

**Written by Jody C. Isernhagen
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**

**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
September 2005**

TABLE OF CONTENTS

CHARTING STARS: VOICES FROM THE FIELD

Section 1: Introduction	5
Section 2: Year Four Summary	9
Section 3: Year Four Primary Research Study	27
Section 4: Aspiring for Excellence: Configurations.....	31
Theme 1: <i>Total Support</i>	37
Theme 2: <i>Assessment Literacy</i>	47
Theme 3: <i>Data</i>	55
Theme 4: <i>Instructional Impact</i>	65
Theme 5: <i>External Support</i>	75
Theme 6: <i>Accountability</i>	79
Theme 7: <i>Leadership</i>	83
Theme 8: <i>Professional Development</i>	87
Theme 9: <i>Challenges and Obstacles to Success</i>	91
Theme 10: <i>Next Steps for Success</i>	99
Appendices:	
A. Researchers for the Project.....	107
B. Effective Leadership for Learning (Isernhagen & Dappen).....	111
C. 2002-2004 Nebraska Criterion Referenced Tests (Dappen & Isernhagen)	119
D. Student Performance on Nebraska’s STARS Writing (Dappen & Isernhagen)	127
E. Statewide Writing Assessment Scoring (Anderson)	133
F. Educators’ Perceptions of STARS (Boss, Endorf, Hefflebower, & Warrick).....	149
G. Nebraska Public School Superintendents’ Perceptions (Clarke).....	157
H. Validation Framework	169
I. IRB Approval Letter.....	177

J. 2004-2005 Primary Research Survey Sample 181

K. 2004-2005 Primary Research Interview Protocol 187

TABLES AND FIGURES

Tables		Page
1	Purposes and Advantages Identified by Educators for Joining a Configuration	11
2	STARS District Return Rate by Configuration 2004-2005	30
3	Average Means for Survey Categories by Configuration 2004-2005	30

Figures		
1	Voices of Nebraska Educators: Average Means by STARS Category.....	12
2	Percentage of Nebraska School Districts by Configuration 2004-2005	28
3	Teacher Response Range by Configuration STARS Survey 2004-2005: Total Support	37
4	Principal Response Range by Configuration STARS Survey 2004-2005: Total Support	39
5	Superintendent Response Range by Configuration STARS Survey 2004-2005: Total Support	42
6	ESU Response Range by Configuration STARS Survey 2004-2005: Total Support.....	44
7	Teacher Response Range by Configuration STARS Survey 2004-2005: Assessment Literacy	47
8	Principal Response Range by Configuration STARS Survey 2004-2005: Assessment Literacy	50
9	Superintendent Response Range by Configuration STARS Survey 2004-2005: Assessment Literacy	52
10	ESU Response Range by Configuration STARS Survey 2004-2005: Assessment Literacy	54
11	Teacher Response Range by Configuration STARS Survey 2004-2005: Data.....	55
12	Principal Response Range by Configuration STARS Survey 2004-2005: Data	58
13	Superintendent Response Range by Configuration STARS Survey 2004-2005: Data.....	61
14	ESU Response Range by Configuration STARS Survey 2004-2005: Data	63
15	Teacher Response Range by Configuration STARS Survey 2004-2005: Instructional Impact.....	65
16	Principal Response Range by Configuration STARS Survey 2004-2005: Instructional Impact.....	68
17	Superintendent Response Range by Configuration STARS Survey 2004-2005: Instructional Impact.....	72
18	ESU Response Range by Configuration STARS Survey 2004-2005: Instructional Impact.....	74
19	Teacher Response Range by Configuration STARS Survey 2004-2005: External Support	75
20	Principal Response Range by Configuration STARS Survey 2004-2005: External Support.....	77
21	Superintendent Response Range by Configuration STARS Survey 2004-2005: External Support.....	77
22	ESU Response Range by Configuration STARS Survey 2004-2005: External Support.....	78

CHARTING STARS: VOICES FROM THE FIELD

introduction

This is the fourth annual report of the Comprehensive Evaluation Project (CEP), an independent evaluation of Nebraska's School-based Teacher-led Assessment and Reporting System (STARS). The CEP was originally contracted between the Nebraska Department of Education and the University of Nebraska-Lincoln, College of Education and Human Sciences in 2001. The CEP is 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 and Dr. Leon Dappen, Assistant Professor, served as Secondary Investigator. All researchers for the Comprehensive Evaluation Project are listed in Appendix A.

