extent to which they used various types of software in four aspects of the STARS: developing and documenting assessment procedures; administering assessments to students; interpreting and reporting assessment results; and aggregation, management, and storage of assessment data. Tables 26-29 and Figures 12-15 show the results for each of the aspects. Respondents rated each software type on a scale of “1” representing “not at all” to “5” representing “to a great extent.” Results are reported as mean responses for each type of software.

The first item addressed use of software in developing and documenting assessment procedures. Results indicated that productivity software such as word processing and spreadsheet applications were used extensively in developing and documenting assessment procedures. Dedicated and local customized software were used to some extent, but less frequently than productivity software.

Follow-up analysis for differences between smaller and larger school districts showed significant differences in the extent of software used in developing and documenting assessment procedures. The mean scores were significantly higher for large schools for word processing (larger schools $M = 4.52$, smaller schools $M = 3.92$), $t(166.446) = -3.967$, $p < .0001$, and dedicated software (larger schools $M = 3.38$, smaller schools $M = 2.81$), $t(205) = -2.387$, $p = .018$.

Table 26
Software Use in Developing and Documenting Assessment Procedures

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Word Processing	212	4.09	1.20	4.25	3.93
Spreadsheet	212	4.07	1.19	4.23	3.91
Dedicated Software	208	2.97	1.58	3.18	2.75
Dedicated Web Application	202	2.88	1.68	3.11	2.65
Local Custom Software/Web	199	2.31	1.51	2.52	2.10
Other	38	1.63	1.48	2.10	1.16

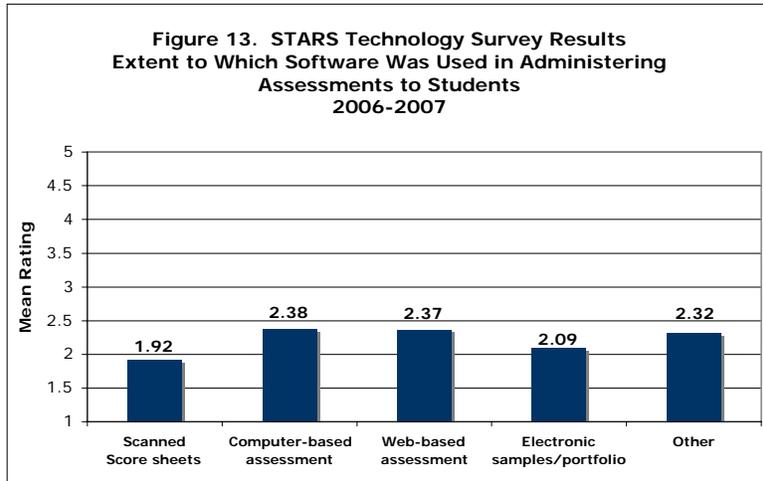


The second item addressed the use of software in administering assessments to students. Results from this item indicated that scanned score sheets, computer-based assessment, Web-based assessment, and electronic samples/portfolios were used infrequently in administering assessments to students.

Follow-up analysis for differences between smaller and larger school districts showed significant differences in the extent of software used in administering assessment. Larger schools indicated greater use of scanned score sheets (large $M = 2.75$, small $M = 1.56$), $t(90.287) = -5.788$, $p < .0001$. Smaller schools indicate greater use of computer-based assessment (large $M = 2.00$, small $M = 2.55$) $t(131.455) = 2.290$, $p = .024$, and Web-based assessment (large $M = 1.73$, small $M = 2.64$), $t(146.482) = 3.900$, $p < .0001$.

Table 27
Software Use in Administering Assessments to Students

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Scanned Score Sheets	204	1.92	1.46	2.12	1.72
Computer-based Assessment	208	2.38	1.71	2.62	2.15
Web-based Assessment	200	2.37	1.75	2.61	2.12
Electronic Samples/Portfolio	199	2.09	1.40	2.29	1.90
Other	53	2.32	1.85	2.82	1.82



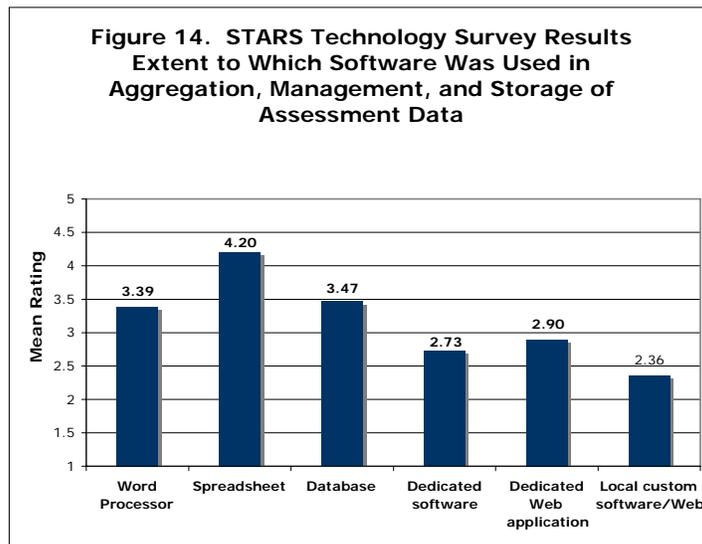
The third item addressed software use in the aggregation, management, and storage of assessment data. Spreadsheets, word processing, and databases were the most extensively used types of software in the aggregation, management, and storage of assessment data. Dedicated software, Web applications, and local custom software were used less extensively.

Follow-up analysis for differences between smaller and larger school districts showed significant differences in the extent of software used in aggregation, management, and storage of assessment data. Database software was used more frequently by large schools (M= 3.97) than small schools (M= 3.27), $t(123.820) = -2.961, p = .004$. The mean scores were significantly higher for large schools (M = 3.32) than smaller schools (M = 2.50) for dedicated software, $t(192) = -3.081, p < .0001$. Local custom software was also used more frequently by large schools (M=3.14) than small schools (M=2.02), $t(94.496) = -4.205, p < .0001$.

Table 28

Software Use in Aggregation, Management and Storage of Assessment Data

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Word Processor	205	3.39	1.47	3.59	3.18
Spreadsheet	208	4.20	1.23	4.36	4.03
Database	203	3.47	1.63	3.69	3.24
Dedicated Software	195	2.73	1.72	2.97	2.49
Dedicated Web Application	200	2.90	1.71	3.13	2.66
Local Custom Software/Web	186	2.36	1.65	2.60	2.12



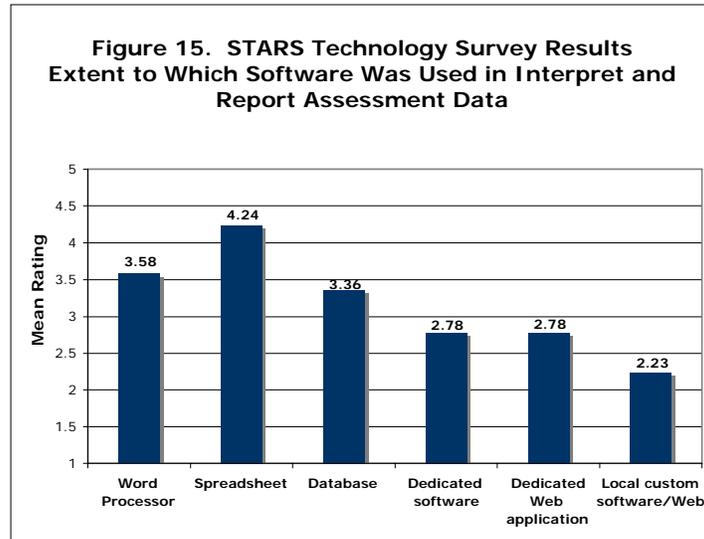
Item 4 in this section asked respondents to assess their use of the same set of software in the interpretation and reporting of assessment data. Mean responses and confidence intervals are reported in Table 29. This information is also depicted graphically in Figure 15.

Table 29

Software Use in Interpretation and Report of Assessment Data

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Word Processor	204	3.58	1.40	3.77	3.39
Spreadsheet	208	4.24	1.11	4.39	4.08
Database	206	3.36	1.67	3.58	3.14
Dedicated Software	196	2.78	1.66	3.01	2.54
Dedicated Web Application	197	2.78	1.68	3.02	2.55
Local Custom Software/Web	183	2.23	1.55	2.46	2.01

Significant differences were also found between large and small schools in this category. Large schools showed greater use of database (large M = 3.97; small M = 3.12), $t(123.687) = -3.664$, $p < .0001$, dedicated (large M = 3.41; small M = 2.51), $t(193) = -3.547$, $p < .0001$, and local custom software (large M = 2.88; small M = 1.98), $t(84.918) = -3.474$, $p = .001$, than smaller schools.

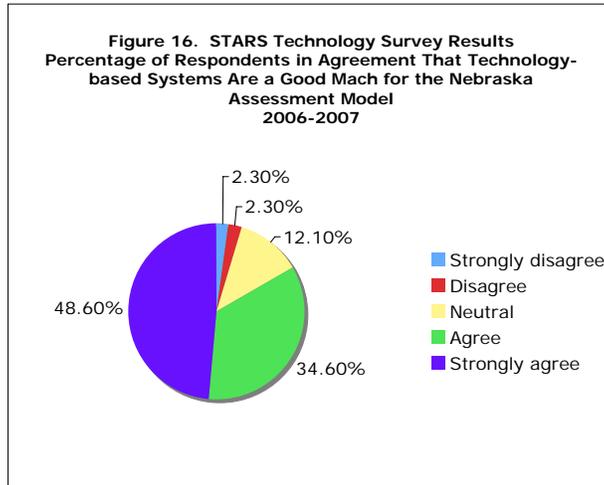


Survey participants were also asked to indicate their agreement with the following statement: “Available technology-based assessment systems are a good match for the Nebraska Assessment Model.” Eighty-three percent (83%) of the respondents indicated that they agreed or strongly agreed with the statement. Twelve percent (12%) were neutral and five percent (5%) either disagreed or strongly disagreed. The results are provided in Table 30 and depicted in Figure 16.

Table 30

Technology-based Assessment Systems are a Good Match for the Nebraska Assessment Model
(n=214)

	Frequency	Percentage
Strongly Disagree	5	2.3
Disagree	5	2.3
Neutral	26	12.1
Agree	74	34.6
Strongly Agree	104	48.6

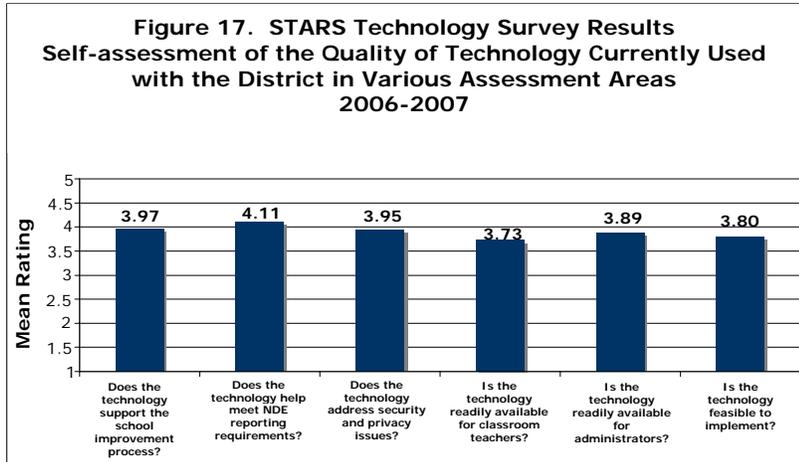


Participants were asked to rate the quality of the technology their district used in the STARS assessment process in seven areas including: supporting the school improvement process, meeting NDE reporting requirements, addressing security and privacy issues, being readily available for classroom teachers, being readily available for administrators, and being feasible to implement. Participants rated each area on a scale of “1” to “5” with “1” representing “low quality” and “5” representing “high quality.” Table 31 and Figure 17 show the mean responses for each of the areas.

Table 31

Self-assessment of the Quality of Technology Currently Used with the District in Various Assessment Areas

	N	Mean	SD	<u>Confidence Interval</u>	
				Upper	Lower
Does the Technology Support the School Improvement Process	214	3.97	1.07	4.12	3.83
Does the Technology Help in Meeting NDE Reporting Requirements	213	4.11	1.02	4.25	3.98
Does the Technology Address Security and Privacy Issues	212	3.95	1.17	4.10	3.79
Is the Technology Readily Available For Classroom Teachers	212	3.73	1.20	3.89	3.57
Is the Technology Readily Available For Administrators	211	3.89	1.17	4.04	3.73
Is the Technology Feasible to Implement	213	3.80	1.11	3.95	3.65



Follow-up analysis for differences between large and small schools did show differences in two of these areas. Respondents from small schools produced a higher mean rating for both availability of technology in classrooms and for administrators. For classroom teacher availability, the small schools mean was 3.86 while the large schools mean was 3.41, $t(209) = 2.499$, $p = .013$. For technology availability for administrators, the small schools mean was 4.00 while the large schools mean was 3.62, $t(208) = 2.177$, $p = .031$.

Impact of Technology on Assessment

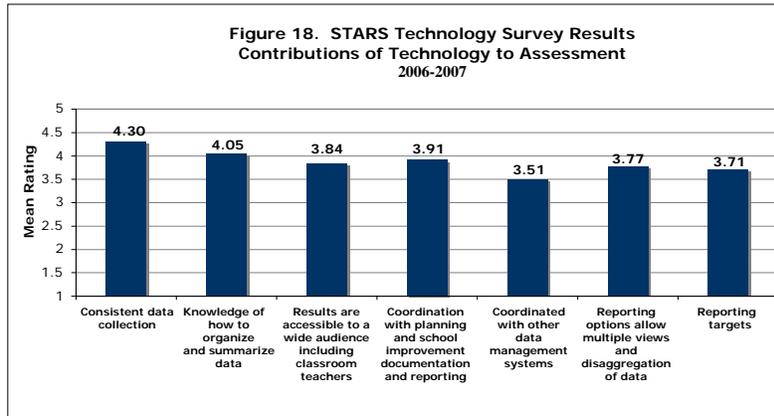
This section of the survey posed a number of questions targeted at gathering participants' perceptions of the impact that technology has on assessment. Items included the potential contributions of technology, the impact of technology on time involved in STARS, the impact on the complexity of the STARS process, the impact on the value of assessment data, whether technology assists teachers in understanding the data collection process, and whether technology assists teachers in utilizing data for instructional decision-making.

Respondents were presented with a series of potential contributions made by currently available technology and asked to rate the extent to which the technology contributed to the assessment process. Each potential contribution was rated on a five-point scale with "1" representing "not at all" and "5" representing "to a great extent." Table 32 presents the mean, standard deviation and confidence interval for each of the seven potential contributions. Ratings ranged from 3.23 to 4.01 indicating that participants saw technology moderately contributing to the assessment process. Consistent data collection had the highest rating. Figure 18 shows mean ratings for each of the potential contributions in a bar chart.

Small and large schools differed on rating for one potential technology contribution. Small schools ($M = 3.96$) rated *Results are Accessible to a Wide Audience Including Classroom Teachers* more highly than did large schools ($M = 3.57$), $t(209) = 2.334$, $p = .021$.

Table 32
Contributions of Technology

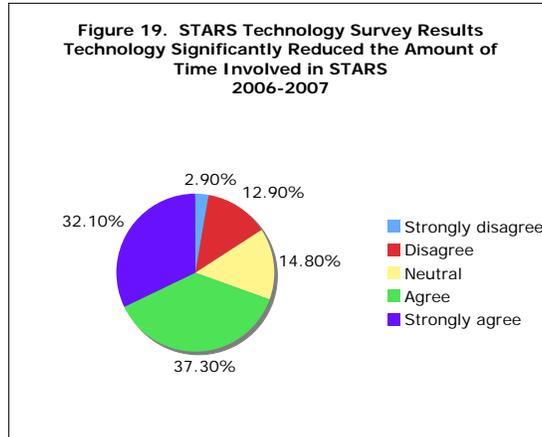
	N	Mean	SD	<u>Confidence Interval</u>	
				Upper	Lower
Consistent Data Collection	212	4.30	.87	4.42	4.18
Knowledge of How to Organize And Summarize Data	211	4.05	1.03	4.19	3.91
Results are Accessible to a Wide Audience Including Classroom Teachers	212	3.84	1.11	3.99	3.69
Coordination with Planning and School Improvement Documentation and Reporting	210	3.91	1.03	4.05	3.77
Coordinated with Other Data Management Systems	211	3.51	1.21	3.67	3.34
Reporting Options Allow Multiple Views and Disaggregating of Data	211	3.77	1.19	3.93	3.61
Reporting Targets	208	3.71	1.11	3.86	3.56



When asked to rate their agreement with the statement “The use of technology significantly reduces the amount of time involved in STARS” on a five-point scale from “strongly disagree” (1) to “strongly agree” (5), a majority of respondents either agreed or strongly agreed. Frequency and percent of respondents selecting each category are reported in Table 33 and depicted graphically in Figure 19. This response would seem to indicate that the respondents saw technology as one way to address the time demands noted in previous STARS evaluation data.

Table 33
Technology Significantly Reduces the Amount of Time Involved in STARS (n=289)

	Frequency	Percentage
Strongly Disagree	6	2.9
Disagree	27	12.9
Neutral	31	14.8
Agree	78	37.3
Strongly Agree	67	32.1

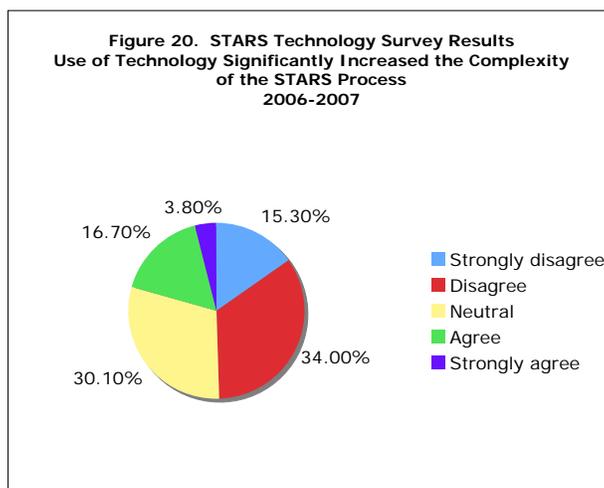


One potential side effects of employing technology is the addition of complexity to the assessment process. To gather survey participants' perceptions of the impact of technology on the complexity of STARS they were asked to indicate their agreement with the statement: "The use of technology significantly increases the complexity of the STARS process." Participants responded on a scale of one to five (1 = strongly agree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree). Results indicated that only about 21% of the respondents agreed or strongly agreed with the statement. Frequencies and percentages are listed in Table 34 and shown in Figure 20.

Table 34

Use of Technology Significantly Increased the Complexity of the STARS Process (n=287)

	Frequency	Percentage
Strongly Disagree	32	15.3
Disagree	71	34.0
Neutral	63	30.1
Agree	35	16.7
Strongly Agree	8	3.8



Perceptions of the impact of technology on the value of assessment data were assessed by asking

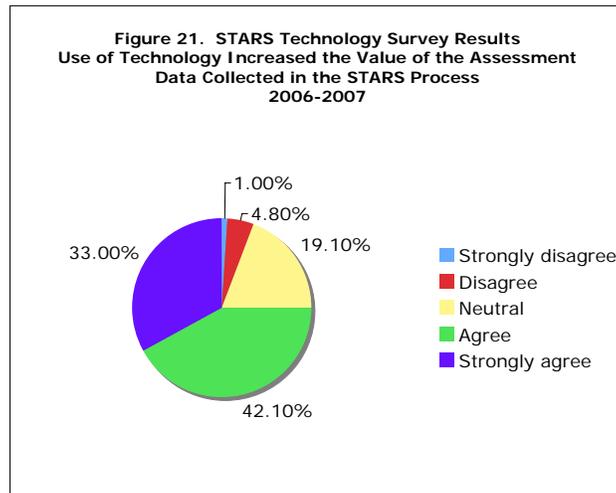
participants to rate their agreement with the following statement: “The use of technology increases the value of the assessment data collected in the STARS process.”

This item used a similar five-point scale (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree). A clear majority (75%) indicated that they agreed or strongly agreed with the statement. Results are show in Table 35 and Figure 21.