Overview

Nebraska schools are aspiring for excellence in assessment and school improvement. A unique and innovative standards, assessment, and accountability system beginning in 2000 was established by the Nebraska Legislature in LB 812.

The first year's comprehensive evaluation study focused on the effectiveness of Nebraska's unique approach to standards, assessment, and accountability. In the second year, the study provided an analysis of the sustainability of the process, curriculum and instruction, school leadership, and professional development of STARS. The third year study focused on the collection, interpretation, and use of assessment data, how important the data were, what schools were learning from the data, and what schools were doing with the data. Throughout this process of evaluation, a growing number of educators were asking new questions about Nebraska's assessment and accountability model because they felt it was the right thing to do for Nebraska's children.

Summary of Fourth Year Study

Nebraska has built a School-based Teacher-led Assessment and Reporting System (STARS) that required each of Nebraska's 495 school districts to develop a local assessment system to measure student performance on local or state standards. This process honors educators and trusts their professional judgment, but also demands work and a great deal of leadership based on a clear vision for learning. As a result of this vision of excellence, school leaders

and teachers joined together to share expertise, time, and resources to develop local assessment systems that truly honor the teaching taking place in classrooms and the best interest of students.

For the last four years, Nebraska educators have created new learning through professional development and dialogue, developed and refined assessments, collected and analyzed data, and implemented best instructional practices with the support of external agencies. During this process educators in school districts began to share and dialogue with others in the state to identify better ways of delivering STARS. As a result of this dialogue, three configurations of delivery emerged.

Year four's primary study examined these three configurations—collaborations, consortia, and individual districts—to determine the impact of STARS on assessment literacy, use of data in classroom settings, leadership, and support from external agencies. Nebraska was comprised of 495 school districts when the year four study began. Of the 495 participating school districts, 72 were collaborations, 289 were consortia, and 134 were individual districts.

In addition to the primary study, six partnerships conducted pertinent studies of the implementation of the STARS process in 2004-05 as a part of the Comprehensive Evaluation. Six quantitative studies were conducted that generated additional recommendations for STARS:

Effective Leadership for Learning conducted by Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln and Dr. Leon Dappen, Assistant Professor, University of Nebraska-Omaha (Appendix B).

An Examination of 2002 to 2004 Nebraska Criterion Referenced Tests, Norm Referenced Tests, and District Portfolio Ratings for Math at Grades 4, 8, and 11 conducted by Dr. Leon Dappen, Assistant Professor, University of Nebraska-Omaha and Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln (Appendix C).

A First Look at Student Performance on Nebraska's STARS Statewide Writing Test at Grades 4, 8, and 11 conducted by Dr. Leon Dappen, Assistant Professor, University of Nebraska-Omaha and Dr. Jody Isernhagen, Associate Professor, University of Nebraska-Lincoln (Appendix D).

Statewide Writing Assessment Scoring: Effective Professional Development for the Classroom Teacher conducted by Sue Anderson, NDE Coordinator of Statewide Writing Assessment (Appendix E).

Educators' Perceptions of STARS conducted by UNL Graduate Students: Toby Boss, ESU #6 Staff Developer; Dan Endorf, York High School Principal; Tammy

Heflebower, Staff Development Director, ESU #6; and Phil Warrick, Waverly High School Principal; (Appendix F).

Nebraska Public School Superintendents' Perceptions of Nebraska's Assessment/Accountability System conducted by Dr. Wanda M. Clark, Instructional Research Administrator, Omaha Public Schools (Appendix G).