Table 35

Use of Technology Increases the Value of the Assessment Data Collected in the STARS Process (n=288)

	Frequency	Percentage
Strongly Disagree	2	1.0
Disagree	10	4.8
Neutral	40	19.1
Agree	88	42.1
Strongly Agree	69	33.0

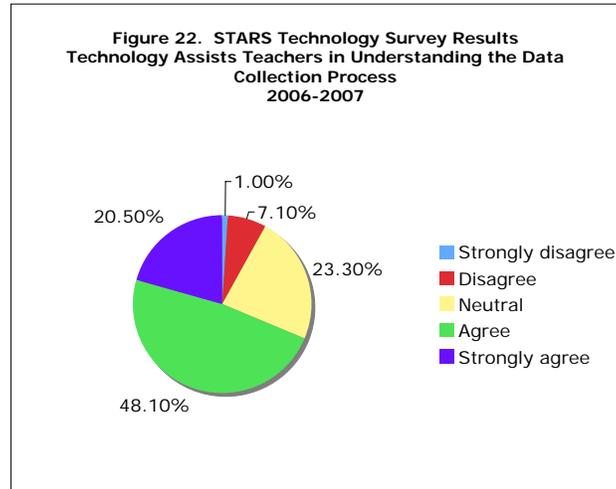


Perceptions of the impact of technology on teachers’ understanding of the data collection process were assessed by asking participants to rate their agreement with the following statement: “The use of technology assists teachers in understanding the data collection process.” This item also used a five-point scale (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree). A majority (68.6%) indicated that they agreed or strongly agreed with the statement. Frequencies and percentages are provided in Table 36 and depicted in Figure 22.

Table 36

Technology Assists Teachers in Understanding the Data Collection Process (n=289)

	Frequency	Percentage
Strongly Disagree	2	1.0
Disagree	15	7.1
Neutral	49	23.3
Agree	101	48.1
Strongly Agree	43	20.5

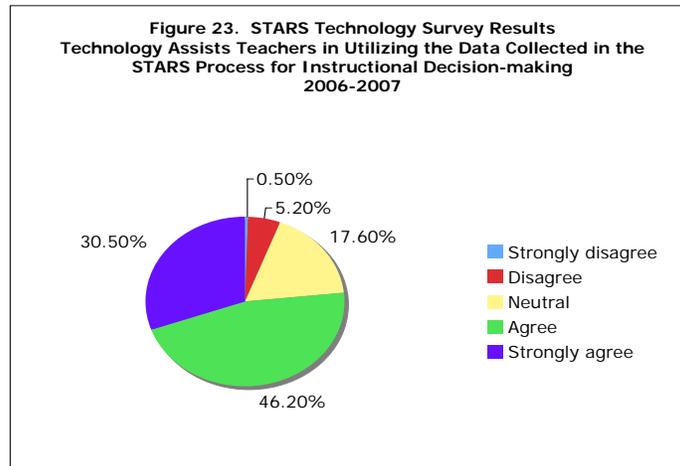


Nebraska’s approach to assessment is designed to provide opportunities to develop and implement assessment that will support instructional decision making on the part of teachers. Participants were asked to rate their agreement with the following statement on technology’s impact in this area: “Technology assists teachers in using the data collected in the STARS process for instructional decision making.” A scale of one to five (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree) was again used for this item. A clear majority (77%) agreed or strongly agreed with this statement. Frequencies and percentages are shown in Table 37 and depicted in Figure 23.

Table 37

Technology Assists Teachers in Utilizing the Data Collected in the STARS Process for Instructional Decision-Making (n=288)

	Frequency	Percentage
Strongly Disagree	1	0.5
Disagree	11	5.2
Neutral	37	17.6
Agree	97	46.2
Strongly Agree	64	30.5



Items in this section of the survey examined perceptions of the impact of technology on the assessment process. Results indicated that respondents were positive about the potential of technology to contribute to the process. A majority of respondents indicated that technology reduced the amount of time involved, increased the value of assessment data, assisted teachers in understanding and using assessment data for instructional decision making and, at the same time, did not increase the complexity of the assessment process. Respondents on average also indicated that they saw a number of potential contributions of technology as having a moderate to extensive impact. Technology’s contribution to consistent data collection was rated as having the greatest impact among the list.

Future Systems and Issues

The final section of the survey was designed to provide information helpful in planning, allocating resources, and developing policies related to implementing technology in the STARS assessment process. Questions addressed the potential barriers in using technology, perceptions of the quality of currently available technology, and the desired features and characteristics of technology-based data collection, management and reporting systems. Participants were also asked to rate the likely impact of a number of potential areas of State support or resources.

In order to assess perceptions of potential barriers to using technology, participants were presented with a list of ten possibilities. Each possibility was rated on a scale of “one” to “five” with “one” representing “low” and “five” representing “high.” Mean ratings for “Initial Cost” (4.09) and “Staff Time to Learn System” (4.05) were rated as the most significant barriers. Maintenance costs (3.9), teacher training (3.84), technical expertise to maintain (3.80), and administrator time to learn (3.77) were rated nearly as high. Mean responses for all items in the list were at a level of moderate or above. Means, standard deviations, and confidence intervals are presented in Table 38. Figure 24 presents a comparison of mean scores for each potential barrier.

It should be noted that the analysis for differences based on school size showed that the mean score for a number of these potential barriers differ significantly in small and larger schools. Mean ratings for the degree to which initial costs are a barrier was greater in large schools (M = 4.31) than small schools (M=3.99), $t(207) = -1.988$, $p = .048$. Maintenance costs were also rated as a greater barrier by large schools (M=4.13) than small schools (M=3.80) $t(159.256) = -2.380$,

p = .019. Interoperability with other systems rating as a barrier to using technology was higher for larger schools (M = 4.02) than for small schools (3.65), t (129.079) = -2.211, p = .029. Finally, administrator time to learn systems was also perceived as a greater barrier by large schools (M=4.07) than small schools (M=3.65), t (143.499) = -2.6545, p = .009.

Table 38
Potential Barriers Related to Using Technology to Support Assessment
(1 = Low, 5 = High)

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Initial Cost	210	4.09	1.05	4.23	3.95
Maintenance Costs	209	3.89	1.05	4.04	3.75
Commitment To Single Vendor	207	3.35	1.26	3.52	3.18
Knowledge to Choose Best System	207	3.71	1.05	3.85	3.56
Technical Expertise to Maintain	209	3.80	1.17	3.96	3.65
Fear of Obsolescence	209	3.40	1.19	3.56	3.24
Teacher Training	209	3.84	1.06	3.99	3.70
Interoperability with Other Systems	209	3.77	1.16	3.92	3.61
Staff Time to Learn System	208	4.05	1.03	4.19	3.91
Administrator Time to Learn System	207	3.77	1.56	3.93	3.61

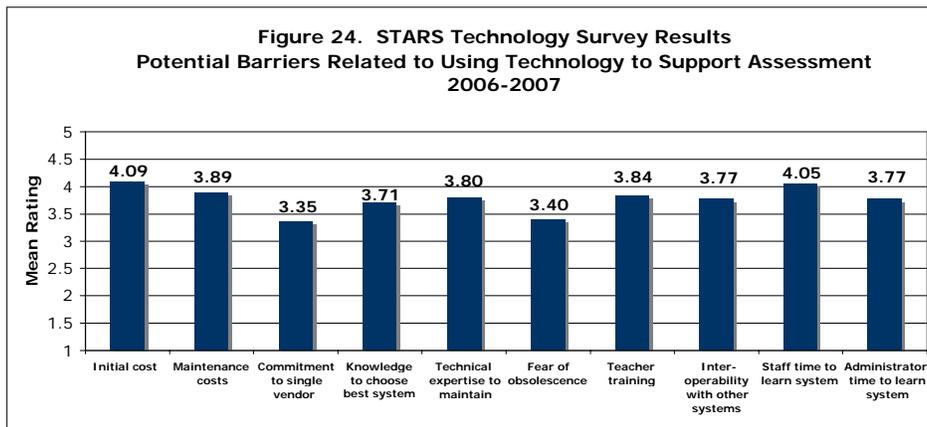


Table 39 and Figure 25 present results from the survey participants’ ratings of the quality of currently available technology used in collection, management and reporting of assessment data. Respondents rated each of five aspects of current technology on a scale of “one” to “five” with “one” being “low” and “five” being “high.” Mean scores for all five aspects were in a moderate range from 3.17 to 3.65.

Follow-up analysis for differences between smaller and larger school districts showed significant differences in the quality of currently available technology used to support the collection, management, and reporting of assessment data. Smaller schools (M = 3.63) reported higher ratings for system quality than did larger schools (M = 3.23), t (204) = 2.602, p = .018. The mean scores were also significantly higher for smaller schools (M = 3.62) than for larger schools (M = 2.95) for classroom level reporting features, t (204) = 3.957, p = .010. Mean time for smaller

schools ($M = 3.31$) were also higher than larger schools (2.80) for interface with State reporting, $t(203) = 2.829, p = .005$.

Table 39

Quality of Currently Available Technology Used to Support the Collection, Management, and Reporting of Assessment Data (1 = Low, 5 = High)

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Systems Are Comprehensive	206	3.51	1.09	3.66	3.36
System Quality	207	3.52	1.04	3.66	3.38
Classroom Level Reporting Features	207	3.43	1.15	3.59	3.27
District Level Reporting Features	207	3.65	1.06	3.80	3.51
Interface with State Reporting	206	3.17	1.19	3.33	3.00

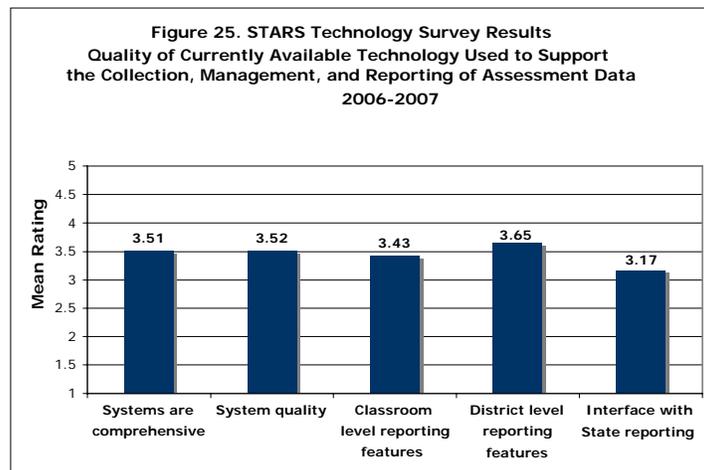


Table 40 presents the results of survey participants' ratings of the desirability of various characteristics of technology-based data collection, management, and reporting systems. Respondents rated each of six characteristics of technology-based systems on a "one" to "five" scale with "one" representing "low" and "five" representing "high." All six of the characteristics were rated at a high level with averages ranging from 4.07 to 4.44. Ease of data entry, usefulness of results for instructional decisions, and match with reporting requirements had the highest average ratings among the listed characteristics. Results would indicate that all of these characteristics should be considered when evaluating or recommending technology-based assessment systems. A relative comparison of mean responses is also shown in Figure 26.

One characteristic was perceived differently by small and large schools. Comprehensive features was see as more desirable by large schools ($M = 4.70$) than small schools ($M = 4.32$), although both ratings were relatively high, $t(143.261) = -3.347, p = .001$.

Table 40

Desirability of Characteristics in a Technology-based Data Collection, Management, and Reporting System (1 = Low, 5 = High)

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Ease of Data Entry	209	4.44	.84	4.56	4.33
Usefulness of Results for Instructional Decisions	209	4.44	.85	4.55	4.32
Comprehensive Features	208	4.12	.91	4.24	4.00
Match with Reporting Requirements	207	4.43	.86	4.55	4.31
Ease of Learning System	207	4.22	.97	4.35	4.09
Interoperability with Other Systems	207	4.07	1.08	4.22	3.92

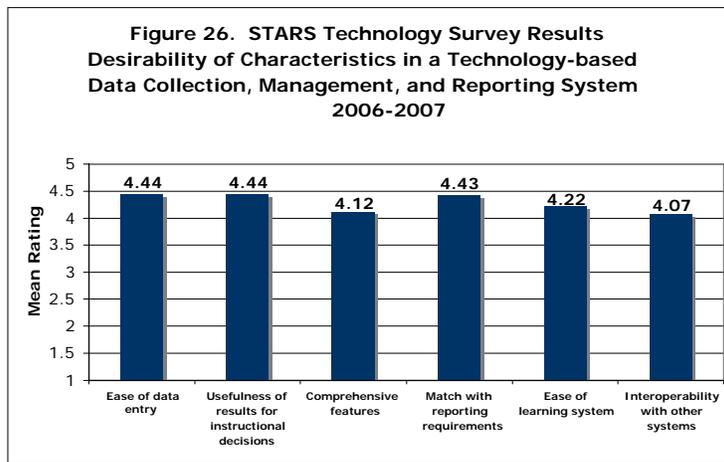


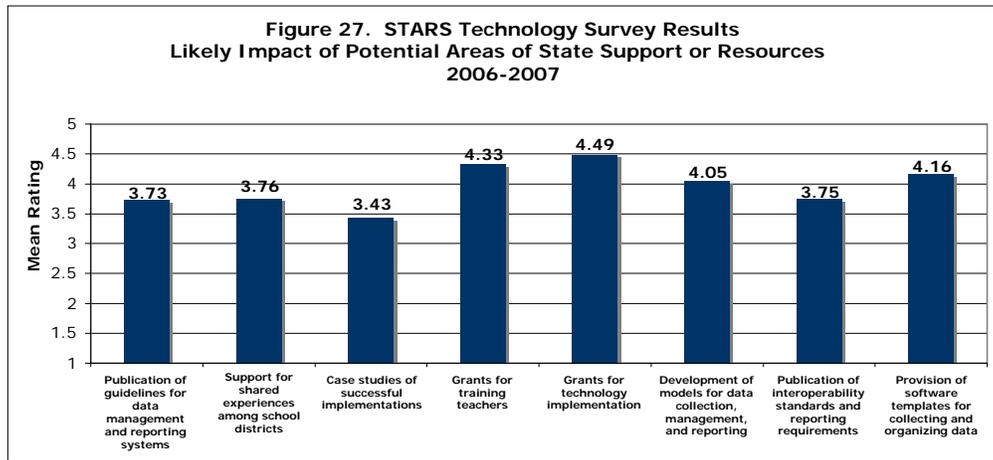
Table 41 presents data on survey participants’ ratings of the likely impact of potential areas of State support or resources. Respondents rated a series of eight potential areas of support or resources on a scale of “one” (low) to “five” (high). Grants for technology implementation (4.49) and for training teachers (4.33) received the highest ratings. The provision of models for data collection, management, and reporting (4.05), and software templates for collecting and organizing data (4.16) were the next most highly rated as likely areas of impact. Figure 27 shows the relative mean rankings of each of the potential areas of State support or resources.

One potential area of State support or resources was rated differently by small and large schools. Grants for implementation were rated more highly by large schools (M=4.66) than small schools (4.42), $t(128.765) = -2.236, p = .027$. Both groups rated grants for technology implementation as having the highest likely impact.

Table 41

Likely Impact of Potential Areas of State Support or Resources (1 = Low, 5 = High)

	N	Mean	SD	Confidence Interval	
				Upper	Lower
Publication of Guidelines for Data Management and Reporting Systems	208	3.73	1.00	3.86	3.59
Support for Shared Experiences Among School Districts	208	3.76	.97	3.89	3.63
Case Studies of Successful Implementations	207	3.43	1.09	3.58	3.28
Grants for Training Teachers	209	4.33	.80	4.43	4.22
Grants for Technology Implementation	209	4.49	.74	4.59	4.39
Development of Models for Data Collection, Management, And Reporting	208	4.05	.90	4.17	3.93
Publication of Interoperability Standards and Reporting Requirements	208	3.75	.95	3.88	3.62
Provision of Software Templates For Collecting & Organizing Data	206	4.16	.92	3.28	4.03



SUMMARY AND DISCUSSION

Knowledge of the characteristics of typical respondents was helpful in interpreting the results. Responses indicated demographics similar to those expected from the initial list of STARS contacts and coordinators in local districts and ESUs. Respondents were primarily from smaller schools (70%) but were often part of larger consortiums for the purpose of implementing STARS. Respondents were often superintendents, principals, or assessment coordinators, with teachers and curriculum coordinators also being represented. The majority indicated funds were available for technology, particularly if it reduced other administrative costs or contributed to school improvement. In a new item, introduced in the 2007 administration of the survey, sixty-seven percent of respondents agreed that the ARMS grants directly impacted their ability to invest in assessment technology. The most frequently reported investment of funds was for management purposes (42%).

The survey provided useful information on the current uses of technology. The majority of respondents indicated that basic productivity tools such as word processing and spreadsheets were the most extensively used tools in a number of aspects of the assessment process. Dedicated assessment software, Web-based software, and locally developed software or Web applications were used less extensively. When broken down by school size, the rated extent of use of dedicated and locally developed software was proportionally higher in larger schools than in smaller schools possibly because of differences in resources. Scanning, computer and Web-based systems, and electronic portfolios had low ratings for use.

Several survey items targeted general perceptions regarding current technology and the assessment process. A majority (87%) of respondents agreed or strongly agreed that the technology-based systems were a good match to the Nebraska assessment model. This is the only item in the survey that differed markedly from the 2006 survey results. In the 2006 survey, only 47% of the respondents agreed or strongly agreed with the statement. The 2007 result is substantially more positive than the earlier survey. Respondents also rated the rating of the quality of the technology used by their district in the assessment process at a moderately to moderately high level. These results would seem to indicate that there is a potential to improve and increase the use of technology in the assessment process in many Nebraska schools.

Data from the survey clearly showed that survey participants felt technology has the potential to positively impact the assessment process. A majority of respondents indicated that technology reduced the amount of time involved, increased the value of assessment data, assisted teachers in understanding and using assessment data for instructional decision making and, at the same time, did not increase the complexity of the assessment process. Respondents also indicated that they saw a number of potential contributions of technology, such as consistent data collection, as having a moderate to extensive impact. It appeared that among school personnel who work with the STARS assessment process technology was viewed as having a positive impact.

Survey data also provided information useful for guiding future policies, support, and resources. The most significant barriers were costs of technology, staff and administrator time to learn systems, and teacher training. Interoperability was also rated as an issue in larger school districts. Quality ratings of current systems regarding their comprehensiveness, quality, reporting features, and interface with state reporting were moderate. Interestingly, ratings of classroom-level reporting features and interface with state reporting, although still moderate, were higher in small schools. A number of features of technology-based assessment systems were rated as highly desirable including: usefulness of results for instructional decisions, ease of data entry, match with reporting requirements, and ease of learning systems. Survey participants felt that the areas of state support with the greatest potential for impact were grants for technology implementation and teacher training. These areas were closely followed by the provision of exemplars in the form of models for data collection, management and reporting, and software templates for collecting and organizing data.

Results from the survey indicated that STARS assessment leaders in Nebraska schools recognized the potential impact of technology on the assessment process, but at this point do not extensively make use of dedicated or locally developed software, although this may be occurring in larger districts or consortiums. Schools saw grant support for implementing technology and training school staff as likely to have significant impacts on the quality, usefulness, and value of the assessment process.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

**Study VIII: Statewide Writing Assessment Scoring:
Effective Professional Development
for the Classroom Teacher**



**Statewide Writing Assessment Scoring: Effective professional Development
for the Classroom Teacher**

Sue Anderson, Director of Professional Development
Educational Service Unit #3

INTRODUCTION

Statement of the Problem

“Statewide Writing Assessment Scoring: Effective Professional Development for the Classroom Teacher” is a study undertaken by the Nebraska Department of Education. This study was intended to ascertain the perceptions of scoring participants for the 2007 Statewide Writing Assessment regarding the relationship of their scoring experience to their instructional practices in the teaching of writing and to determine the effectiveness of the procedures used for the training of scoring participants for the assessment.

The benefits of this study are two-fold. First, the relationship that scoring participants, referred to in this analysis as raters, perceive between the scoring experience and their classroom practices is critical to understanding the effects, if any, of a state-level assessment on the improvement of student achievement in writing. Secondly, the responses provided by raters related to the training procedures are beneficial to the Nebraska Department of Education in the implementation processes of the Statewide Writing Assessment.

Purpose of the Study

The purposes of the project of study were to learn raters' perceptions of:

- 1) the effect of the scoring experience on their classroom practices, and
- 2) the effectiveness of the training they received during the scoring of the 2007 Nebraska Statewide Writing Assessment.

Research Questions

The purposes of the study provided the basis for the research questions to be answered. They are:

- 1) What impact does participation in the scoring of the Statewide Writing Assessment have on raters' classroom practices in the teaching of writing?
- 2) What are raters' perceptions of the effectiveness of the training they received during the scoring of the 2007 Nebraska Statewide Writing Assessment?

REVIEW OF RELATED LITERATURE

Since 1975, when Newsweek magazine published "Why Johnny Can't Read," the national attention has been focused on how to improve student academic achievement. The "Nation At Risk" report in 1983 declared that American educational standards were too low, and in 1989 a national-level education summit convened by President George H. W. Bush advocated state academic standards and standardized tests. The 1994 Educate America Act and the Goals 2000 program during the Clinton administration placed accountability at the center of the school reform movement (Gallagher, 2007). With the enactment of the federal No Child Left Behind Act in 2002, states are now fully engaged in measuring and reporting student achievement against academic standards through extensive standardized testing programs.

Persistent throughout this 30-year debate by national and state officials is the issue of whether accountability through high stakes testing improves student achievement. The merits, or lack thereof, of state level writing tests have not been omitted from the debate. Of state testing programs that include performance assessments, Linda Mabry (1999) reports that direct writing assessments, which ask students to produce a sample of writing, are the most prevalent and longest running.

While that may be true, much has been written about how such writing assessments may not be accurate or adequate measures of how well students can write. The primary reasons cited are standardization measures that some believe reduce student writing to formulaic texts that demonstrate little creativity or original thought. Standardization measures of large-scale writing assessments that come under the most frequent attack are requiring all students to write on the same topic, placing a time limit on the writing response, and using rubrics to score the writing.

Mabry reports that all states administering direct writing assessments use rubrics to score student performance. In her discussion of rubrics she maintains that they "promote reliability by standardizing scoring, but they also standardize writing" by limiting the writing to "pre-determined standards." She further argues that rubrics "overwhelm the curriculum," and the result is that teacher-driven writing curricula, which may be superior, are overridden, an example of what she calls "the de-professionalism of teachers effectively being prevented from making decisions about curriculum and pedagogy for their students" (1999, p. 35).

Opponents of state writing assessments also argue that they limit the local school curriculum to the kinds of writing that will be tested to the exclusion of a variety of writing genre and modes. Among Hillocks' (2002) conclusions in *The Testing Trap: How State Writing Tests Control Learning* is that when states determine categories of writing to assess, those categories receive priority of emphasis by schools and teachers. Such prioritization and concern over how well student performance will stand up to public scrutiny often limits the local writing curriculum to what will be tested. Hillocks echoes Mabry's (1999) contention that writing assessments, more often than not, foster "formulaic" writing that rarely reflects a thoughtful process in the expression of original ideas or in presenting evidence to support an idea. Wiggins (1998) warns that many state writing assessments may undermine good writing by scoring only for focus, organization, style and mechanics because such scoring does not usually ask judges to consider whether the writing is powerful or moving. Freedman and Daiute (2001) maintain that writing assessments do not always acknowledge issues of culture, process, or purpose, and when they carry high stakes, teachers may feel forced to spend instructional time on writing rules out of context, a practice that has not been associated with the support of writing development.

Hillocks (2002) examined the state writing assessments of Illinois, Texas, New York, Kentucky, and Oregon. In spite of his conclusion that state writing tests limit the local curriculum, Hillocks reports that teachers in each of the five states believe their state assessment supports a desirable writing program, approve of the state scoring rubric, support teaching to the test, and think the assessment improves student writing. However, most teachers surveyed believe their state writing test is not a comprehensive measure of student ability. Only teachers from Kentucky believe that the state test, which is a portfolio of writing that students create over time, is a valid measure of their writing skills.

Hillocks' (2002) findings, related to the scoring of the state writing tests he examined, showed a wide range of processes. Illinois and Texas contract with commercial testing companies for the scoring. New York relies on a training of trainer model, training 2-3 teachers from each district who are then responsible for training teachers and conducting the scoring of the test within their respective districts. Oregon recruits teachers to score the test at regional scoring sites throughout the state, and Kentucky's writing assessment, which is a portfolio of student writing generated over time, is scored at the local school district level by classroom teachers. In Hillocks' study teachers representing states that use their own teachers to score the state writing test, report higher levels of approval of the test and its relationship to their classroom practices.

Spandel and Stiggins (1997) looked at an important way to bridge the gap between large-scale assessment and classroom instruction in their discussion of large-scale assessments that enlist classroom teachers as raters. They maintain that teachers who assess students' work in the large-scale setting give themselves an education both in how to write and how to assess. They further argue that teachers who participate as raters in large-scale assessment gain a broader perspective of student writing from throughout a state or district, which is very different from seeing only the writing of their students. But they cited that the major contribution large-scale writing assessment has made to writing instruction is higher expectations about student performance. They also maintain that a state writing assessment, perhaps more than any other single stimulus, can prompt decision makers to ask important questions that they may not have asked before, questions that examine when, where, and how writing is being taught.

In direct opposition to the position on rubrics taken by Mabry (1999), Spandel (2005) argues that an equally important contribution large-scale writing assessment has made to instruction is the use of clearly defined rubrics that promote the reliability of scoring. Rubrics are generally defined as scoring tools that contain criteria and a scale for evaluating performances or products. Rubrics are frequently accompanied by examples of products or performances to illustrate the various score points on the scale. Arter and McTighe (2001) describe the best rubrics used to assess student work as those that capture what teachers look for when judging quality. They also maintain that rubrics used by teachers to score student work in large-scale assessments help them clarify instructional goals. During the process of scoring many papers, they say, teachers build their skill and confidence to identify the features of effective writing and return to their teaching more confident to judge their students' work and more confident in their ability to teach students to be better writers.

Nebraska's Statewide Writing Assessment was first implemented in 2001 as a result of state legislation (79-760) which mandated the assessment and reporting of student performance by all public schools in meeting state standards in reading, mathematics, and writing. Policy guidelines for the writing assessment issued by the State Board of Education in 2001 require that the assessment be a direct assessment of student writing requiring students in grades four, eight and eleven to respond to a single writing prompt or topic within a prescribed amount of time.

The scoring process of Nebraska's Statewide Writing Assessment requires each student's response to be read and scored by two raters who use a rubric to assign a single holistic score. The rubric provided by the Nebraska Department of Education identifies the scoring criteria as ideas and content, organization, voice or tone, word choice, sentence fluency, and conventions. Raters assign a score based on how well the writing meets these criteria overall. The rubric criteria are aligned to the qualities of writing identified in the Nebraska content standards for writing.

In "Charting STARS – Sustainability as Challenge and Opportunity" Gallagher (2003) reports the results of year two of a research study and comprehensive evaluation of Nebraska's School-based Teacher-led Assessment and Reporting System (STARS). Among the major findings of the analyses of a survey administered to teachers on their perceptions and classroom practices related to the state writing assessment were that 69% placed more emphasis on practice writing assessments, 73% placed more emphasis on sharing assessment criteria in class, and 73% placed more emphasis on explicit instruction in six trait writing.

In addition to these findings, Gallagher reports that 88% of teachers agreed or strongly agreed that the six traits scoring rubric used to score the state writing assessment is useful for instruction; 75% agreed or strongly agreed that the state writing assessment supports learning objectives they have for their students; 72% agreed or strongly agreed that the results of the state writing assessment are useful for teachers; and 65% agreed or strongly agreed that the six traits are the most important features of writing.

METHODOLOGY

Population and Sample

The population of this study consisted of 390 raters in the scoring of the 2007 Nebraska Statewide Writing Assessment, with 131 raters at Grade 4, 129 raters at Grade 8, and 130 raters at Grade 11. A scoring session was conducted for each grade level.

Raters were selected to participate in the scoring sessions based on their teaching experience, including grade level(s) taught and years of teaching, their previous writing assessment scoring experience, prior training in the Six Trait writing assessment model, and the geographic region of the state they represented. Approximately 50% of the raters for each scoring session represented the eastern region of the state, 30% of the raters were from the central region, and 20% were from the western region. Raters in the Grade 4 scoring session represented 46 school districts; while raters in Grades 8 and 11 represented 47 districts each.

Data Collection and Analysis

The data for this study was collected from rater responses to a survey entitled 2007 Statewide Writing Assessment Scoring Rater Evaluation (Appendix L) created by the Assistant Director of Statewide Assessment at the Nebraska Department of Education. The survey items were intended to collect information about the raters, their perceptions of the effectiveness of the training provided during the scoring sessions, and the relationship between the raters' scoring experience and their classroom practices in the teaching of writing.

The survey was designed to be completed by raters in no more than 10 minutes. They were asked, but not required, to complete the survey on the final scoring day of each scoring session just prior to the start of scoring. Surveys were placed at each scoring table, and raters were asked by the scoring room leader to complete the survey before the re-training procedures and scoring began. Table leaders assigned to scoring tables collected the surveys when raters had completed them.

All survey responses were coded and tabulated at the Nebraska Department of Education. Analysis included the tabulation of frequency of responses for all items and the calculation of mean scores for items asking raters to indicate their levels of agreement and disagreement. The survey contained 29 questions across the following categories: rater information, scoring training, the scoring experience and its relationship to classroom practices, and comments related to reasons why raters wanted to participate in the scoring sessions.

Rater information items pertained to gender, level of education, previous writing assessment scoring experience at the state level, previous scoring experience at the regional or local level, participation in six trait writing training prior to the scoring, years of classroom teaching experience, and current teaching status including grade level(s) and subject area(s).

Survey items related to the raters' scoring experience and their classroom practices asked them to indicate their levels of agreement and disagreement with how the scoring experience affirmed their current instructional practices, increased their understanding of the qualities of good writing, increased their confidence to provide meaningful feedback to students, and increased their confidence as teachers of writing. Raters indicated their levels of agreement and disagreement with items in this section of the survey on a scale of "strongly disagree" to "strongly agree."

A single open-ended item on the survey asked raters to explain their reason(s) for wanting to participate in the scoring session(s). A final item invited raters to provide additional comments on any aspect of the scoring experience they wished to share.

Survey items related to scoring training asked raters to respond to the effectiveness of the content of the training, strategies to ensure scoring accuracy, clarity of training materials, and the knowledge and skill of scoring leaders and table leaders. Raters indicated their levels of agreement and disagreement with items in this section of the survey on a scale of “strongly disagree” to “strongly agree.”

FINDINGS

Analyses of the survey responses provided general rater information, rater perceptions of the effectiveness of the training they received as part of the scoring process, and the effect of the scoring experience on their knowledge and classroom practices as teachers. Of 390 raters participating in the scoring across all grade levels, 374 raters responded to the survey, for a response rate of 96%.

General Rater Information

Among raters across all grade levels, 90% were females, with 93% females at Grade 4, 90% at Grade 8, and 86% at Grade 11. Among all raters, 4% had Bachelor degrees, 54% had Masters’ degrees, 4% had Education Specialist degrees, and 2% had Doctoral degrees. Across all grade levels raters had an average of 18 years of teaching experience. At Grade 4 the average number of years of teaching experience was 17, at Grade 8 it was 20 years, and at Grade 11 it was 18 years.

The majority of raters across all grade levels, 64%, were currently teaching, with 73% of those at Grade 4, 63% at Grade 8, and 57% at Grade 11. Among raters currently not teaching, 6% were school administrators, 36% retired teachers, 33% substitute teachers, 1% student teachers, and 24% with education-related responsibilities or experiences (i.e. curriculum specialists, guidance counselors, assessment directors). Raters’ years of teaching experience reported by grade level and the percentage currently teaching are included in Table 42.

Table 42

Raters’ Years of Teaching Experience by Grade Level and Percentage of Raters Currently Teaching

	Years of Teaching Experience	Raters Currently Teaching
Grade 4	17	73%
Grade 8	20	63%
Grade 11	18	57%
Average of All Grade Levels	18	64%

Among the of raters across all grade levels who were currently teaching, 43% were teaching at the grade level being assessed, with 47% teaching Grade 4, 42% teaching Grade 8, and 41% teaching Grade 11. Raters for the Grade 4 assessment who were currently teaching but not at that grade level reported teaching within a broader range of elementary grades, K-6. The same

was true for raters currently teaching but not at Grade 8 or Grade 11, with the majority teaching within a broader range of middle school grades, 6-8, and high school grades, 9-12, respectively.

An average of 49% raters across all grade levels currently teaching reported the subjects or content areas they teach as those directly related to writing (i.e. English, language arts, reading, literature, English as a Second Language) with 50% of Grade 4 raters teaching content directly related to writing, 50% of Grade 8 raters, and 46% of Grade 11 raters. Other subject or content areas taught by raters included social studies, math, science, and special education.

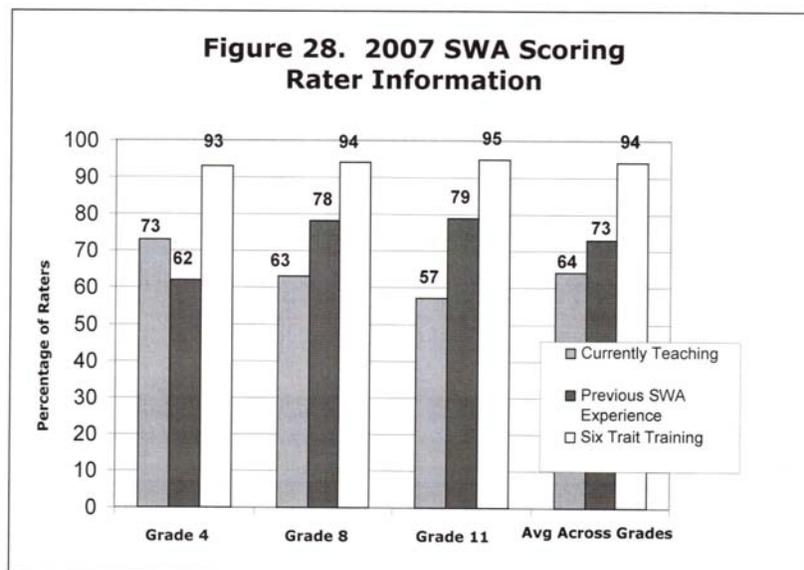
Across all grade levels a majority of raters, 73%, had participated in previous Statewide Writing Assessment scoring sessions with 62% at Grade 4, 78% at Grade 8, and 79% at Grade 11. Of the raters who had previously participated in state level scoring, 16% participated in 2001, the pilot year of the assessment, 13% in 2002, and 20% in 2003, 47% in 2004, 53% in 2005, and 66% in 2006. Raters’ previous Statewide Writing Assessment scoring experience by grade level and year are included in Table 43.

Table 43

Raters’ Previous Statewide Writing Assessment Scoring Experience by Grade Level and Year

	2001	2002	2003	2004	2005	2006
Grade 4	11%	12%	8%	37%	44%	58%
Grade 8	21%	18%	34%	55%	58%	67%
Grade 11	17%	9%	17%	50%	57%	72%
Average of All Grade Levels	16%	13%	20%	47%	53%	66%

A majority of raters across all grade levels, 83%, also reported previous experience within their school districts or at regional Educational Service Units in large scale writing assessment scoring, with 81% at Grade 4, 85% at Grade 8, and 83% at Grade 11. Similarly, a majority, 94%, reported having participated in some form of training in the Six Trait writing assessment model prior to the 2007 state scoring. A composite of rater experience is shown in Figure 28.



Relationship Between Raters’ Scoring Experience and Their Classroom Practices

Survey items related to the raters’ scoring experience and their own classroom practices asked them to indicate their levels of agreement and disagreement with how being a rater affirmed their current instructional practices, increased their understanding of the qualities of good writing, increased their confidence to provide meaningful feedback to students, and increased their confidence as teachers of writing. A single constructed response item in this section of the survey also asked raters to explain why they wanted to be a rater for the Statewide Writing Assessment. Many of the responses by raters to this item re-iterated or elaborated on those provided for the selected response items in this section of the survey. Some of the responses to this item are included in addition to the results provided in the tables.

Among raters across all grade levels, a majority, 94%, agreed or strongly agreed that being a rater increased their understanding of the qualities of good writing, and 97% reported that being a rater built their confidence in making accurate ratings of the quality of student writing. Similarly, 94% of raters across all grade levels agreed or strongly agreed that the scoring experience helped them develop more accurate grade-level expectations for student writing. For each of these three items, more Grade 4 raters than Grade 8 or Grade 11 raters agreed or strongly agreed. See Tables 44 and 45 for raters’ levels of agreement or disagreement related to these aspects of being a rater.

Table 44

Being a rater has increased my understanding of the qualities of good writing

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	1%	28%	71%
Grade 8	0%	1%	3%	22%	74%
Grade 11	1%	2%	9%	29%	59%
Average of All Grade Levels	0%	1%	4%	26%	68%

Rater Comments:

“I wanted to experience reading other papers from students outside of my district. I also wanted to refresh my understanding of the qualities of what a good paper involves.”
(Grade 4 Rater)

“The rating enables me to thoroughly understand the rubric.”
(Grade 8 Rater)

“The process has affirmed my belief that by using a good rubric people of varying backgrounds can assess writing”
(Grade 11 Rater)

Table 45

Being a rater has increased my confidence in making accurate ratings of the quality of student writing

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	1%	20%	79%
Grade 8	0%	0%	4%	21%	75%
Grade 11	0%	0%	6%	27%	67%
Average of All Grade Levels	0%	0%	3%	23%	74%

Rater Comments:

“Being a rater has greatly improved my ability to rate without bias. It is very beneficial to see other writers, not just my own. Seeing other students’ writing helps me teach my students differently.” (Grade 4 Rater)

“I enjoy scoring 8th grade (especially descriptive) because I teach descriptive writing in my classroom. This helps me evaluate my own students’ writing better and gives me an idea where they are headed for state assessment writing in 8th grade.” (Grade 8 Rater)

“Being a rater helps me to better instruct my students in preparation for the State Writing Assessment. It also makes me a better assessor of my own students’ papers.” (Grade 11 Rater)

A majority of raters across all grade levels, 90%, agreed or strongly agreed that being a rater contributed to their ability to help their students become better writers and affirmed their current writing instruction practices. In addition, 94% of raters across all grade levels agreed or strongly agreed that being a rater helped them develop more accurate grade-related expectations for student writing. Grade 4 raters, more than Grade 8 or Grade 11 raters, agreed or strongly agreed that the scoring experience provided these benefits. Related to these benefits also was the belief by the majority that being a rater gave them new ideas they could use in the teaching of writing. See Tables 46, 47, 48, and 49 and comments related to these items.

Table 46

Being a rater has contributed to my ability to help my students become better writers

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	5%	26%	69%
Grade 8	0%	0%	11%	19%	70%
Grade 11	1%	1%	13%	33%	52%
Average of All Grade Levels	0%	0%	10%	26%	64%

Rater Comments:

“The knowledge I receive as a rater about the qualities of a good paper helps me to be able to teach my students how to be great writers. I also enjoy speaking with teachers from other districts about the tools they use to teach their students writing.” (Grade 4 Rater)

“I wanted to be a rater so that I can make sure I’m stressing the best form, process and style of writing to my students.” (Grade 8 Rater)

“As a first year teacher, I have made it a goal to become as proficient in the six traits and assessing as possible. State grading, as well as the regional assessment, have given me tools I can take into the classroom.” (Grade 11 Rater)

Table 47

Being a rater has affirmed my current classroom writing instruction practices

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	4%	35%	61%
Grade 8	0%	1%	14%	16%	69%
Grade 11	0%	1%	11%	32%	56%
Average of All Grade Levels	0%	0%	10%	28%	62%

Rater Comments:

“I wanted to become a rater because I feel it’s a good way to reaffirm my knowledge in my classroom content and teaching methods.” (Grade 11 Rater)

[Being a rater helps me] “to serve as a resource within our district and to confirm that what we are doing is on track. My experience here confirms this.” (Grade 11 Rater)

“I wanted to be a rater because I wanted to see if my classroom writing practices were valid processes in comparison to the rest of the state.” (Grade 4 Rater)

Table 48

Being a rater has given me new ideas I can use in the teaching of writing in my classroom

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	1%	14%	35%	50%
Grade 8	1%	1%	16%	31%	51%
Grade 11	1%	7%	21%	31%	40%
Average of All Grade Levels	1%	3%	17%	32%	47%

Rater Comments:

“I like the professional development I receive by rating the assessment. As a classroom teacher, I always get ideas how to motivate and teach my students to become better writers.” (Grade 8 Rater)

“It also helps me bring back helpful info for my own students.” (Grade 8 Rater)

“State grading, as well as the regional assessment, have given me tools I can take into the classroom.” (Grade 11 Rater)

“The involvement has helped/informed and given me ideas that I can share with classroom teachers.” (Grade 11 Rater)

“I sat with a reading coordinator and learned so much about what we could use as texts and programs to help our reading curriculum in our school.” (Grade 11 Rater)

Table 49

Being a rater has helped me develop more accurate grade-related expectations for student writing

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	0%	29%	71%
Grade 8	0%	1%	7%	22%	70%
Grade 11	0%	2%	9%	33%	56%
Average of All Grade Levels	0%	1%	5%	28%	66%

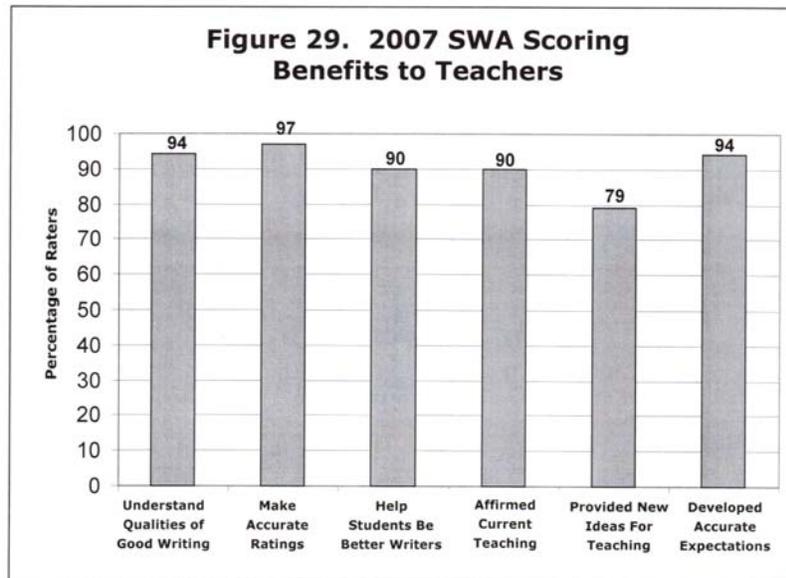
Rater Comments:

“My main reason for being a rater was to be able to see other 4th grade students' writing so that I could compare where my students fall, and therefore be able to be more effective in my writing instruction in the classroom.”
 (Grade 4 Rater)

“It is a valuable experience for teachers to gather together and have discussions about children's writing. It gives me a better picture of how my kids compare with other kids from around the state.”
 (Grade 4 Rater)

“I want to be a rater because I believe reading student work from all areas of the state at the grade level I teach helps me be more accurate in my expectations in my classroom. I am reassured that my practices are accurate.”
 (Grade 8 Rater)

A composite of rater responses regarding classroom practices is shown in Figure 29.