Pat Roschewski, Director of Statewide Assessment for the NDE, and the CEP designed a Validation of Nebraska's Standards, Assessment, and Accountability System (Appendix H) for 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.

This 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 2004-05 school year. All of the studies contributed to a comprehensive picture of assessment practices and support systems within the state that provide assistance to improving student learning in Nebraska.

Year Four Comprehensive Evaluation Format

The report is divided into sections beginning with an introduction of the total report (Section 1), followed by a summary of the complete findings of all studies conducted during the fourth year study (Section 2). The third section (Section 3) focuses on the methodology of the primary research study followed by educators from the field describing their views of collaborations, consortia, and individual districts (Section 4). Additional sections describe each theme from the primary research study in detail (Themes 1-10). The appendices provide papers from each of the supplementary studies conducted during the fourth year study.

Acknowledgements

To each of the educators that welcomed our visits and helped arrange the interviews, and for those that participated in the interviews, we offer a special thanks for sharing your expertise and STARS experiences.

We wish to thank all of the researchers that conducted studies included in the Fourth Year Report. We wish to acknowledge the support provided by the NDE staff, including Commissioner of Education, Douglas Christensen; Pat Roschewski, Director of Statewide Assessment; Sue Anderson, Coordinator of the Statewide Writing Assessment; Dottie Heusman, National Assessment of Educational Progress (NAEP) Coordinator; and Jackie Naber, Office Administrator of Statewide Assessment.

We would like to express our appreciation to our team of interviewers, Larry Bornschlegl, Ron Klemke, and Bob Whitehouse for conducting the many interviews in school districts across the state. I would personally like to acknowledge the many hours and expertise

contributed by Leon Dappen and Shirley Mills to the development and completion of the primary research project. I could not have done it without the two of you.

For helping to bring the project to fruition, we thank Susan Wilson, Cindy DeRyke, Susan McCoy, and Tammie Herrington. 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 support of the Comprehensive Evaluation Project.

Charting Stars: Voices from the Field

Year Four Report Stars Comprehensive Evaluation

Summary

“The key is to give the highest priority to structures that allow educators to work as members of true research teams and thus to become, as Fullan notes, ‘scientist who continuously develop their intellectual and investigative capacities.’”

Schmoker (2004, p. xvi)

INTRODUCTION

Through professional development efforts across the state over the past four years, Nebraska educators have acquired a new language previously unknown at the classroom level. Teachers now write and revise the STARS assessments used to measure Nebraska content standards. Their assessments are based on the six indicators of high quality assessment as designed by the Buros Center for Testing in partnership with the Nebraska Department of Education. The six quality indicators for assessments used in each district must 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, and 6) have appropriate mastery levels. Each year Nebraska school districts develop an assessment portfolio that is based on the clear guidelines for writing assessments that measure state or local standards. The district portfolio outlines the district assessment process used to meet the quality criteria. The portfolios are examined and rated by outside experts to verify that school districts are designing and using quality assessments.

FOURTH YEAR PRIMARY RESEARCH STUDY

Overview

The fourth year of the STARS Comprehensive Evaluation Project (CEP) focused upon the impact of STARS on assessment literacy, use of data in classroom settings, leadership, and support from external agencies. The study concentrated on the three configurations—collaborations, consortia, and individual districts. These three configurations became apparent as schools struggled to implement the STARS process in a very diverse landscape across Nebraska. Each school district determined the best configuration to assist their district with implementation of the STARS process. Pat Roschewski, Director of Statewide Assessment, for the Nebraska Department of Education (NDE), defined the configurations in the following manner:

Collaboration – A group of districts that work together to develop and/or share individual assessments and submit individual assessment portfolios to NDE. All portfolios for districts participating in collaboration are sent to the same reviewer.

Consortium – A group of districts that work the STARS process together, use all the same assessments, and submit a single assessment portfolio to NDE. There are three consortium models: Online consortium, Panhandle consortium, and “Other” consortia.

Individual –A district that develops its own assessment system and submits independently its District Assessment Portfolio to the NDE.