Raters reported that being a rater helped them better understand how writing is related to thinking and reasoning, with 85% of raters across all grade levels citing this understanding. Eighty-two percent (82%) of raters also agreed that being a rater gave them a greater appreciation of writing’s uses in content area learning. For both of these items, Grade 4 and Grade 8 raters more than Grade 11 raters, agreed or strongly agreed. See Tables 50 and 51 for responses to these items by grade level.

Table 50

Being a rater has helped me better understand how writing relates to thinking and reasoning

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	4%	4%	42%	50%
Grade 8	0%	4%	11%	32%	53%
Grade 11	2%	5%	17%	33%	43%
Average of All Grade Levels	1%	4%	11%	36%	49%

Rater Comments:

“Being a rater has expanded and affirmed my conviction that writing IS thinking. I at times need this validation as I’m seen as the ‘crazy expert’ in my building simply because I continue to learn and share my experiences.” (Grade 8 Rater)

“You learn so much about the thinking and reasoning of the grade levels [from being a rater].” (Grade 11 Rater)

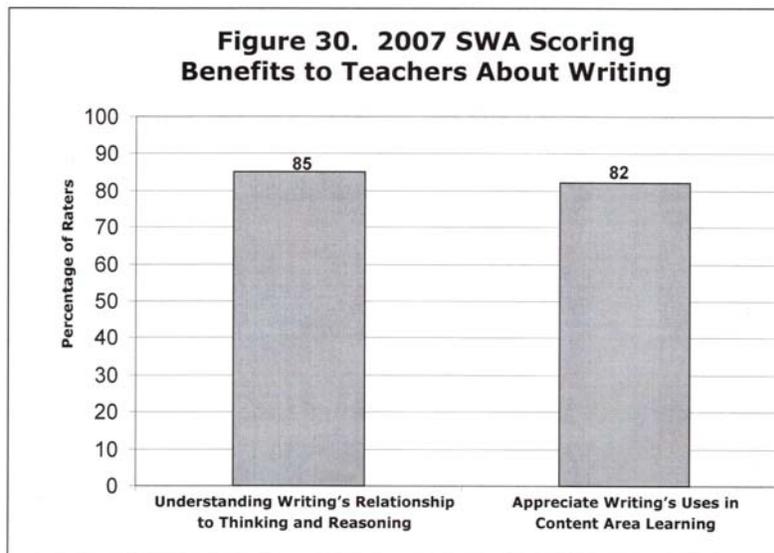
Table 51

Being a rater has given me greater appreciation of writing’s uses in content area learning

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	1%	3%	11%	32%	53%
Grade 8	0%	3%	10%	33%	54%
Grade 11	1%	6%	19%	31%	43%
Average of All Grade Levels	1%	4%	13%	32%	50%

Rater Comments:

“As a past rater, I’ve always returned to my classroom with a renewed sense of worth, accomplishment, and interest in my content area.” (Grade 11 Rater)



A composite of rater responses regarding writing is shown in Figure 30.

Building teacher confidence was cited by a majority of raters across all grade levels as a result of participating in the Statewide Writing Assessment scoring. Ninety percent (90%) of raters indicated that being a rater increased their general confidence as a teacher of writing, and 89% reported that being a rater built their confidence in giving students helpful feedback on their writing. In addition, 86% of raters agreed or strongly agreed that being a rater gave them assurance in talking with parents and others about writing and its importance. See Tables 52, 53, and 54 for responses to these items by grade level.

Table 52

Being a rater has increased my general confidence as a teacher of writing

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	3%	32%	65%
Grade 8	0%	0%	10%	24%	66%
Grade 11	1%	1%	16%	27%	55%
Average of All Grade Levels	0%	0%	10%	28%	62%

Rater Comments:

“I wanted to participate in state scoring to enhance my skills as a teacher. The insights I gain as a rater translate directly to my teaching of writing.” (Grade 4 Rater)

“Helps me professionally - I can go back to my classroom with more confidence in my knowledge and ability in teaching.” (Grade 8 Rater)

“I wanted to see the standards at the state level. Being a rater has improved my skills as a writing teacher.” (Grade 8 Rater)

“My first reason was to become a better writing teacher - it worked! It's also useful to discuss the writing process with other educators.” (Grade 8 Rater)

“Every rater I have ever spoken to feels that this experience gives them so many ideas, enriches their background and helps them develop and improve their instruction.” (Grade 11 Rater)

Table 53

Being a rater has built my confidence in giving my students helpful feedback on their writing

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	4%	28%	68%
Grade 8	0%	0%	10%	26%	64%
Grade 11	1%	1%	15%	28%	55%
Average of All Grade Levels	0%	0%	10%	27%	62%

Rater Comments:

“Scoring the writing assessments is a very effective method to improve the writing instruction. It allows an instructor to improve in feedback with a student by giving positive comments. This is a very important process.” (Grade 8 Rater)

“I have been involved in the writing scoring for 3 years now. Each year it has given me a better and better idea of approaches to writing. It makes it easy to discuss writing with my kids. It also validates to them, the students, that I do know what I'm talking about.”

(Grade 4 Rater)

Table 54

Being a rater has given me assurance in talking to parents and others about writing and its importance

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	1%	6%	36%	57%
Grade 8	0%	1%	15%	20%	64%
Grade 11	0%	3%	15%	28%	54%
Average of All Grade Levels	0%	2%	12%	28%	58%

Rater Comments:

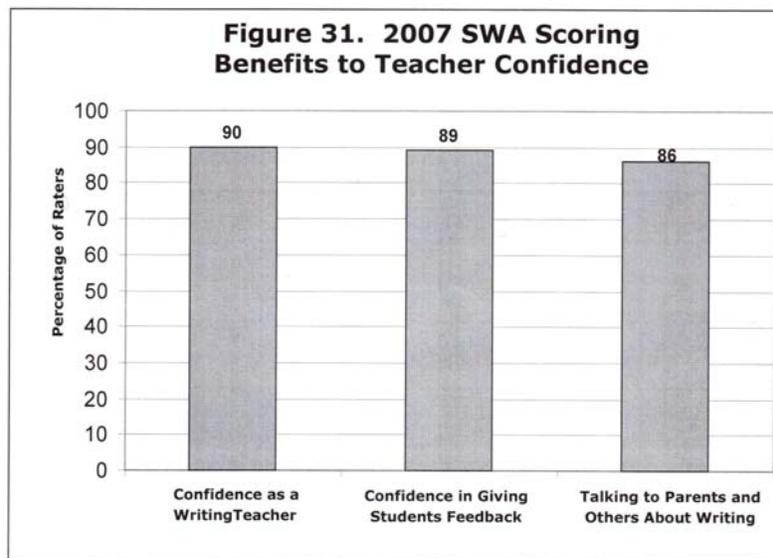
“The experience of being a rater has given me the depth of knowledge to conference with students and parents regarding strengths and areas of concern for a student's writing.”

(Grade 4 Rater)

“Having scored for several years, I lead professional development for writing and scoring at my school. The experience has definitely given me the confidence and knowledge base to conduct these meetings.”

(Grade 4 Rater)

A composite of rater scores relating to teacher confidence is shown in Figure 31.



Raters’ Response to the Training Content and Procedures

The scoring sessions for the Statewide Writing Assessment were held at Educational Service Unit #3 in Omaha in March of 2007. Each scoring session consisted of three days. Primary training for raters was included in the first day of each scoring session with retraining or recalibration procedures occurring throughout the scoring sessions.

The content of the primary training included a general overview of scoring procedures and the holistic scoring process, explanation of the scoring rubric and activities designed to reinforce rater understanding of the content of the rubric and strategies for using the rubric to assign scores, use of anchor papers to exemplify performance levels, practice scoring of blind scored papers, and consensus scoring rounds of actual papers by table groups. The same training format, including all these components, was used for each grade level. The content of the training varied from grade level to grade level only with relationship to the grade specific writing prompts and mode specific anchor papers. The length of this training for raters in each scoring session was three hours. Across all grade levels 98% of raters agreed or strongly agreed that the training content was appropriate with 100% at Grade 4, 99% at Grade 8, and 97% at Grade 11. Raters' level of agreement or disagreement with the appropriateness of the training content is included in Table 55.

Table 55
Raters' Level of Agreement or Disagreement Related to the Appropriateness of the Training

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	0%	0%	27%	73%
Grade 8	1%	0%	0%	30%	69%
Grade 11	0%	0%	3%	29%	68%
Average of All Grade Levels	0%	0%	1%	29%	70%

In addition to the primary training raters received, scoring room leaders conducted re-training and re-calibration activities to ensure accuracy of scoring throughout the remainder of each scoring session. The content of these procedures included reviewing the rubric and anchor papers and scoring of additional blind scored papers for purposes of rater re-calibration. Across all grade levels 97% of raters agreed or strongly agreed these re-training procedures were effective with 97% at Grade 4, 98% at Grade 8, and 95% at Grade 11. For raters' levels of agreement or disagreement on the effectiveness of the re-training procedures see Table 56.

Table 56
Raters' Level of Agreement or Disagreement on the Effectiveness of Re-training Procedures

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	3%	0%	21%	76%
Grade 8	0%	2%	0%	24%	74%
Grade 11	0%	5%	0%	26%	69%
Average of All Grade Levels	0%	3%	0%	24%	73%

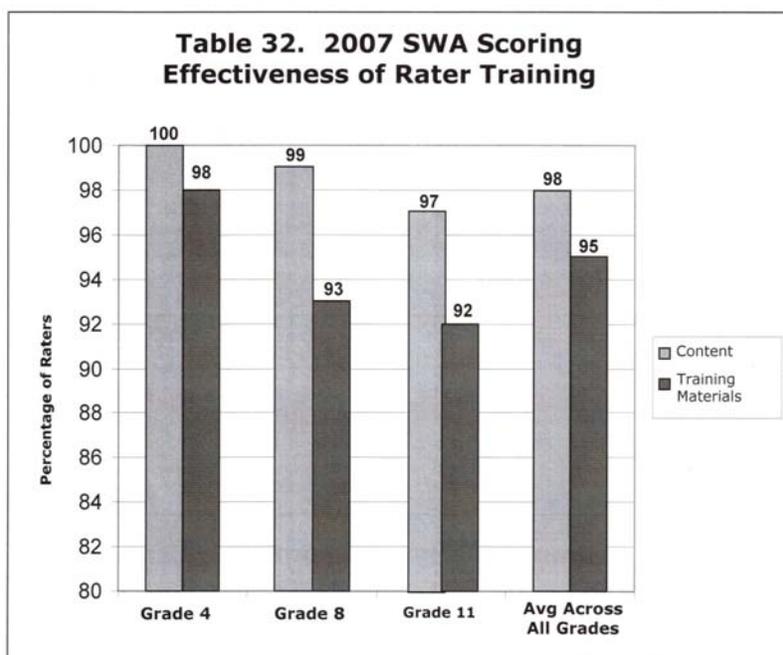
A variety of scoring materials were provided by the Nebraska Department of Education for use by raters throughout the training and scoring. These materials included the "2007 Nebraska Statewide Writing Assessment Scoring Manual," which contains explanations of all scoring procedures, mode-specific rubrics, anchor papers for each grade level, information about the characteristics of writing modes, score point scale worksheets, and individual scoring keys. Use of these materials by raters was explained during the training and scoring by scoring room leaders and through a video presentation by the NDE Coordinator of the Statewide Writing Assessment. Ninety-five percent (95%) of raters across all grade levels agreed or strongly

agreed the scoring materials were clear and understandable. For rater levels of agreement or disagreement related to the clarity of the scoring materials see Table 57.

Table 57
Raters' Levels of Agreement or Disagreement Related to the Clarity of Training/Scoring Materials

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Grade 4	0%	1%	1%	21%	77%
Grade 8	0%	3%	4%	27%	66%
Grade 11	1%	4%	3%	23%	69%
Average of All Grade Levels	0%	3%	3%	24%	71%

A composite of rater responses regarding both content and training materials is shown in Figure 32.



In addition to raters, other scoring personnel in each scoring session included scoring room leaders and table leaders. Scoring leaders and table leaders participated in training sessions specific to their roles and responsibilities on the day prior to each scoring session. For each grade level of scoring in 2007 there were five scoring leaders, one for each of the scoring rooms. The number of raters in each scoring room ranged from 44 to 16, with raters assigned to scoring table groups of 4 in each scoring room.

Scoring room leaders' responsibilities were to facilitate the training on the rubric, modes of writing and anchor papers and modes of writing. In addition, they were responsible to conduct re-training procedures for raters as well as to supervise the over all scoring process in their respective rooms. Scoring room leaders were selected on the basis of their previous state level scoring experiences, knowledge of the scoring process, presentation and training expertise,

familiarity with student writing at the grade level being scored, and their understanding of the characteristics of mode specific writing.

For each scoring session there were 13 table leaders, one for every 2-3 table groups of raters. In each scoring room, table leaders were assigned to assist with re-training procedures for raters at their assigned tables, to monitor the scoring accuracy of raters at their assigned tables, and to conduct adjudication readings of papers that received discrepant scores. Table leaders were selected on the basis of their previous state level scoring experiences, their knowledge of the scoring process, their familiarity with student writing at the grade level being assessed, and their understanding of the characteristics of mode specific writing.

Ninety-eight percent (98%) of raters across all grade levels agreed or strongly agreed that scoring leaders demonstrated expert knowledge of the assessment and scoring procedures. Similarly, 95% of raters across all grade levels agreed or strongly agreed that table leaders demonstrated expert knowledge of the scoring procedures. For raters' levels of agreement and disagreement related to the effectiveness of scoring leaders and table leaders see Table 58.

Table 58
Raters' Levels of Agreement or Disagreement Related to the Expert Knowledge Demonstrated by Scoring Room Leaders and Table Leaders

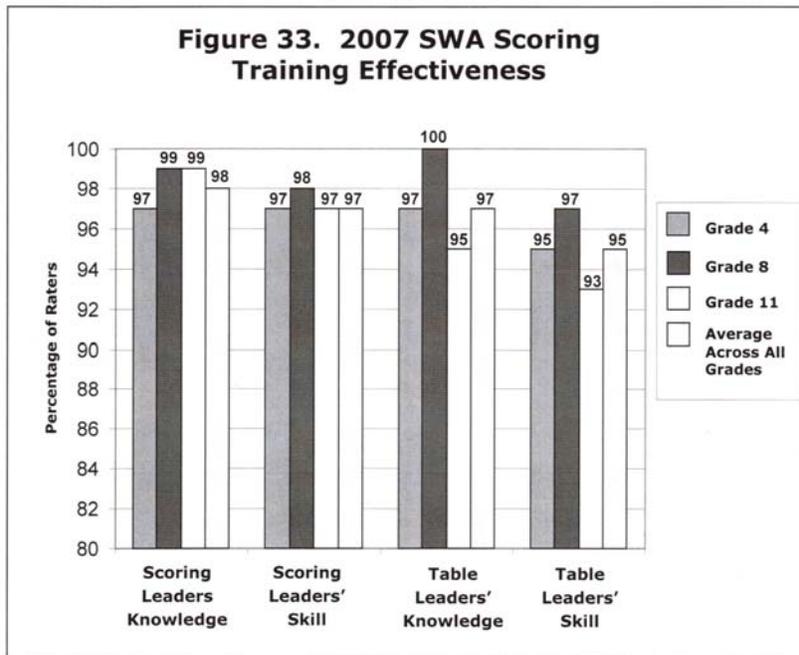
	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Scoring Leaders					
Grade 4	0%	0%	2%	13%	84%
Grade 8	0%	0%	1%	21%	78%
Grade 11	0%	0%	1%	17%	82%
Average of All Grade Levels	0%	0%	1%	17%	81%
Table Leaders					
Grade 4	1%	1%	1%	17%	80%
Grade 8	0%	0%	0%	28%	72%
Grade 11	0%	2%	8%	22%	68%
Average of All Grade Levels	0%	1%	4%	21%	74%

Raters across all grade levels agreed that scoring leaders were skilled in teaching them how to score the assessment and that table leaders were effective in facilitating the scoring process through re-calibration procedures (i.e. read behinds and resolving issues of discrepant scoring). See Table 59 for Raters' levels of agreement or disagreement related to the skill of scoring leaders and table leaders. A composite of the data contained in Tables 58 and 59 is shown in Figure 33.

Table 59

Raters' Levels of Agreement or Disagreement Related to the Skill of Scoring Room Leaders and Table Leaders

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Scoring Leaders					
Grade 4	0%	2%	1%	21%	76%
Grade 8	0%	0%	1%	28%	70%
Grade 11	0%	1%	2%	21%	76%
Average of All Grade Levels	0%	1%	1%	23%	74%
Table Leaders					
Grade 4	1%	1%	3%	19%	76%
Grade 8	0%	0%	3%	24%	73%
Grade 11	0%	2%	5%	18%	75%
Average of All Grade Levels	0%	1%	4%	20%	75%



SUMMARY OF FINDINGS

This study was designed to ascertain the relationship between the 2007 Statewide Writing Assessment scoring experience of raters and their classroom practices as teachers and to determine the effectiveness of the procedures used to train them to score the assessment. Responses collected from raters pertained to broad categories: general rater information, scoring training, and scoring experience related to their teaching practices.

Rater Information

Raters of the 2007 Statewide Writing Assessment were experienced teachers with a majority possessing a master's degree or higher and with an average of 18 years of teaching experience.

Sixty-four percent (64%) of raters were currently teaching at or near the grade levels being assessed and a majority were teaching content areas directly related to writing (i.e. English, language arts, reading, literature). Raters currently not teaching were school administrators, educators with education related responsibilities outside the classroom (i.e. curriculum and assessment directors), substitute teachers, and student teachers.

A majority of teachers, 73%, came to the scoring sessions with previous state level scoring experience, and 83% reported previous scoring experience within their school districts or at regional Educational Service Units. In addition, 83% of raters reported prior training in the six trait writing assessment model.

Relationship Between the Raters' Scoring Experience and Their Classroom Practices

Hillocks' (2002) study of five states' writing test provides important information about the relationship between state level writing tests and their effects on local curriculum decisions. Even though he concludes that the expectations of state writing tests tend to override local decisions of pedagogy which may be based on best practices for the teaching of writing, the teachers he interviewed in each of the state's whose writing assessments he examined generally reported favorable responses to their respective state's test and the relationship they perceived between it and their classroom practices.

What his study does not examine are the perceptions of teachers who were actually involved in the scoring of the assessment about the relationship of the scoring experience and their classroom practices.

Participation in the scoring of the 2007 Statewide Writing Assessment was perceived by a majority of raters across all grade levels as beneficial to their classroom practices in the teaching of writing. Responses by raters were related to increasing their understanding of the qualities of good writing, building their confidence as a teacher of writing, affirming their current classroom practices, providing helpful feedback to students about their writing, and other areas related to the teaching of writing.

Ninety-seven percent (97%) reported that being a rater increased their confidence to assess the quality of student writing, and 94% believed their scoring experience increased their understanding of the qualities of good writing and helped them develop more accurate grade-related expectations for student writing.

Across all grade levels, 90% of raters agreed or strongly agreed that being a rater affirmed their current teaching practices, contributed to their ability to help students become better writers, and increased their general confidence as a teacher of writing. A majority of raters, 89%, agreed that being a rater gave them confidence to give students helpful feedback about their writing, and 86% reported that being a rater gave them assurance in talking to parents and others about writing and its importance.

Eighty-five percent (85%) of the raters across all grade levels had a better understanding of how writing relates to thinking and reasoning, and gave 82% of raters a greater appreciation of writing's uses in content area learning. Seventy-nine percent (79%) of raters agreed or strongly agreed that being a rater gave them new ideas they could use in the teaching of writing.