Research Design and Methodology

The fourth year primary study was a mixed-methods design, selected for use to strengthen the study results. The study began with a broad survey in order to generalize results to a large statewide population of educators at all levels. Secondly, open-ended interviews were conducted to collect detailed views about the STARS process in each of the configurations. Researchers under the direction of Jody Isernhagen, at the University of Nebraska-Lincoln (UNL), conducted this study.

In the research study, approximately 2000 surveys were sent to teachers, principals, superintendents, assessment coordinators, and Educational Service Unit staff developers to determine the impact of STARS on assessment literacy, use of data in classroom settings, leadership, and support from external agencies. Nebraska educators across the state were given voice during the 169 interviews conducted following the survey during the 2004-05 school year. The interviews were designed and conducted based on the results of the survey in an effort to better understand the impact of the implementation of the STARS process.

RESULTS

Purposes for Joining a Configuration

Educators identified the purposes and advantages they believed were essential for belonging to a collaboration, consortium, or individual district configuration on the survey. Those purposes and advantages are identified by priority in Table 1.

Table 1: Purposes and Advantages Identified by Educators for Joining a Configuration

1. To acquire support and opportunities for collaboration with other professionals
2. To use time wisely
3. To share money and personnel
4. To help write assessments, align curriculum, and identify standards of achievement for the configuration
5. To ensure district continuity with STARS guidelines
6. To assist in meeting the six quality criteria

Voices from the Configurations

I look at school before we had STARS and the fact that teachers were pretty much freelancing, that they could teach whatever they wanted just wasn't right. And I think it's given us some direction. . . .It really makes that connection between the assessment process and the learning process and how they all go together. And I feel very lucky to have had those opportunities. . . .This process is good and getting better. . . .But without the state department setting the bar, it would never have forced us to get to where we are today.

Voices of Nebraska Teachers

Educators in collaborations shared the “wonderful opportunity to dialogue between and among other educators.” Working together, they produced quality assessments that they were able to use within their district. This could not have been done without the support of ESU staff developers as they guided the development of the assessment process and assisted educators in analyzing their own data to improve instruction on a daily basis. Educators in collaborations have developed a common language that bonds them together, and have built an interdependent community of learners.

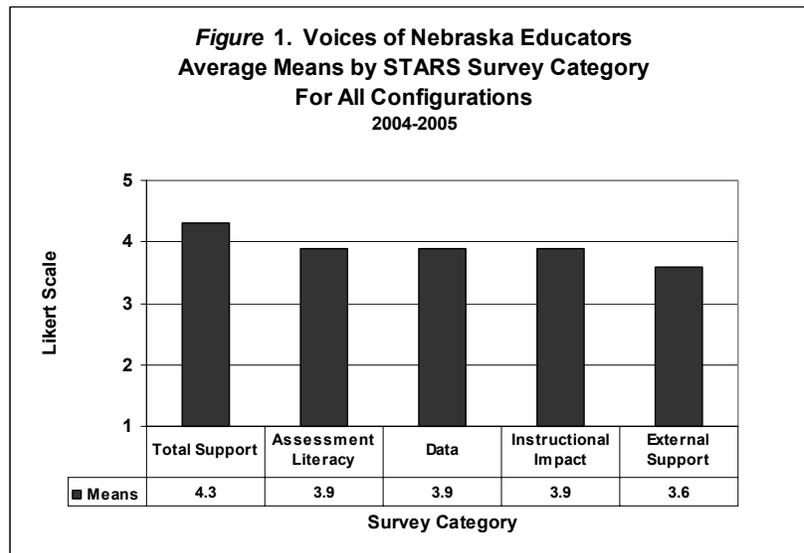
Educators in consortia alluded to their consortium as “a lifesaver” serving many small schools that would be “lost without it.” The bigger value of belonging to a consortium was “being around other teachers and finding out what they’re doing in the classroom and how they’re approaching standards.” Educators in the Online consortium appreciated the immediate feedback provided to teachers and students because of the technology used to administer and score their assessments. “Other” consortia “worked together as a team so that they’re not out there alone . . . to share ideas.” Finally, the Panhandle consortium involved educators from all over the area to create and revise assessments and curriculum based on the state standards “so that they’d be testing and teaching the same kinds of things.”