Scoring Training

An overwhelming majority of raters across all grade levels agreed or strongly agreed that all aspects of the scoring training were effective. Ninety-nine percent (99%) of raters agreed or strongly agreed that the general content of the training (i.e. the scoring process, scoring criteria, six traits of writing, writing modes) was appropriate while 97% indicated the strategies to ensure scoring accuracy (i.e. using the rubric to assign scores, using anchor papers to demonstrate performance levels, use of re-calibration papers, and consensus building discussions) were effective. Ninety-five percent (95%) of raters agreed that the training materials were clear and understandable.

A majority of raters, 98%, agreed or strongly agreed that scoring leaders and 97% agreed or strongly agreed that table leaders demonstrated expert knowledge of the assessment and the scoring process. Ninety-seven percent (97%) of raters agreed that scoring leaders were skilled in teaching them to score the assessment, and 95% agreed that table leaders were effective in facilitating the scoring process through re-calibration procedures.

Conclusions

Based on the results of this study, it is possible to draw important conclusions about the raters and their perceptions of the scoring process. First, the scoring experience was perceived by a majority of raters as a valuable professional development opportunity that contributed to their knowledge and understanding of the teaching and assessment of student writing. Secondly, raters' perceptions about the effectiveness of the scoring training were very positive across all grade level scoring groups. What this means is that the training, including the content, materials, training personnel, and procedures, was appropriately designed and delivered in such a way that raters felt they were adequately prepared to score the assessments.

Recommendations

The results of this study provide valuable information to the Nebraska Department of Education related to its ongoing implementation of the scoring of the Statewide Writing Assessment. The positive responses of raters related to the scoring training should serve as a recommendation to continue to develop and utilize high quality training plans and materials and to continue to enlist highly qualified individuals to implement the training procedures. The positive responses of raters to the scoring experience in relationship to their classroom practices should also serve as a recommendation to the NDE to continue to utilize the state's teachers for this assessment initiative as a way to assist in developing classroom expertise in the teaching of writing statewide. It should also serve as a recommendation to local school districts to send teachers to the state level scoring sessions. For in doing so, their students will ultimately reap the benefits of the teachers' learning derived from the scoring experience.

REFERENCES

- Arter, J. and McTighe, J. (2001). *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance*. Thousand Oaks, California: Corwin Press.
- Gallagher, C. W. (2007). *Reclaiming assessment: A better alternative to the accountability*

agenda. Portsmouth, New Hampshire: Heinemann.

Heinemann. (2003). *Charting STARS-sustainability as challenge and opportunity: Research study and comprehensive evaluation of Nebraska's school-based, teacher-led assessment and reporting system (STARS)* (Tech Rep. No 2). Lincoln, Nebraska: University of Nebraska-Lincoln. Teachers College and College of Arts and Sciences.

Hillocks, G., Jr. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.

Freedman, S. W., & Diaute, C. (2001). "Instructional methods and learning in teaching writing." In *Subject-specific instructional activities and methods*. Vol. 8. New York: Elsevier Science.

Mabry, L. (1999). "Writing to the rubric." *Phi Delta Kappan*. May 1999, pp. 673-679.

National Writing Project and Nagin, C. (2003). *Because writing matters: Improving student writing in our schools*. San Francisco, California: Jossey-Bass.

Pike-Baky, M., & Fleming, G. (2005). *Prompted to write: Building on-demand writing skills*. San Francisco: Jossey-Bass.

Spandel, V. (2005). *Creating writers through 6-Trait writing assessment and instruction*. Boston: Pearson Education, Inc.

Spandel, V., & Stiggins, R. (1997). *Creating writers: Linking writing assessment and instruction*. New York: Addison Wesley Longman.

Wiggins, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

**CHARTING STARS:
PATHWAYS TO EXCELLENCE
Section 3: Research Studies**

**Study IX: The Effect of an Intensive Writing Instruction
Assessment Preparation Program on Student Writing
Proficiency**



**THE EFFECT OF AN INTENSIVE WRITING INSTRUCTION
ASSESSMENT PREPARATION PROGRAM ON
STUDENT WRITING PROFICIENCY**

Dr. Edward J. Bennett, Assistant Principal
Omaha Central High School, Omaha, Nebraska

INTRODUCTION

The Nebraska legislature holds school districts accountable for the quality of writing instruction through State Statute 79-760 which requires all public schools to participate in yearly Statewide Writing Assessments (NSWA). The results of the NSWA, which is given to all students in grades 4, 8, and 11, are published annually by all major newspapers in the state and are posted on the Nebraska Department of Education's website. The scores are used to determine the percent of proficient writers at the specified grade levels in each school. These proficiency levels also are one factor in determining if schools make Adequate Yearly Progress (AYP). Failure to make AYP can result in severe consequences for a school, including loss of funding and eventual surrender of district control.

STUDY PURPOSE

The purpose of this pretest-posttest two group comparative survey study was to determine whether an intensive writing instruction and assessment preparation program (IWIAPP) was effective in helping 8th grade students improve their performance on the Nebraska Statewide Writing Assessment (NSWA).

RESEARCH DESIGN AND METHODOLOGY

To help students succeed on the NSWA, the research district developed its own Fall District Writing Assessment (FDWA) which is administered to all district students in grades 4, 8, and 11 each September. District personnel score the FDWA in a consistent fashion to the NSWA, and scores are reported to individual buildings to help inform the writing instruction in the weeks between FDWA and NSWA. From a study design perspective, FDWA and NSWA created a natural pretest-posttest opportunity in which the effectiveness of a school's writing preparation program could be measured and analyzed.

The students who took part in this study were 8th-graders attending a middle school serving 950 students in grades 5 through 8. Although all students took part, Special Education students, ESL students, and all students not attending the research school during 7th grade were not included in the unit of analysis. In addition to FDWA and NSWA scores, the study also analyzed California Achievement Test language mechanics and language expression subtest Normal Curve Equivalence (NCE) scores from grades 7 and 8 and classroom behavior as measured by office referrals and tardy to class frequencies.

The six week intensive writing instruction and assessment preparation program (IWIAPP) was administered to all students whose scores ranged from 2-4 on the FDWA ($N = 74$) during regular classroom time by both of the research school's 8th-grade language arts teachers. Identified students spent approximately 90 minutes per week participating in the program. The three unique characteristics of the program were (a) explicit instruction on the writing process (b) explicit instruction by the teacher on one trait of the Six Traits of Writing per week (c) consistent use of the Six Trait based State Assessment Scoring Rubric by students and teachers. Students in the IWIAPP program were required to draft, revise, and finish at least three original pieces of descriptive writing during the 6 weeks. The goal was that each of those pieces, after thorough revising and editing, met or exceeded the holistic score of 4.33 when scored by the teacher in a manner consistent with the FDWA and the NSWA.

The components of the program were based, in part, on Gersten and Baker's 2001 meta-analysis which identified three key components of writing programs proven to help struggling writers: (a) adhering to a basic framework of planning, writing, and revision, (b) explicitly teaching critical steps in the writing process, and (c) providing feedback guided by the information explicitly taught. Additional support for process writing was provided by an analysis of the 1998 NAEP Writing Assessment results which revealed that 8th and 12th grade students who were always asked to write more than one draft of a paper had higher average scores than peers who were sometimes or never asked to do so (National Center for Education Statistics, 1999). Voss (2001) found that students whose teachers used the process approach to writing instruction scored significantly higher on a state-mandated writing assessment.

Research in recent years has indicated that teaching writing using the process approach can be enhanced when students and teachers use a common language when discussing writing (Spandel & Stiggins, 1997). In fact, the National Council of Teachers of English included the need for a common language as one of its most important beliefs about the teaching of writing (NCTE, 2004). Calkins (1994) defined this need by pointing out how students cannot talk easily and clearly about how to improve their own or other's writing unless they know the language of editing and revision. In fact, the lack of a common language can actually slow students' progress (Perchemlides & Coutant, 2004). A common language enables students and teachers to communicate using the same terminology, turning both into writing teachers and coaches and helping both come to a better understanding of how writing works (Spandel, 2005).

During the past twenty years, Northwest Regional Educational Laboratory (NWREL) has taken the lead in developing, specifying, and articulating a common language method that describes good writing. That method, known as the Six Traits of Writing, has served as a reliable reference to help guide writers through the entire writing process, while simultaneously improving their writing skills and confidence (Isernhagen & Kozisek, 2000; Kowalewski, Murphy, & Starns, 2002). Perchemlides and Coutant (2004) have gone so far as to say that they have yet to come across a teaching approach that better addresses a need for a common language.

Reeves (2002) asserted that the proper use of rubrics makes final work submitted to teachers far more reflective of proficient performance. It also sends a clear message to the students that they are responsible for frequent evaluation and improvements to their work prior to submission (Reeves, 2002). Apparently when teachers conceptualize the classroom progress of students from the perspective of a rubric, as opposed to the perspective of accumulated total points, they think more in terms of the knowledge and skill strengths and weaknesses of their students (Marzano, 2000). This perspective is apparently passed on to students who in turn improve their performance (Marzano, 2000). Grading students with rubrics also seems to improve grading accuracy when compared to the traditional point total grading method. Correlations between rubric scores and outside standardized test scores are higher than correlations of grades with standardized tests (Marzano, 2000).

The IWIAPP model's basic framework for planning, writing, and revision as part of the writing process was provided by enriched POWERUPS Notemaking Sheets (Henderson, 2005). POWERUPS is an acronym for eight critical steps in the writing process: prewriting, organizing, writing, evaluating, revising, underlining, proofreading, and submitting. Each POWERUPS Notemaking Sheet includes blank lines for writing as well as helpful hints on the writing process and opportunities for self-evaluation in the left and bottom margins. For all three descriptive writing drafts, students used a POWERUPS Notemaking Sheet.

Feedback on student writing was provided by the students themselves, their peers, and their teacher as part of a writer's workshop format (Atwell, 1987). The workshop included a daily mini-lesson by the teacher on one of the Six Traits. At least 30 minutes was set aside during each class purely for quiet writing time, and at least 15 minutes during each class for peer conferencing and informal conferencing with the teacher. In addition, the teacher held at least one formal, individual writing conference with each IWIAPP student during the program to suggest explicit revisions in both content and form of student writing as recommended by Gersten and Baker (2001). Each teacher also held many individual informal writing conferences with students while circulating throughout the room during writing workshop time. Dictionaries

and thesauruses were provided for students as well as plenty of writing utensils, style manuals, and adolescent novels. Students were encouraged to choose their own topics for their three writing pieces, but the teacher also provided topic choices for students who needed them. The teacher approved all writing choices to ensure they were descriptive in nature.

To assist students each week as they improved their skills on each trait, a laminated copy of the State Assessment Scoring Rubric was provided to each IWIAPP student. Laminated copies were also taped to each of the student tables in each classroom. The rubric clarifies the description of what effective writing looks like for each trait at each of the four scoring levels and provides a common language for discussing student writing. The language of the State Assessment Scoring Rubric is based on the Six Traits of Writing and is also embedded into the POWERUPS Notemaking Sheets.

Students whose scores ranged from 5-8 on the FDWA ($N = 86$) comprised the control group of the study. During regular classroom time, the control group students engaged in less structured assignments and followed a more standard language arts curriculum. Specifically, they wrote drafts of descriptive writing independently, worked on a portion of a poetry unit, and engaged in free reading of novels of their choosing.

STUDY FINDINGS

Pretest to posttest gains made by IWIAPP students ($N = 74$) revealed significant improvement in writing performance on NSW. The pretest FDWA score ($M = 3.54$, $SD = .80$) compared to the posttest NSW writing score ($M = 5.11$, $SD = .95$) was statistically significantly different, $t(72) = 12.13$, $p = .0001$ (one-tailed), $d = 1.79$. Furthermore, 93% of IWIAPP students improved their scores from FDWA to NSW.

The control group of 8th-grade students ($N = 86$) also demonstrated significant pretest to posttest gains. The pretest FDWA writing score ($M = 5.88$, $SD = .85$) compared to the posttest NSW writing score ($M = 6.13$, $SD = .78$) was statistically significantly different, $t(84) = 2.37$, $p = .01$ (one-tailed), $d = .30$.

Despite the significant gains made from FDWA to NSW, California Achievement Test language mechanics and language expression subtest Normal Curve Equivalence (NCE) scores dipped significantly for both groups in three of the four subtests. However, the achievement qualitative description of these scores at posttest remained in either the average or above average range for all participants. Furthermore, while there was a statistical difference between the groups of students for behavior measures, during the second semester levels of office referrals and tardies would be considered low for middle school students.

DISCUSSION AND RECOMMENDATIONS

While not all IWIAPP students demonstrated proficiency at posttest, the overwhelming majority improved, which indicates the IWIAPP intervention succeeded in improving student writing. The results of the study also serve to underscore the value of the three underlying components of the IWIAPP intervention: explicit instruction in the writing process, explicit instruction in the use of a common language, and frequent references to a scoring rubric clarifying the definitions of effective writing. While it is unclear which of these components had the greatest impact on student performance, the results leave little doubt that the intervention as a whole revealed the

codes of writing to the participants in a way that they internalized and applied with overwhelming success. The study's results should encourage local and state education officials to consider implementing the IWIAPP intervention to help more struggling writers in middle schools across the research district and the state.

REFERENCES

- Atwell, N. (1987). *In the middle: Writing reading, and learning with adolescents*. Portsmouth, NH: Heinemann.
- Calkins, L. M. (1994). *The art of teaching writing* (Rev. ed). Portsmouth, NH: Heineman.
- Gersten, R. & Baker, S. (2001). Teaching expressive writing to students with learning disabilities: A meta-analysis. *Elementary School Journal*, 97, 475-500.
- Henderson, S. (2005, October). *POWERUPS: An on-demand descriptive writing process guide*. Document presented at Secondary Language Arts Curriculum Day for the Omaha Public Schools, Omaha, NE.
- Isernhagen, J. & Kozisek, J. (2000). Improving students' self-perceptions as writers. *Journal of School Improvement*, 2, 3-4.
- Marzano, R. (2000). *Transforming classroom grading*. Aurora, CO: ASCD.
- National Center for Education Statistics. (1999). *The NEAP 1998 writing report card for the nation and the states (NCES 1999-462)*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- National Council of Teachers of English. (2004). *NCTE beliefs about the teaching of writing*. Retrieved Jan. 3, 2006, from www.ncte.org/about/over/positions/category/write/118876htm.
- Perchemlides, N. & Coutant, C. (2004). Growing beyond grades. *Educational Leadership*, 62(2), 54-60.
- Reeves, D. (2002). *Making standards work*. Advanced Denver, CO: Learning Press.
- Spandel, V. (2005). *Creating writers through 6-trait writing assessment and instruction*. Boston: Allyn and Bacon.
- Spandel, V. & Stiggins, R. (1997). *Creating writers: Linking writing assessment and instruction* (2nd ed.). White Plains, NY: Longman.
- Voss, V. V. (2001). The effect of writing process training for teachers on the Texas Assessment of Academic Skills reading and writing scores of students in grades 3-8. (Doctoral dissertation, Texas A&M University-Commerce, 2001). *Dissertation Abstracts International*, 63(01), 75A.

**CHARTING STARS:
PORTRAITS OF EXCELLENCE**

Section 5: Appendices

Appendix A
IRB Approval Letter

January 8, 2007

Dr. Jody Isernhagen
Leon Dappen
132 TEAC
(0360)

IRB # 2006-08-501 EP

TITLE OF PROJECT: **Comprehensive Evaluation of School-based Teacher-led Assessment and Reporting System**

Dear Dr. Isernhagen:

This letter is to officially notify you of the final approval of your project by the Institutional Review Board (IRB) for the Protection of Human Subjects. It is the Board's opinion that you have provided adequate safeguards for the rights and welfare of the participants in this study. Your proposal seems to be in compliance with this institution's Federal Wide Assurance 00002258 and the DHHS Regulations for the Protection of Human Subjects (45 CFR 46).

Date of EP Review: 09/16/06.

You are authorized to implement this study as of the Date of Final Approval: 1/8/07. This approval is Valid Until: 9/21/07.

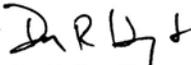
We wish to remind you that the principal investigator is responsible for reporting to this Board any of the following events within 48 hours of the event:

- Any serious event (including on-site and off-site adverse events, injuries, side effects, deaths, or other problems) which in the opinion of the local investigator was unanticipated, involved risk to subjects or others, and was possibly related to the research procedures;
- Any serious accidental or unintentional change to the IRB-approved protocol that involves risk or has the potential to recur;
- Any publication in the literature, safety monitoring report, interim result or other finding that indicates an unexpected change to the risk/benefit ratio of the research;
- Any breach in confidentiality or compromise in data privacy related to the subject or others; or
- Any complaint of a subject that indicates an unanticipated risk or that cannot be resolved by the research staff.

For projects which continue beyond one year from the starting date, the IRB will request continuing review and update of the research project. Your study will be due for continuing review as indicated above. The investigator must also advise the Board when this study is finished or discontinued by completing the enclosed Protocol Final Report form and returning it to the Institutional Review Board.

If you have any questions, please contact Shirley Horstman, IRB Administrator, at 472-9417 or email shorstman1@unl.edu.

Sincerely,



Dan R. Hoyt, Chair
for the IRB



Shirley Horstman
IRB Administrator

Appendix B
Researchers for the Project

Researchers for the Primary Sixth Year Project 2006-2007

Principal Investigator

Dr. Jody C. Isernhagen is an Associate Professor in Educational Administration at the University of Nebraska-Lincoln. She received her doctoral degree from Virginia Tech and has been a teacher, assistant principal, principal, supervisor of elementary education, and superintendent in pre-K through 12 schools. Dr. Isernhagen serves as the primary investigator for the STARS Process and is the primary instructor for the School Improvement Specialist Program, a joint program between the North Central Association on Accreditation and School Improvement (NCA CASI). She serves as the State Accreditation and North Central Accreditation External Leader for four school districts in Nebraska. Dr. Isernhagen was awarded the College of Education and Human Sciences Distinguished Teaching Award.

Secondary Investigators

Sue Anderson is the Director of Professional Development for Educational Service Unit #3 in Omaha and is currently a doctoral student in the College of Education and Human Sciences at the University of Nebraska-Lincoln.

Edward J. Bennett is an Assistant Principal at Omaha Central High School, Omaha, Nebraska. He received his Doctorate from the University of Nebraska at Omaha.

Dr. Shirley J. Mills is an Assistant Professor at the University of Texas-Pan American and served as a secondary investigator for the STARS Comprehensive Evaluation. She taught in Nebraska for 38 years prior to receiving her Doctorate in Leadership and Higher Education from the University of Nebraska-Lincoln in 2005.

Jadi Miller is the Principal at Everett Elementary School, Lincoln, Nebraska and is currently a doctoral student in the College of Education and Human Sciences at the University of Nebraska-Lincoln..

Dr. Allen L. Steckelberg is an Associate Professor in the Department of Teaching, Learning, and Teacher Education at the University of Nebraska-Lincoln. Dr. Steckelberg teaches and coordinates instructional technology and serves as the Technology Coordinator for the College of Education and Human Sciences. Areas of teaching and research include technology in education, Web-based instructional and educational management, and paraprofessionals in school programs.

Lan Li is a doctoral student in Instructional Technology at the University of Nebraska-Lincoln. She teaches courses in Instructional Technology and is currently doing research on improving student learning with technology-mediated peer assessment.

Project Administrative Coordinator

Susan Wilson is on the staff at the University of Nebraska-Lincoln and holds an Associate of Science degree in Business Administration from the College of St. Mary, Omaha.

Interviewers

Larry Bornschlegl, Ed. Specialist, retired from Hastings Public Schools in 2002 after having spent 31 years in Nebraska as a secondary principal. The Nebraska State Association of Secondary School Principals named him the Region IV “Principal of the Year” in 2001. In 2002 he received the “Award of Excellence” from the Nebraska Association of Middle Level Education. In 2004 he received the “Distinguished Service Award” from the Nebraska State Association of Secondary School Principals. He remains active in education.

Ronald Klemke, M.A. retired after 33 years of service to education in Nebraska Panhandle schools. Ron served as a teacher and coach in wrestling, track, and golf, and was inducted into the “Nebraska Scholastic Wrestling Coaches’ Hall of Fame” in 1991. He has served as Chairman of the Board of County Commissioners in Garden County since 2001. He is a member of the Oshkosh Economic Development Board and is active in volunteer work with Saint Elizabeth’s Parish and the Grand Island Diocese Deanery Pastoral Council.

Appendix C
Validation Framework

Validation of Nebraska's Standards, Assessment, and Accountability System

"An Accountability System can be said to have validity when the evidence is judged to be strong enough to support inferences that:

- 1) the components of the system are aligned to the purposes and are working in harmony to help the system accomplish those purposes; and,
- 2) the system is accomplishing what was intended (and did not accomplish what was not intended.)"

The accountability system will be validated as "working" if evidence is found that the following purposes are goals of the accountability system are being met.

Goal One Educators can appropriately and accurately assess and report student performance on content standards using local assessment systems.

Goal Two Student performance and assessment data are being used for school improvement efforts.

Goal Three Student performance in reading, writing, and mathematics have improved because of school improvement efforts based upon assessment and performance data.

Goal Four Does the Nebraska accountability system (both state and AYP) misclassify schools or classify school systems correctly?

Goal Five The consequences of the accountability system are positive ones and contribute to the achievement of the districts in the state accountability goals.

Goal One

Educators can appropriately and accurately assess and report student performance on content standards using local assessment systems.

Indicators for Study	Research Questions	Method	Who?
District Assessment Portfolio	<p>Are the Six Quality Criteria the right criteria for evaluating local assessment?</p> <p>What is the correlation (if any) between DAP ratings and NRT data? CRT data?</p> <p>In districts where assessment ratings are high, what are school leaders doing? What are their characteristics?</p>	<p>Inter-rater reliability of DAP scoring</p> <p>Outside Evaluator of Portfolio Process</p> <p>Evaluation of Process</p> <p>On-site visitations</p> <p>Surveys of local educators</p> <p>Surveys, case studies</p>	<p>External Review Team</p> <p>Nebraska Peer Reviewers</p> <p>NDE</p> <p>UNL Evaluation</p> <p>UNL evaluation</p>
Local district assessments	Are the local assessments used of sufficient quality to accurately measure student performance?	<p>Examination of assessments</p> <p>Peer review sessions</p>	<p>Peer Review Teams</p> <p>Trained teacher teams</p> <p>External Review Team</p>
Consortiums, Collaborations, Individual Districts	How do educators working in groups handle local assessment differently from educators in independent districts?	Survey/case studies	<p>UNL evaluation</p> <p>Nebraska Peer Reviewers</p>

Goal Two

Student performance data generated from assessment and graduation rates are being used to inform local school improvement.

Indicators for Study	Research Questions	Method	Who?
School Improvement Plan	Are districts basing their school improvement goals on collected data from assessment?	Review of school improvement plans External visitations	NDE NDE/Review teams
Professional development	What are reading and math teachers saying about professional development? What are the characteristics of Nebraska's professional development?	Survey Case studies	UNL Evaluation UNL Evaluation UNL Evaluation
Instructional change	How is instruction changing in Nebraska schools as a result of assessment data?	External Visitations Case studies	NDE UNL Evaluation
Leadership	How are local leaders modeling effective leadership assessment practices? How are the leaders in effective schools aligned in their thinking about assessment practices? What do principals need to do the local assessment work?	Case studies Survey Survey Leaders of Learning Focus Group	UNL Evaluation UNL Evaluation UNL Evaluation NDE/UNL

Goal Three

Student performance in reading, mathematics, and writing, as well as graduation rate will improve as a result of data-informed school improvement efforts.

Indicators for Study	Research Questions	Method	Who?
<p>Student performance on reading standards.</p> <p>Student performance on mathematics standards</p> <p>Disaggregated Student Performance</p>	<p>Have NRT and CRT reading data improved from 2001, 2002, 2003, 2004? 2005? 2006? 2007?</p> <p>Have STARS 2002 mathematics data improved from 2002 – 2004 mathematics data? 2005? 2006? 2007?</p> <p>Have NRT and CRT mathematics data improved from 2001, 2002, 2003, 2004? 2005? 2006? 2007?</p> <p>Have NRT and CRT data improved for all groups of students from 2001-2007?</p>	<p>Correlational Studies</p> <p>Correlational Studies</p> <p>Correlational Studies</p>	<p>Research in Action</p> <p>Research in Action</p> <p>Research in Action</p> <p>Research in Action UNL</p>
<p>Student performance on Statewide Writing</p>	<p>Have statewide writing results improved? Grade 4 - 2002/2004/2006 Grade 8 – 2003/2004/2006 Grade 11 – 2004/2006</p> <p>Are Statewide Writing results and reading results comparative?</p> <p>How are statewide writing data being used in classrooms in districts and for programmatic change?</p>	<p>Correlational Studies</p> <p>Case studies</p>	<p>Research in Action UNL</p> <p>UNL UCLA CREEST</p> <p>UNL Evaluation</p>
<p>Graduation Rate</p>	<p>Has the graduate rate improved from 2003-04 to 2005-07?</p>	<p>Correlational Studies</p>	<p>NDE Research in Action</p>

Goal Four

Does the Nebraska accountability system (both state and AYP) classify or misclassify schools?

Indicators for Study	Research Questions	Method	Who?
Internal Data Validation Decision Verification	What does NDE do internally to insure the accuracy of its data? How does NDE know its accountability decisions are accurate?	AAR Business Rules Flow Chart Internal Audit	NDE/Research in Action

Goal Five

The consequences of the accountability system are positive ones that contribute to the achievement of the other four accountability goals.

Indicators for Study	Research Questions	Method	Who?
Intended consequences Emerging consequences	What kinds of things are happening in schools where accountability goals aren't being met? What kinds of things are happening in schools where accountability goals are being met?	Case Studies Visitations Survey Case Studies Survey	UNL Evaluation NDE UNL Evaluation NDE
Unintended Consequences	What have been the intended and unintended consequences in: a) Schools meeting accountability goals b) Schools not meeting accountability goals	Case studies Survey	UNL case studies Visitations – NDE ESU

Appendix D
2006-2007 Study I: Nebraska-led Portfolio Peer
Review Survey

Nebraska-Led Portfolio Peer Review Survey STARS Comprehensive Evaluation Project

INSTRUCTIONS: Please provide the following demographic information by responding to the questions or marking the appropriate category for each area.

1. Mark all the position(s) below that you hold in your school district:

Superintendent Curriculum Coordinator Assessment Coordinator Other: _____

Principal: ELEM MS HS

Teacher: ELEM MS: Subject _____ HS: Subject _____

2. My Gender: Male Female

3. Years of experience with assessment portfolio: _____

4. Total Years of Experience in Education: _____

	None of the time	Very little of time	Some of the time	Most of the time	All of the time
Directions: Please circle the number that best describes your response to each statement.					
ALIGNMENT					
1. Our district assessment items/tasks reflect a match to the appropriate standards.	1	2	3	4	5
2. Our district assessment items/tasks reflect the content and skills found within the standards.	1	2	3	4	5
3. Our district had assessment items reviewed by district personnel.	1	2	3	4	5
4. Our district had assessment items reviewed by external personnel.	1	2	3	4	5
5. Our district has a list of specifications mapping the assessment items to the standards in order to show which items assess which standards.	1	2	3	4	5
6. There is a documentation process for alignment of assessments to standards.	1	2	3	4	5
7. Our district involved staff in the alignment of the assessments to standards.	1	2	3	4	5
8. Our district supports teachers working collaboratively to ensure assessments measure the standards.	1	2	3	4	5
SUFFICIENCY					
9. Our district measures all academic content standards in the assessment items/tasks.	1	2	3	4	5
10. Our assessment items/tasks are distributed across all performance levels.	1	2	3	4	5
11. Our assessment items/tasks use a variety of appropriate formats.	1	2	3	4	5
12. Our assessment items/tasks include higher order thinking skills.	1	2	3	4	5
13. Our district reviewed assessment items/tasks for sufficiency results.	1	2	3	4	5
CLARITY					
14. Our assessment directions for students are clear.	1	2	3	4	5
15. Our assessment directions for teachers are clear.	1	2	3	4	5
16. Our assessment directions for students are standardized across the district.	1	2	3	4	5
17. Our assessment directions for teachers are standardized across the district.	1	2	3	4	5
18. Our district sends individual reports to parents each school year.	1	2	3	4	5
19. Our district provides parents with reports that give an appropriate explanation of assessments results.	1	2	3	4	5
20. Our district provides individual student reports in appropriate language(s).	1	2	3	4	5
21. All district/school reports are appropriately disaggregated.	1	2	3	4	5
APPROPRIATENESS					
22. Our assessments are appropriate for the assessed grade level.	1	2	3	4	5
23. Our assessments demonstrate an increase of expectation from one grade level to the next.	1	2	3	4	5
24. Our assessments were screened for fairness, bias, and sensitivity.	1	2	3	4	5
25. Our assessments indicate our expectations for the students.	1	2	3	4	5
26. Our assessment plan provides for appropriate accommodations where necessary.	1	2	3	4	5
27. Our assessments have been administered with appropriate accommodations.	1	2	3	4	5
28. Our assessments were reviewed by internal or external groups.	1	2	3	4	5

**Nebraska-Led Portfolio Peer Review Survey
STARS Comprehensive Evaluation Project**

	None of the time	Very little of time	Some of the time	Most of the time	All of the time
Directions: Please circle the number that best describes your response to each statement.					
SCORING PROCEDURES					
29. Our performance level descriptors are clear and specific for each assessment.	1	2	3	4	5
30. Our performance level descriptors clearly differentiate for each proficiency level.	1	2	3	4	5
31. Our district consistently applies performance level descriptors to the cut scores for each assessment or standard.	1	2	3	4	5
32. Our performance level descriptors indicate increased expectations from one grade level to the next.	1	2	3	4	5
33. Our assessments have established scoring guidelines and directions.	1	2	3	4	5
34. Our subjectively scored assessments have clearly defined rubrics.	1	2	3	4	5
35. Our subjectively scored assessments have inter-rater reliability and decision consistency methods that are within acceptable ranges.	1	2	3	4	5
36. Our students are given instruction about behavioral objectives during the assessments.	1	2	3	4	5
37. Our district has taken test security measures to ensure results are not compromised.	1	2	3	4	5
38. Our district has monitoring procedures in place for inclusion, standardization, and security.	1	2	3	4	5
39. Our district provides training for those administering the assessments.	1	2	3	4	5
40. Our participation rates are documented.	1	2	3	4	5
41. Our district has local procedures in place for assuring appropriate accommodations for ELL students, students with disabilities, and students on 504 plans.	1	2	3	4	5
42. Our district has local assessment policies in place to assure comparability and consistency across the district.	1	2	3	4	5
SUMMARIZING THE REVIEW PROCESS					
43. I feel prepared to present my district portfolio to my peer reviewers.	1	2	3	4	5
44. I was provided time within the teaching day to prepare the district assessment portfolio.	1	2	3	4	5
45. Compensation is provided to prepare the district assessment portfolio when completed outside of the regular school day.	1	2	3	4	5
46. I have had adequate help in preparing the district assessment portfolio.	1	2	3	4	5
47. I have the necessary information to prepare the district assessment portfolio.	1	2	3	4	5
48. Please rate each area of the district portfolio review process according to your district's areas of strength and weakness with "1" representing the weakest rating and "5" the strongest rating					
a. Alignment	1	2	3	4	5
b. Sufficiency	1	2	3	4	5
c. Clarity	1	2	3	4	5
d. Appropriateness	1	2	3	4	5
e. Scoring Procedures	1	2	3	4	5
49. Please identify your district's portfolio review strengths and weakness by ranking the following components in priority order from "1" to "5," with "1" ranking weakest and "5" ranking strongest. Use each number only once.					
a. Alignment					
b. Sufficiency					
c. Clarity					
d. Appropriateness					
e. Scoring Procedures					

PLEASE COMPLETE AND RETURN BEFORE YOU COMPLETE YOUR FIRST NEBRASKA LED PORTFOLIO PEER REVIEW. Use the enclosed self-addressed postage-paid envelope or mail to:

Jody Isernhagen, Principal Investigator
STARS Comprehensive Evaluation
141 Teachers College Hall
PO Box 880360, Lincoln, NE 68588-0360

Appendix E
2006-2007 Study I: Nebraska-led Portfolio Peer
Review Interview Protocol

STARS PORTFOLIO INTERVIEW PROTOCOL
2006-2007

Qualitative Research Purpose: *Explore and understand the perceptions of educators about the STARS Portfolio Peer Review.*

Date of interview: _____ Time of interview: _____

Location of interview: _____

Interviewer: _____

Participant Profile

Participant: _____

District and School: _____

Position: ___ Superintendent ___ Assessment Coordinator ___ Principal @ HS MS ELEM
 ___ Teacher ELEM MS HS Subject: _____

OTHER: _____

Years at present position and site: _____ Total Years in Education: _____

Introduction:

1. Thank you for taking the time to visit with me today.
2. I am serving as an interviewer for the STARS Comprehensive Evaluation conducted by the University of Nebraska-Lincoln. This research is being conducted so that the Nebraska Department of Education has a better understanding of how the STARS portfolio process is being implemented in school districts and schools across the state. Information gained from this research is used to improve the process and to provide insight into next steps.
3. First, I want to assure you that this interview is strictly confidential. Information provided by school and district staff is reported or released in aggregated form only. Districts, schools and individuals are not identified.
4. I have an Informed Consent form outlining your rights as a research participant. You are free to decide not to participate in this study or to withdraw from the study at any time without adversely affecting your relationship with the investigators, the University of Nebraska-Lincoln, or the Nebraska Department of Education. Contact persons for the project and the Institutional Review Board are provided on the Informed Consent Form in case you have questions or concerns. I have a copy for you to sign and one for you to keep for your use.
5. I am going to record this interview so that the interview can be transcribed (a typed copy of the interview will be made) and we have an accurate rendering of your responses.
6. It is important that I maintain the integrity of your words and intentions; therefore, I may ask you to review the transcription if I have any difficulties with the interpretation.
7. We are interested in finding out about the perceptions that you hold regarding the STARS assessment portfolio peer review process and its implementation in your school or district. Questions about alignment, sufficiency, clarity, appropriateness, scoring procedures, and any other topics of interest that would help improve the process.
8. Please feel free to discuss your views openly. From time to time, I may have additional questions to further understand a concept that you have shared.
9. Let's begin. Please state your name, school, district and indicate permission to record this interview by repeating this statement, "I (your name) at (school/district name) willingly give my permission to record this interview."

Interview Questions

DIRECTIONS: Place a check when the participant mentions each probe so that you do not repeat the probe.

1. How have you been supported with the preparation of the assessment portfolio?

Probes

- _____ a. What type of training did you receive to prepare your assessment portfolio for the review and who provided the training?
- _____ b. Did the training you received prepare you for the actual peer review process?

Descriptive Notes:	Reflective Notes

2. When the peer review process began, what were your initial thoughts?

Probes

- _____ a. How did you know that your expertise was valued?
- _____ b. Share some of your thoughts regarding the first part of the review?

Descriptive Notes:	Reflective Notes

3. Tell us about how your assessments matched and measured your standards?

Probes

- _____ a. Tell us about the changes you will make to your alignment process due to the interaction with the peer reviewers.
- _____ b. What recommendations could you offer to other districts about matching and measuring the assessments to standards to improve the process (alignment)?

Descriptive Notes:	Reflective Notes

4. Share about the process of distributing assessment items across all performance levels (sufficiency)?

Probes

- _____ a. How did you assure that assessment performance descriptors were clear, differentiated, and increased expectations for each grade?
- _____ b. Tell us about any changes you may make in how you will distribute assessment items across all performance levels due to your interaction with the peer reviewers (sufficiency).
- _____ c. What recommendations could you offer to make the sufficiency process easier for other districts?

Descriptive Notes:	Reflective Notes

5. Talk about how your district provided directions that were clear and appropriate for all teachers and students (clarity)?

Probes

- _____ a. How do you report assessment results to students and parents?
- _____ b. How do you report to students and parents about special populations?
- _____ c. What recommendations could you offer to make assessment directions clearer for other districts administering assessments?

Descriptive Notes:	Reflective Notes

6. Share how your district ensured that the assessments were free of bias?

Probes

- _____ a. Tell us about any changes you may make in how you will ensure that assessments are from bias.
- _____ b. What recommendations could you offer to make the bias review process easier for other districts?

Descriptive Notes:	Reflective Notes

7. How did your district ensure that your assessments were appropriate for grade levels across the district?

Probes

- _____ a. How did you ensure that assessment expectations increased from one grade level to the next?
- _____ b. How did you plan for and administer the needed accommodations for students?
- _____ c. Tell us about the changes you will make to your assessments due to the interaction with the peer reviewers.
- _____ d. What recommendations could you offer to make assessments meet the standards of appropriateness easier for other districts?

Descriptive Notes:	Reflective Notes

8. Share how your district ensured that assessments were reliable and consistent.

Probes

- _____ a. Tell us about the methods you used for meeting reliability and consistency?
- _____ b. How will the review process help you improve your assessments?
- _____ c. Tell us about the changes you will make to your scoring procedures due to the interaction with the peer reviewers.
- _____ d. What recommendations could you offer to make the scoring procedures easier for other districts?

Descriptive Notes:	Reflective Notes

9. What new learnings have you had due to your involvement in the Nebraska-led Assessment Portfolio Peer Review?

Probes

- _____ a. You have stated (one, two or what ever has been stated) new learnings due to your involvement in the peer review process. Are there others?
- _____ b. What was the value of your new learnings to you as a professional and to your school or district?
- _____ c. Did the process meet your expectations?

Descriptive Notes:	Reflective Notes

10. Please share anything that you believe will strengthen the Nebraska-led Assessment Portfolio Peer Review.

Descriptive Notes:

Reflective Notes

Appendix F
2006-2007 Study II: STARS Leader
Research Survey

STARS Leader Survey (2006-07)

INSTRUCTIONS: Please provide the following demographic information by responding to the questions or marking the appropriate category for each area.

District Name: _____

Mark the primary position you hold in your school district:

Superintendent
 Assessment/Curriculum Coordinator
 Secondary Principal/Assistant Principal
 Middle School Principal/Assistant Principal
 Elementary Principal/Assistant Principal
 Other: Specify Job Title: _____

My Gender: Male Female

Total Years of Experience: _____ **Years of Experience in my Current Role:** _____

<i>Directions:</i> Please click on the number that best describes your rating for how STARS is implemented in your district.	None of time	Very little of	Some of time	Most of time	All of the time
District Support					
1. My district supports "school-based teacher-led" assessment.	1	2	3	4	5
2. My district encourages a culture of continuous school improvement.	1	2	3	4	5
3. My district provides services needed for the implementation of STARS.	1	2	3	4	5
4. My district provides on-going assessment training for STARS.	1	2	3	4	5
5. My district provides resources for learning teams as a means of professional development.	1	2	3	4	5
6. My district provides a record keeping system for recording STARS data.	1	2	3	4	5
7. My district has assisted teachers throughout the K-12 system for developing STARS assessments.	1	2	3	4	5
8. My district is committed to external communication regarding student achievement.	1	2	3	4	5
9. My district has adopted a guiding assessment philosophy, mission, and beliefs regarding STARS.	1	2	3	4	5
10. My district provides personnel policies that reflect an expectation of assessment competence.	1	2	3	4	5
11. My district provides policies at the district/school level that contribute to STARS assessment practices.	1	2	3	4	5
12. My district provides individual support to ensure understanding of the STARS process for new teachers	1	2	3	4	5
Assessment Literacy					
13. In my district/school assessments are aligned to the state/district standards.	1	2	3	4	5
14. In my district/school, teachers are committed to improving their own assessment competence.	1	2	3	4	5
15. In my district/school teachers write their own STARS assessments.	1	2	3	4	5
16. In my district/school teachers select assessment items from a common bank/pool.	1	2	3	4	5
17. In my district/school administrators support "school-based teacher-led" assessment.	1	2	3	4	5
18. In my district/school teachers support "school-based teacher-led" assessment.	1	2	3	4	5
19. In my district/school I participate in the development of the STARS assessment portfolio sent to NDE.	1	2	3	4	5
20. In my district/school teachers that work with students of special populations are involved in STARS.	1	2	3	4	5
21. In my district/school teachers participate in learning teams to collaborate about STARS.	1	2	3	4	5
22. I participate in a learning team within my district or school.	1	2	3	4	5
23. In my school/district, new teachers are provided time, training, and resources to develop assessment literacy	1	2	3	4	5
Data					
24. In my district/school I use the STARS data results to make instructional decisions.	1	2	3	4	5
25. In my district/school I disaggregate the STARS data.	1	2	3	4	5
26. In my district/school I use STARS disaggregated data to make instructional decisions about students in special populations.	1	2	3	4	5
27. Schools in my district provide a record keeping system for STARS data.	1	2	3	4	5
28. My district/school provides assistance for analyzing STARS assessment data.	1	2	3	4	5
29. In my school/district, new teachers are trained to use data for improving instruction	1	2	3	4	5

STARS Leader Survey (2006-07)

Leadership					
30. In my district/school I am committed to academic excellence that is communicated through a clear vision.	1	2	3	4	5
31. In my district/school my vision defines how STARS assessment fits into effective teaching and learning.	1	2	3	4	5
32. In my district/school I define success as high achievement for all learners.	1	2	3	4	5
33. In my district/school I am assessment literate and committed to assessment literacy for all.	1	2	3	4	5
34. In my district/school I provide a schedule or plan that coordinates testing of students.	1	2	3	4	5
35. In my school/district, I am committed to encouraging new teachers to become school leaders.	1	2	3	4	5
Instructional Impact					
36. In my district/school curriculum is aligned to the state standards.	1	2	3	4	5
37. In my district/school a curriculum implementation plan describes roles and responsibilities and ensures standards-based instruction.	1	2	3	4	5
38. In my district/school consistency in achievement expectations for all educators is ensured.	1	2	3	4	5
39. In my district/school teachers are held accountable for teaching the adopted curriculum.	1	2	3	4	5
40. In my district/school new teachers are given specialized assistance in understanding the STARS process.	1	2	3	4	5
41. In my district/school achievement standards are held high for all students .	1	2	3	4	5
42. In my district/school teachers understand and apply the principles of sound grading practices.	1	2	3	4	5
43. In my district/school assessment is woven into instruction.	1	2	3	4	5
44. Students in my district/school are involved in understanding their own progress and achievement status.	1	2	3	4	5
45. In my district/school instruction is modified when students do not perform well on the STARS assessments.	1	2	3	4	5
46. In my district/school lesson plans are aligned to STARS assessment results.	1	2	3	4	5
47. In my district/school instructional differentiation is provided to meet individual needs of students.	1	2	3	4	5
48. I use STARS assessment results to question, modify and adjust my district/school instructional decisions.	1	2	3	4	5
49. In my district/school I have evidence that lesson plans are aligned to standards.	1	2	3	4	5
50. In my district/school I have evidence that standards are taught.	1	2	3	4	5
51. In my district/school I have evidence that all students have an opportunity to learn.	1	2	3	4	5
52. In my district/school students that do not master the standards are re-taught.	1	2	3	4	5
53. In my district/school I have visited other school sites to examine successful instructional interventions.	1	2	3	4	5
54. In my district/school administrators and teachers collectively use data to develop support systems for students not meeting the standards.	1	2	3	4	5
55. In my district/school leaders are focused on instruction and continuous improvement.	1	2	3	4	5
56. In my district/school opportunities are provided for teachers to share STARS assessment results with parents.	1	2	3	4	5
57. In my district/school teachers use STARS data to develop interventions.	1	2	3	4	5
58. In my district/school administrators assist teachers in making instructional decisions based on STARS data.	1	2	3	4	5
59. In my school/district, new teachers are involved in curriculum review so they better understand how curriculum, assessment, and school improvement are aligned.	1	2	3	4	5
External Support					
60. I participate in professional development for STARS provided by external experts.	1	2	3	4	5
61. I have attended the Nebraska Assessment Conference at least once.	1	2	3	4	5
62. The ESU provides on-going assessment training for me.	1	2	3	4	5
63. The ESU conducts data retreats to help me understand STARS assessment results.	1	2	3	4	5
64. The ESU assists my district/school in the identification of appropriate interventions.	1	2	3	4	5
65. The ESU provides training in instructional strategies that produce increased student achievement.	1	2	3	4	5
66. The ESU supported my district/school in preparing evidence for the Nebraska-led peer review of the STARS.	1	2	3	4	5
67. I encourage new teachers in my school/district in increase their knowledge of STARS by participating in professional development at the ESU.	1	2	3	4	5

Appendix G
2006-2007 Study II: STARS Teacher
Survey

STARS Teacher Survey 2006-07

Leadership					
32. In my school I am committed to academic excellence that is communicated through a clear vision.	1	2	3	4	5
33. In my school my vision defines how STARS assessment fits into effective teaching and learning.	1	2	3	4	5
34. In my school I define success as high achievement for all learners.	1	2	3	4	5
35. In my school I am assessment literate and committed to assessment literacy for all.	1	2	3	4	5
36. In my school I provide an assessment schedule that coordinates testing for my students.	1	2	3	4	5
37. In my school/district, I am committed to encouraging new teachers to become school leaders.	1	2	3	4	5
Instructional Impact					
38. My curriculum is aligned to the state standards.	1	2	3	4	5
39. In my school a curriculum implementation plan describes roles and responsibilities and ensures standards-based instruction.	1	2	3	4	5
40. In my school consistency in achievement expectations for all educators is ensured.	1	2	3	4	5
41. In my school I am held accountable for teaching the adopted curriculum.	1	2	3	4	5
42. In my school new teachers are given specialized assistance in understanding the STARS process.	1	2	3	4	5
43. In my school achievement standards are held high for all students .	1	2	3	4	5
44. In my school I understand and apply the principles of sound grading practices.	1	2	3	4	5
45. In my school I am responsible for weaving assessment into instruction.	1	2	3	4	5
46. Students in my school are involved in understanding their own progress and achievement status.	1	2	3	4	5
47. I modify my instructional strategies when students do not perform well on the STARS assessments.	1	2	3	4	5
48. I aligned my lesson planning to STARS assessment data results.	1	2	3	4	5
49. I use instructional differentiation to provide for the individual needs of students.	1	2	3	4	5
50. I use STARS assessment results to question, modify and adjust my own teaching.	1	2	3	4	5
51. I align my lesson planning to the standards.	1	2	3	4	5
52. I record the standards that I teach to ensure that all students have an opportunity to learn.	1	2	3	4	5
53. When students do not master a standard, I reteach it.	1	2	3	4	5
54. I have visited other school sites to examine successful instructional interventions.	1	2	3	4	5
55. Administrators and teachers collectively use data to develop support systems for students not meeting the standards.	1	2	3	4	5
56. Administrators in my school assist teachers in making instructional decisions based on STARS data.	1	2	3	4	5
57. Teachers in my school are focused on instruction and continuous improvement.	1	2	3	4	5
58. I share STARS assessment results with parents of students in my classroom.	1	2	3	4	5
59. Teachers in my school use STARS data to develop interventions.	1	2	3	4	5
60. My administrator and I focus upon student achievement results with standards in teacher evaluation conferences.	1	2	3	4	5
61. In my school/district, new teachers are involved in curriculum review so they better understand how curriculum, assessment, and school improvement are aligned.	1	2	3	4	5
External Support:					
62. I participate in professional development for STARS provided by external experts.	1	2	3	4	5
63. I have attended the Nebraska Assessment Conference at least once.	1	2	3	4	5
64. The ESU provides on-going assessment training for me.	1	2	3	4	5
65. The ESU conducts data retreats to help me understand STARS assessment results.	1	2	3	4	5
66. The ESU assists my school in the identification of appropriate interventions.	1	2	3	4	5
67. The ESU provides training in instructional strategies that produce increased student achievement.	1	2	3	4	5
68. The ESU supported my school in preparing evidence for the Nebraska-led peer review of the STARS.	1	2	3	4	5
69. I encourage new teachers in my school/district to increase their knowledge of STARS by participating in professional development at the ESU	1	2	3	4	5

Appendix H
2006-2007 Study II: STARS Research Interview
Protocol-Leaders

STARS RESEARCH INTERVIEW PROTOCOL
LEADERS
2006-2007

Qualitative Research Purpose: Explore and understand the similarities and differences in perceptions of the STARS process and its impact.

Date of interview: _____ Time of interview: _____

Location of interview: _____

Interviewer: _____

Participant Profile

Participant: _____

School District: _____ ESU # _____

Position: Superintendent _____ Principal _____ Other _____

Years at present position/site: _____ Total Years in Education: _____ Gender: M F

Introduction:

1. Thank you for taking the time to visit with me today.
2. I am serving as an interviewer for the STARS Comprehensive Evaluation conducted by the University of Nebraska-Lincoln (UNL). This research is being conducted so that the Nebraska Department of Education (NDE) has a better understanding of how the STARS process is being implemented in school districts and schools across the state. Information gained from this research is used to improve the process and to provide insight into next steps.
3. First, I want to assure you that this interview is strictly confidential. Information provided by participants is reported or released in aggregated form only. Districts, schools, individuals, and ESU's are not identified. Are you participating in this interview willingly?
(Interviewer only: If the participant indicates they are not a willing participant, thank them for coming in to speak with you and dismiss them with the assurance their declination will not affect their relationship with their school, district, UNL, or the NDE.)
4. I have an Informed Consent Form outlining your rights as a participant. You are free to decide not to participate in this study or to withdraw from the study at any time without adversely affecting your relationship with the investigators, the University of Nebraska-Lincoln, the Nebraska Department of Education, or your school district. Contact persons for the project and the Institutional Review Board are provided on the Informed Consent Form in case you have questions or concerns. I have a copy for you to sign and one for you to keep for your use.
5. I am going to record this interview so that the interview can be transcribed (a typed copy of the interview will be made) and we have an accurate rendering of your responses.
6. It is important that I maintain the integrity of your words and intentions; therefore, I may ask you to review the transcription if I have any difficulties with the interpretation.
7. We are interested in finding out about the perceptions that you hold regarding the STARS assessment process and its implementation in your district. Questions about your leadership in the assessment process and its impact upon instruction and curriculum are specific topics of interest.
8. Please feel free to discuss your views openly. From time to time, I may have additional questions to further understand a concept that you have shared.
9. Let's begin. Please state your name, the name of your building/district and indicate your willingness to provide permission to record this interview by repeating this statement, "I (your name) at (school/district name) willingly give my permission to record this interview."

Interview Questions

DIRECTIONS: Place a check when the participant mentions each probe so that you do not repeat the probe.

1. How does your school/district help new teachers and/or teachers new to the STARS process learn how to use the STARS six quality criteria for assessment?

Six Quality Criteria:

1. match and measure the standards
2. provide opportunity for students to have learned the content
3. be free of bias
4. be written at the appropriate level
5. be reliable and consistently scored
6. have appropriate mastery levels

Probes

- _____ What has been your experiences with the six quality criteria?
- _____ What creative ways has your school/district used to provide the needed time for STARS?
- _____ How do you develop your assessment literacy skills (training, conversations, learning teams)?

Descriptive Notes	Reflective Notes

2. How do you involve staff in the leadership of the STARS process in your school/district?

Probes

- _____ How are new teachers participating in the leadership of the STARS process?
- _____ Does your school mission and beliefs support high achievement for all learners?
- _____ What is your role and responsibility in the planning for the administration of STARS assessments?
- _____ What policies/plans are in place to support STARS assessment? (Ex: vision, philosophy of assessment, high achievement for all, assessment literacy, testing procedures)

Descriptive Notes	Reflective Notes

3. How are you increasing teacher competence to maximize student learning in your school/district?

Probes:

- _____ What evidence do you collect to ensure teacher competence?
- _____ How have you maximized the use of time for instruction in classrooms?
- _____ How has your school/district helped teachers develop clear and appropriate instructional targets based on STARS assessment results?
- _____ How does your school/district choose interventions for increasing student learning?
- _____ How does your school/district choose interventions for increasing student learning for those specific students who do not meet the standards?

Descriptive Notes	Reflective Notes

4. How have you assisted teachers in integrating STARS assessment results into instruction to inform teaching and learning?

Probes

- _____ What are some of the ways you collect evidence that STARS assessment results inform teaching and learning?
- _____ How do you collect evidence that teachers are modifying their instruction in cases where students did not perform well on a STARS assessment?
- _____ How do you collect evidence that rubrics are used in assessment and instruction?
- _____ How has STARS impacted the grading practices for your school/district?

Descriptive Notes	Reflective Notes

5. What changes have you made within your school/district to ensure all students experience academic growth?

Probes

- _____ How do you track student success with standards?
 _____ What do you do for students who are not successful?
 _____ Can you share some examples of the challenges you face with student academic growth?

Descriptive Notes	Reflective Notes

6. How do you incorporate the STARS data in your school improvement process?

Probes

- _____ What types of data do you use in your school improvement process?
 _____ In what ways do you participate in conversations with others about the improvement of instruction? (Professional Learning Communities, Content Area Teams, etc.)
 _____ How are you involved in your school improvement process in your building? In your district?

Descriptive Notes	Reflective Notes

7. How has STARS impacted classroom instruction for special populations (Special Ed, ELL, Title I, Gender)?

Probes

- _____ What are you doing to increase student achievement for these special populations?
 - _____ Examples of programs
 - _____ Examples of strategies/interventions/accommodations
 - _____ Examples of tools

Descriptive Notes	Reflective Notes

8. How are you assisting students of mobility to master the STARS standards and perform successfully on STARS assessments?

Probes

- _____ How do you assess students of mobility when they arrive?
- _____ What is your data telling you about students of mobility?
- _____ What kind of STARS assessment records do you send to other schools for students of mobility?

Descriptive Notes	Reflective Notes

9. How have the ways in which you monitor and report student performance to students, parents, and community changed since the inceptions of STARS?

Probes

- _____ How are you involving students, parents, and community in improving teaching and learning (meetings, open houses, media, PTA/PTO)?
- _____ How do you communicate your district/school STARS assessment results to your stakeholders?

Descriptive Notes	Reflective Notes

10. What comments, recommendations, or final observations would you like to make about STARS that we have not discussed?

Descriptive Notes	Reflective Notes

Appendix I
2006-2007 Study II: STARS Research Interview
Protocol-Teachers

**STARS RESEARCH INTERVIEW PROTOCOL
TEACHERS
2006-2007**

Qualitative Research Purpose: Explore and understand the similarities and differences in perceptions of the STARS process and its impact.

Date of interview: _____ Time of interview: _____

Location of interview: _____

Interviewer: _____

Participant Profile

Participant: _____

School District: _____ ESU # _____

Position: Grade Level _____ Subject Area _____

Years at present position/site: _____ Total Years in Education: _____ Gender: M F

Introduction:

1. Thank you for taking the time to visit with me today.
2. I am serving as an interviewer for the STARS Comprehensive Evaluation conducted by the University of Nebraska-Lincoln (UNL). This research is being conducted so that the Nebraska Department of Education (NDE) has a better understanding of how the STARS process is being implemented in school districts and schools across the state. Information gained from this research is used to improve the process and to provide insight into next steps.
3. First, I want to assure you that this interview is strictly confidential. Information provided by participants is reported or released in aggregated form only. Districts, schools, individuals, and ESU's are not identified. Are you participating in this interview willingly?
(Interviewer only: If the participant indicates they are not a willing participant, thank them for coming in to speak with you and dismiss them with the assurance their declination will not affect their relationship with their school, district, UNL, or the NDE.)
4. I have an Informed Consent Form outlining your rights as a participant. You are free to decide not to participate in this study or to withdraw from the study at any time without adversely affecting your relationship with the investigators, the University of Nebraska-Lincoln, or the Nebraska Department of Education, or your school district. Contact persons for the project and the Institutional Review Board are provided on the Informed Consent Form in case you have questions or concerns. I have a copy for you to sign and one for you to keep for your use.
5. I am going to record this interview so that the interview can be transcribed (a typed copy of the interview will be made) and we have an accurate rendering of your responses.
6. It is important that I maintain the integrity of your words and intentions; therefore, I may ask you to review the transcription if I have any difficulties with the interpretation.
7. We are interested in finding out about the perceptions that you hold regarding the STARS assessment process and its implementation in your district. Questions about your leadership in the assessment process and its impact upon instruction and curriculum are specific topics of interest.
8. Please feel free to discuss your views openly. From time to time, I may have additional questions to further understand a concept that you have shared.
9. Let's begin. Please state your name, the name of your building/district and indicate your willingness to provide permission to record this interview by repeating this statement, "I (your name) at (school/district name) willingly give my permission to record this interview."

Interview Questions

DIRECTIONS: Place a check when the participant mentions each probe so that you do not repeat the probe.

1. How does your school/district help new teachers and/or teachers new to the STARS process learn how to use the STARS six quality criteria for assessment?

Six Quality Criteria:

1. match and measure the standards
2. provide opportunity for students to have learned the content
3. be free of bias
4. be written at the appropriate level
5. be reliable and consistently scored
6. have appropriate mastery levels

Probes

- _____ What has been your experiences with the six quality criteria?
- _____ What creative ways has your school/district used to provide the needed time for STARS?
- _____ How do you develop your assessment literacy skills (training, conversations, learning teams)?

Descriptive Notes	Reflective Notes

2. How do you participate in the leadership of the STARS process in your school?

Probes

- _____ How are new teachers participating in the leadership of the STARS process?
- _____ Does your school mission and beliefs support high achievement for all learners?
- _____ What is your role and responsibility in the planning for the administration of STARS assessments?
- _____ What policies/plans are in place to support STARS assessment? (Ex: vision, philosophy of assessment, high achievement for all, assessment literacy, testing procedures)

Descriptive Notes	Reflective Notes

3. How are you increasing your own teaching competence to maximize student learning in your classroom?

Probes

- _____ How have you maximized the use of time for instruction in your classroom?
- _____ How has your school/district helped you develop clear and appropriate instructional targets based on STARS assessment results?
- _____ How are you involved in choosing interventions for increasing student learning?
- _____ How are you involved in choosing interventions for increasing student learning for those specific students who do not meet the standards?

Descriptive Notes	Reflective Notes

4. How have you integrated STARS assessment results into instruction to inform teaching and learning?

Probes

- _____ What are some ways that your STARS assessment results inform your teaching and learning?
- _____ How do you modify your instruction in cases where students did not perform well on a STARS assessment?
- _____ How do you use rubrics in assessment and instruction?
- _____ How has STARS impacted your grading practices?

Descriptive Notes	Reflective Notes

5. What changes have you made within your classroom to ensure all students experience academic growth?

Probes

- _____ How do you track student success with standards?
 _____ What do you do for students who are not successful?
 _____ Can you share some examples of the challenges you face with student academic growth?

Descriptive Notes	Reflective Notes

6. How do you incorporate the STARS data in your school improvement process?

Probes

- _____ What types of data do you use in your school improvement process?
 _____ In what ways do you participate in conversations with others about the improvement of instruction? (Professional Learning Communities, Content Area Teams, etc.)
 _____ How are you involved in your school improvement process in your building? In your district?

Descriptive Notes	Reflective Notes

7. How has STARS impacted classroom instruction for special populations (Special Ed, ELL, Title I, Gender)?

Probes

- _____ What are you doing to increase student achievement for these special populations?
 _____ Examples of programs
 _____ Examples of strategies/interventions/accommodations
 _____ Examples of tools

Descriptive Notes	Reflective Notes

8. How are you assisting students of mobility to master the STARS standards and perform successfully on STARS assessments?

Probes

- _____ How do you assess students of mobility when they arrive?
 _____ What is your data telling you about students of mobility?
 _____ What kind of STARS assessment records do you send to other schools for students of mobility?

Descriptive Notes	Reflective Notes

9. How have the ways in which you monitor and report student performance to students, parents, and community changed since the inceptions of STARS?

Probes

- _____ How are you involving students, parents and community in improving teaching and learning (meetings, open houses, media, PTA/PTO)?
- _____ How do you communicate your STARS assessment results to your stakeholders?

Descriptive Notes	Reflective Notes

10. What comments, recommendations, or final observations would you like to make about STARS that we have not discussed?

Descriptive Notes	Reflective Notes

Appendix J
2006-2007 Study VI: Use of Data to Inform Decisions by
Elementary Building Principals Survey

School Improvement Practices in Nebraska Elementary Schools

Directions: Select the category that best describes the practices in your building.

1. School improvement goals are generated exclusively from my own personal experience and/or opinions.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

2. As the principal, I analyze student achievement on my own and then inform the faculty of the results and needs I have identified.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

3. As the principal, I work with a representative group of faculty members to analyze student achievement data to determine school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

4. As the principal, I work with the entire faculty to analyze student achievement data to determine school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

5. I allocate time for all teachers to work collaboratively to analyze student achievement data.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

6. In working with staff members to analyze data, the process
 - a. Was facilitated entirely by me.
 - b. Was facilitated with assistance from others in the building.
 - c. Was facilitated with assistance from others in the district.
 - d. Was facilitated by someone in the district other than myself.
 - e. Did not take place with staff in my building.

7. In analyzing data, my Educational Service Unit (ESU) provided (select all that apply)
 - a. Training about data analysis.
 - b. Consultation with ESU staff about my data.
 - c. On-site consultation working with my school's staff members.
 - d. A data retreat for all schools in the area.
 - e. Detailed data reports about my school.
 - f. No support for data analysis.

8. I feel capable of analyzing data
 - a. Yes
 - b. No

9. My answer to question #8 was based on (select all that apply):
 - a. Coursework that I have completed
 - b. Training that I have received
 - c. Support that I receive at my district
 - d. Support that I receive at my ESU
 - e. Other, please specify

10. As the principal, I alone decide the focus of schoolwide improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

11. As the principal, I consult with selected teachers and/or community members before I make decisions about schoolwide improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

12. As the principal, I work with a representative teacher group to make decisions about the focus of school-wide improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

13. As the principal, I work with the whole faculty to make decisions about the focus of school-wide improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

14. Parents and/or community members were involved in analyzing data to determine school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

15. In working with staff to identify school improvement goals, the process
 - a. Was facilitated entirely by me.
 - b. Was facilitated with assistance from others in the building.
 - c. Was facilitated with assistance from others in the district.
 - d. Was facilitated by someone in the district other than myself.
 - e. Did not involve other staff in my building.

16. To identify school improvement goals, my Educational Service Unit (ESU) provided (select all that apply)
 - a. Training about writing goals.
 - b. Consultation with ESU staff about my goals.
 - c. On-site consultation working with my school's staff members.
 - d. A data retreat for all schools in my area
 - e. Detailed examples of school improvement goals.
 - f. No support for writing goals.

17. I feel capable of identifying school improvement goals
 - a. Yes
 - b. No

18. My answer to question #17 is based on (select all that apply):
 - a. Coursework that I have completed
 - b. Training that I have received
 - c. Support that I receive at my district
 - d. Support that I receive at my ESU
 - e. Other, please specify

19. A variety of student achievement results are used to identify school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

20. Criterion-referenced and norm-referenced test results are used to identify school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

21. Student work samples, student portfolios, and/or other examples of locally developed assessments are used to identify school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

22. Other types of data, such as discipline referrals, grade retention, and/or enrollment in advanced courses are used to identify school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

23. Data that has been disaggregated by race, gender, free and reduced lunch status, and/or special needs are used to identify school improvement goals.
 - a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

24. In working with data for my school, the data
 - a. Was compiled, disaggregated, and organized entirely by me.
 - b. Was compiled, disaggregated, and organized with assistance from others in the building.
 - c. Was compiled, disaggregated, and organized with assistance from others in the district.
 - d. Was compiled, disaggregated, and organized by others in the district.
 - e. Was not compiled, disaggregated, and organized for my school.

25. In using disaggregated data, my Educational Service Unit (ESU) provided (select all that apply)
 - a. Training about data disaggregation.
 - b. Consultation with ESU staff about my disaggregated data.
 - c. On-site consultation working with my school's staff members.
 - d. A data retreat for all schools in my area.
 - e. Detailed data reports about my school.
 - f. No support for data disaggregation.

26. I feel capable of working with disaggregated data.
 - a. Yes
 - b. No

27. My answer to question #26 is based on (select all that apply):
 - a. Coursework that I have completed
 - b. Training that I have received
 - c. Support that I receive at my district
 - d. Support that I receive at my ESU
 - e. Other, please specify

28. As the principal, I have some control over the types of professional development activities provided for the teachers in my building.

- a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion
29. As the principal, I determine the professional development activities for the teachers in my building.
- a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion
30. The whole school faculty utilized disaggregated student learning results to identify professional development needs.
- a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion
31. As the principal, I collect data informally from staff members to decide the types of professional development offered at the building.
- a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion
32. The school's professional development program is tied to the school improvement goal.
- a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion
33. As the principal, I use data from classroom observations and/or walk-throughs to determine the design of the professional development program.
- a. Strongly agree
 - b. Agree
 - c. Disagree
 - d. Strongly disagree
 - e. No opinion

34. As the principal, I collect and use teacher surveys to determine the design of the professional development program.
- Strongly agree
 - Agree
 - Disagree
 - Strongly disagree
 - No opinion
35. The whole faculty uses data for evaluation of professional development.
- Strongly agree
 - Agree
 - Disagree
 - Strongly disagree
 - No opinion
36. When identifying professional development activities for my school, the professional development
- Was planned entirely by me.
 - Was planned with assistance from others in the building.
 - Was planned with assistance from others in the district.
 - Was planned by someone in the district other than me.
 - Does not take place at the building level.
37. In designing professional development activities for my school, my Educational Service Unit (ESU) provided (select all that apply)
- Training about effective professional development.
 - Consultation with ESU staff about my professional development.
 - On-site consultation working with my school's staff members.
 - Trainings held at the ESU for all schools in my area.
 - No support for professional development.
 - Other, please specify
38. I feel capable of planning professional development.
- Yes
 - No
39. My answer to question #38 is based on (select all that apply):
- Coursework that I have completed
 - Training that I have received
 - Support that I receive at my district
 - Support that I receive at my ESU
 - Other, please specify
40. As the principal, I collect student achievement data related to the school improvement goal.
- Frequently (Once or more per month)
 - Occasionally (Once or more per quarter)
 - Sometimes (Once or more per semester)
 - Infrequently (Once or more per year)
 - Never (Not at all)

41. As the principal, I collect data from teachers related to the school improvement goal.
 - a. Frequently (Once or more per month)
 - b. Occasionally (Once or more per quarter)
 - c. Sometimes (Once or more per semester)
 - d. Infrequently (Once or more per year)
 - e. Never (Not at all)

42. As the principal, I am required to submit reports about student achievement to the district.
 - a. Frequently (Once or more per month)
 - b. Occasionally (Once or more per quarter)
 - c. Sometimes (Once or more per semester)
 - d. Infrequently (Once or more per year)
 - e. Never (Not at all)

43. As the principal, I am required to submit reports about student achievement to the state.
 - a. Frequently (Once or more per month)
 - b. Occasionally (Once or more per quarter)
 - c. Sometimes (Once or more per semester)
 - d. Infrequently (Once or more per year)
 - e. Never (Not at all)

44. As the principal, I provide reports about student achievement to parents and the community.
 - a. Frequently (Once or more per month)
 - b. Occasionally (Once or more per quarter)
 - c. Sometimes (Once or more per semester)
 - d. Infrequently (Once or more per year)
 - e. Never (Not at all)

45. Baseline data is used to monitor student achievement related to the school improvement goal.
 - a. Frequently (Once or more per month)
 - b. Occasionally (Once or more per quarter)
 - c. Sometimes (Once or more per semester)
 - d. Infrequently (Once or more per year)
 - e. Never (Not at all)

46. Celebrations of improvements and accomplishments, based on data, are held in my school.
 - a. Frequently (Once or more per month)
 - b. Occasionally (Once or more per quarter)
 - c. Sometimes (Once or more per semester)
 - d. Infrequently (Once or more per year)
 - e. Never (Not at all)

47. When collecting and reporting student achievement results, the process
 - a. Was facilitated and implemented entirely by me.
 - b. Was facilitated and implemented with assistance from others in the building.
 - c. Was facilitated and implemented with assistance from others in the district.
 - d. Was facilitated and implemented by someone in the district other than me.

48. In collecting and reporting student achievement results for my school, my Educational Service Unit (ESU) provided (select all that apply)
- a. Training about collecting and reporting student achievement results.
 - b. Consultation with ESU staff about collecting and reporting student achievement results.
 - c. On-site consultation working with collecting and reporting student achievement results.
 - d. No support for collecting and reporting student achievement results.
49. I feel capable of collecting and reporting student achievement results.
- a. Yes
 - b. No
50. My answer to question #49 is based on (select all that apply):
- a. Coursework that I have completed
 - b. Training that I have received
 - c. Support that I receive at my district
 - d. Support that I receive at my ESU
 - e. Other _____
51. My gender is:
- a. Male
 - b. Female
52. The highest level of education that I have completed is
- a. Bachelors Degree
 - b. Masters Degree
 - c. Specialist Degree
 - d. Doctoral Degree
53. I am a member of the following professional organizations (select all that apply)
- a. National Association for Elementary School Principals
 - b. Nebraska Council of School Administrators
 - c. Association for Supervision and Curriculum Development
 - d. Nebraska Association for Supervision and Curriculum Development
 - e. National Staff Development Council
 - f. Other (please specify) _____
54. How many students are enrolled in your building?
- a. 1-100
 - b. 101-200
 - c. 201-300
 - d. 301-400
 - e. 401-500
 - f. 501-600
 - g. 601-700
 - h. 701-800
 - i. 801-900
 - j. 901-1000
 - k. more than 1000

55. How many students are enrolled in your district? (Select only one)
- a. 1-500
 - b. 501-1000
 - c. 1001-2000
 - d. 2001-3000
 - e. 3000-4000
 - f. 4001-5000
 - g. 5001-15,000
 - h. more than 15,000
56. What grade levels, not including preschool, are included in your building? _____
57. How many years, including the current year, have you served in your current job?
- a. 1-5
 - b. 6-10
 - c. 11-15
 - d. 16-20
 - e. 21-25
 - f. 26+
58. How many total years, including this year, have you served as an administrator?
- a. 1-5
 - b. 6-10
 - c. 11-15
 - d. 16-20
 - e. 21-25
 - f. 26+
55. What was the topic of the most recent professional development in your school?
- _____
56. Was the most recent professional development for your staff, selected based on district needs or selected based on school needs? (Select only one)
- a. District-based
 - b. Building-based

Appendix K
2006-2007 Study VII: Role of Technology in Supporting
STARS Survey

Role of Technology in Supporting Nebraska STARS

Charting STARS: Voices from the Field, Year-Four Report, an evaluation of the Nebraska School-based Teacher-led Assessment and Reporting System (STARS) summarized challenges and obstacles to the successful implementation of STARS. It is clear from responses across teachers, principals, superintendents, and ESUs that the time the process takes, the difficulties in the collection, reporting and management of data, and assessment literacy are key issues impacting the implementation of STARS.

Technology-based systems are often used to address issues of time, accuracy in collection and management of large amounts of data, and clarity in reporting and sharing information. These systems can also be designed to support participants in learning and using the systems properly. This survey is designed to assess the extent of technology use by schools as they participate in the STARS process; identify critical issues in using technology-supported planning, assessment and reporting systems; and provide direction for the development of policies and resources that will support more efficient and effective implementation of STARS.

Demographic Information

1. How many students are enrolled in your district?
 - a. 1 – 599
 - b. 600 – 999
 - c. 1000 – 1999
 - d. 2000 – 4999
 - e. over 5000

2. What is your job title?
 - a. Superintendent
 - b. Principal
 - c. Assessment Coordinator
 - d. Teacher
 - e. Other _____

3. What responsibility(ies) do you have in the assessment process? Select all that apply.
(IRB NOTE: The website survey will request a YES or NO answer for Items 3a. through 3g.)
 - a. Assessment coordination **(Yes/No)**
 - b. Development of assessment plans and quality indicators **(Yes/No)**
 - c. Administration of assessments **(Yes/No)**
 - d. Data aggregation **(Yes/No)**
 - e. Supervision of data management **(Yes/No)**
 - f. Reporting summary data internally **(Yes/No)**
 - g. Reporting summary data to NDE **(Yes/No)**

4. How do you participate in the STARS process?
 - a. As an individual school district

- b. As part of a consortium
 - c. As part of a collaboration
5. Does the ESU play a role in the assessment process in your district? (Yes/No)
If yes, what role(s) does the ESU play in your participation in STARS? Choose all that apply.
- a. I am completing the survey as a member of an ESU
 - b. The ESU assists in the development of the assessment process
 - c. The ESU participates in the administration of the assessments
 - d. The ESU aggregates data
 - e. The ESU manages the storage of data
 - f. The ESU reports data to the district
 - g. The ESU reports data to the NDE on behalf of the district
 - h. The ESU provides training for staff, teachers and/or administrators
6. Which of the following best describes your district's ability to invest in technology specifically to support the assessment process? Choose one.
- a. No funds are available for this purpose
 - b. Funds are available if the technology reduces other costs of administration
 - c. New funds are available if it leads to school improvement
 - d. Funds are available to develop new technology
7. Did the ARMS Grants issued through the Nebraska Department of Education directly impact the district's ability to invest in assessment technology?
- 1. Yes.
 - 2. No.
- If **Yes**, specifically how were those funds invested? Please check all that apply.
- a. Assessment administration
 - b. Management
 - c. Scoring
 - d. Student systems
 - e. Reporting
 - f. Training
 - g. Other

Current Uses of Technology

The next series of questions address the extent to which technology is used in your district or ESU to directly support the assessment process. The assessment process includes the collection, management, and reporting of assessment results. It also includes the planning and validation of the assessment system used in the district.

Please rate the extent to which you use each of the following types of technology to complete each of the listed assessment collection, management, and reporting tasks on a scale of 1 to 5 with 1 representing "to no extent" and 5 representing "to a great extent."

IRB NOTE: *The website survey will provide respondents with a five-point scale selection field for each item as shown in Item 1. For Items 1 through 5, the scale range is from “to no extent” to “to a great extent.” The scale range for Item 6 is “low” to “high.”)*

1. Developing and documenting assessment procedures

a. Word Processing	1	2	3	4	5
b. Spreadsheet	1	2	3	4	5
c. Dedicated software	1	2	3	4	5
d. Dedicated Web application	1	2	3	4	5
e. Local custom software/Web	1	2	3	4	5
f. Other	1	2	3	4	5

2. Administering assessments to students
 - a. Scanning
 - b. Computer-based assessment/Web-based assessment
 - c. Electronic samples/portfolio
 - d. Other _____
(Please identify)

3. Aggregation, management, and storage of assessment data
 - a. Word Processor
 - b. Spreadsheet
 - c. Database
 - d. Dedicated software
 - e. Dedicated Web application
 - f. Local custom software/Web

4. Interpreting and reporting assessment results
 - a. Word Processor
 - b. Spreadsheet
 - c. Database
 - d. Dedicated software
 - e. Dedicated Web application
 - f. Local custom software/Web

5. Available technology-based assessment systems are important to the implementation of the Nebraska Assessment Model.

6. Please provide your rating of the quality of the technology your district currently uses in the STARS assessment process in each of the following areas on a scale of 1 to 5, with 1 representing “Low” and 5 representing “High.”
 - a. Does the technology support the school improvement process?
 - b. Does the technology help in meeting NDE reporting requirements?
 - c. Does the technology address security and privacy issues
 - d. Is the technology readily available for classroom teachers?

- e. Is the technology readily available for administrators?
- f. Is the technology feasible to implement?

Impact of Technology on Assessment

The following questions assess your perceptions of the impact of technology on the assessment process.

1. Please indicate the extent to which you believe currently available technology contributes to each of the following on a scale of 1 to 5 with 1 representing “to no extent” and 5 representing “to a great extent.”
(IRB NOTE: The web survey will provide a five-point field as shown in Current Use of Technology Item 1.)
 - a. Consistent data collection
 - b. Knowledge of how to organize and summarize data
 - c. Results are accessible to a wide audience including classroom teachers
 - d. Coordination with planning and school improvement documentation and reporting
 - e. Coordinated with other data management system
 - f. Reporting options allow multiple views and disaggregation of data
 - g. Reporting targets
2. The use of technology significantly reduces the amount of time involved in STARS.
 - a. Strongly Disagree
 - b. Disagree
 - c. Neutral
 - d. Agree
 - e. Strongly Agree
3. The use of technology significantly increases the complexity of the STARS process.
 - a. Strongly Disagree
 - b. Disagree
 - c. Neutral
 - d. Agree
 - e. Strongly Agree
4. The use of technology increases the value of the assessment data collected in the STARS process.
 - a. Strongly Disagree
 - b. Disagree
 - c. Neutral
 - d. Agree
 - e. Strongly Disagree
5. Technology assists teachers in understanding the data collection process.
 - a. Strongly Disagree
 - b. Disagree

- c. Neutral
 - d. Agree
 - e. Strongly Agree
6. Technology assists teachers in utilizing the data collected in the STARS process for instructional decision making.
- a. Strongly Disagree
 - b. Disagree
 - c. Neutral
 - d. Agree
 - e. Strongly Agree

Future Systems and Issues

This portion of the survey gathers data that will help the Nebraska Department of Education make decisions about how to support the STARS process. Please rate each of the following four items on a scale of 1 to 5, with 1 being “Low,” and 5 being “High.”

(IRB NOTE: The website survey will provide respondents with a five-point scale selection field for each item as shown in Item 1.)

1. Please rate each of the following potential barriers related to using technology to support the assessment process:
- | | | | | | |
|--|---|---|---|---|---|
| a. Initial cost | 1 | 2 | 3 | 4 | 5 |
| b. Maintenance costs | 1 | 2 | 3 | 4 | 5 |
| c. Commitment to single vendor | 1 | 2 | 3 | 4 | 5 |
| d. Knowledge to choose best system | 1 | 2 | 3 | 4 | 5 |
| e. Technical expertise to maintain | 1 | 2 | 3 | 4 | 5 |
| f. Fear of obsolescence | 1 | 2 | 3 | 4 | 5 |
| g. Teacher training | 1 | 2 | 3 | 4 | 5 |
| h. Interoperability with other systems | 1 | 2 | 3 | 4 | 5 |
| i. Staff time to learn system | 1 | 2 | 3 | 4 | 5 |
| j. Administrator time to learn system | 1 | 2 | 3 | 4 | 5 |
2. Please rate your perception of the quality of currently available technology used to support the collection, management, and reporting of assessment data.
- a. Systems are comprehensive
 - b. System quality
 - c. Classroom level reporting features
 - d. District level reporting features
 - e. Interface with State reporting
3. Please rate the desirability of each of the following characteristics in a technology-based data collection, management, and reporting system.
- a. Ease of data entry
 - b. Usefulness of results for instructional decisions
 - c. Comprehensive features

- d. Match with reporting requirements
 - e. East of learning system
 - f. Interoperability with other systems
4. Please rate the likely impact of each of the following potential areas of state support or resources:
- a. Publication of guidelines for data management and reporting systems
 - b. Support for shared experiences among school districts
 - c. Case studies of successful implementations
 - d. Grants for training teachers
 - e. Grants for technology implementation
 - f. Development of models for data collection, management and reporting
 - g. Publication of interoperability standards and reporting requirements
 - h. Provision of software templates for collecting and organizing data

Appendix L
2006-2007 Study VIII: Statewide Writing Assessment
Scoring Rater Evaluation

2007 Statewide Writing Assessment Scoring

Rater Evaluation

Thank you for taking a few moments to complete this survey. Your responses will assist the Nebraska Department of Education in the continued implementation of the Statewide Writing Assessment scoring process.

Rater Information

1. Gender (circle) M F
2. Years of classroom teaching experience _____
3. Highest level of education ___ Bachelors ___ Masters ___ Ed Specialist ___ Doctorate
4. Have you previously participated as a rater in the Statewide Writing Assessment? ___ yes ___ no
5. If yes, please indicate *when* (check) and which *grade(s)* (circle)
 ___ 2001 (4, 8, 11) ___ 2002 (4) ___ 2003 (8) ___ 2004 (4, 8, 11) ___ 2005 (4, 8, 11)
 ___ 2006 (4, 8, 11)
6. Have you participated as a rater for your district/building/ESU writing assessment? ___ yes ___ no
7. Have you previously attended a workshop or training in Six Trait Writing? ___ yes ___ no
8. Are you currently a classroom teacher? ___ yes ___ no
 - a. If yes, what grade level(s) do you currently teach _____
 - b. If yes, what subject area(s) do you currently teach _____
9. If you currently are *not* a classroom teacher, what is your status?
 ___ school administrator ___ retired teacher ___ substitute teacher ___ student teacher
 ___ other (please explain) _____

Scoring Training

Please indicate your level of agreement or disagreement with the following items by circling your choice (SD = Strongly Disagree; D = Disagree; N = Neither Agree nor Disagree; A = Agree; SA = Strongly Agree)

1. SD D N A SA The general content of the training for scoring was appropriate (i.e. overview of the scoring process, understanding the scoring criteria, 6 traits of writing, writing mode characteristics).
2. SD D N A SA Strategies implemented during training to ensure scoring accuracy were effective (i.e. strategies for rubric use to assign scores, using anchor papers to demonstrate performance levels, use of re-calibration/practice papers, and consensus building discussions).

3. SD D N A SA The training materials (i.e. rubrics, anchor papers, rater scoring manual, training video) used during the training and scoring sessions were clear and understandable.
4. SD D N A SA The scoring leader demonstrated expert knowledge of the assessment and scoring procedures.
5. SD D N A SA The scoring leader demonstrated skill in teaching raters how to fairly and accurately score the writing assessment.
6. SD D N A SA The table leaders demonstrated expert knowledge of the scoring procedures.
7. SD D N A SA The table leaders were effective in facilitating the scoring process through re-calibration procedures (e.g. read behinds, resolving issues of discrepant scoring).

Scoring Experience

Please indicate your level of agreement or disagreement with the following statements about being a rater by circling your choice. Being a rater has:

1. SD D N A SA ...increased my understanding of the qualities of good writing.
2. SD D N A SA ...helped me develop more accurate grade-related expectations for student writing
3. SD D N A SA ...given me greater appreciation of writing's uses in content-area learning
4. SD D N A SA ...helped me better understand how writing relates to thinking and reasoning
5. SD D N A SA ...given me new ideas I can use in the teaching of writing in my classroom
6. SD D N A SA ...affirmed my current classroom writing instruction practices
7. SD D N A SA ...contributed to my ability to help my students become better writers
8. SD D N A SA ...increased my confidence in making accurate ratings of the quality of student writing
9. SD D N A SA ...built my confidence in giving my students helpful feedback on their writing
10. SD D N A SA ...increased my general confidence as a teacher of writing
11. SD D N A SA ...given me assurance in talking to parents and others about writing and its importance
12. In 2 or 3 sentences, please explain why you wanted to be a rater for the Statewide Writing Assessment.
13. Please share any additional comments/thoughts you would like to share about your experience as a rater or about the scoring process.