Educators in individual districts valued “local control” and the “ability to assess what we actually teach.” The educators believed that the value of belonging to an individual district was the importance and the need of working together. It works very well and is a wonderful system “because it is teacher involved and teacher driven.” “But having the time and freedom to change and adapt the assessments as needed” provided strong ownership in the final product. Individual districts “set up assessments and make sure that curriculum is aligned with standards” and assessments were developed at each grade level. An individual

teacher stated, “I feel I have had a say in what goes into the assessments.” “My voice is heard” as individual districts valued their teachers’ opinions. “Strong ownership in the final product” was developed. “Local control for the development of standards and assessments is very empowering.”

Findings

Survey results indicated strong support for the STARS process across all three configurations. The average mean for all items for each survey category for all configurations ranged from 3.6 to 4.3 on the five-point Likert scale with “1” representing “none of the time” and “5” representing “all of the time.”



The interviews supported the five themes examined on the survey as well as providing five additional themes. The themes helped to paint a picture of what collaborations, consortia, and individual districts looked like as they grew in their knowledge of assessment, collection and interpretation of data, and application of their new learnings to instruction.

Theme 1: Total Support for the STARS Process

In the area of total support, all configurations ranged from 3.90 to 4.58 on the five-point Likert scale with “1” representing “none of the time” and “5” representing “all of the time.” There was a perception of strong district support for the STARS process across all configurations. There were some statistically significant differences in the perceptions regarding district services to meet school needs for the implementation of STARS. Individual districts’ perceptions were lower regarding district provision of services and district providing on-going assessment training for teachers. Consortium districts perceived that their district support was high in all areas. Collaborations believed that their district support for teacher training was higher than that of administrator training.

There are three different types of consortia—Panhandle, Online and “Other.” In the area of total support, all three types of consortia ranged from 3.94 to 4.72 on the Likert scale. The

Panhandle consortium perceived district support for “school-based teacher-led” assessment significantly lower than the Online and “Other” consortium districts.

Educators shared their support for the STARS process as practiced in and perceived from each configuration. Teachers indicated that STARS had made them better teachers regardless of their own personal level of experience. One veteran teacher summed her experiences up nicely by stating,

I guess I think it’s made me a better teacher. And I think after teaching for a number of years, we tend to get in a rut and sometimes think we’ve got it all down; we know what we’re doing and I think we need to be reminded that there’s always room for us to learn.

Principals in all configurations indicated that educators were more comfortable with the STARS process now that many are trained and have had time to implement the change. But . . . “Transition takes some time. They’re more at ease than what they were a couple years ago, which is good.”

Theme 2: Assessment Literacy

In the area of assessment literacy, all configurations ranged from 2.65 to 4.46 on the five-point Likert scale. There were some statistically significant differences in perceptions regarding assessment literacy of all districts participating in the STARS process.

Collaboration and consortium districts were significantly stronger than individual districts regarding local district development of assessments. Collaboration districts were also stronger in teacher collaboration for the design of assessments. Individual districts were significantly stronger than either collaboration or consortia in the development of the STARS portfolio and using the same grade-level assessments to meet the STARS requirements.

In the area of assessment literacy, all consortia ranged from 2.50 to 4.61 on the five-point Likert scale. The Panhandle consortium perceived assessment literacy significantly lower in the areas of district-developed assessments, teacher use of rubrics, and teacher participation in learning teams. Districts in “Other” consortia consistently scored higher in developing assessments and use of rubrics in assessment. The Online consortium scored significantly higher in participating in learning teams but was slightly lower than “Other” consortia in developing standards-based assessments and the use of rubrics.

Assessment literacy was evident in the common vocabulary used by educators in all configurations when speaking about assessment development, quality criteria, data collection, and interpreting and reporting data. One principal said:

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 students to that? I think they’re more concerned with what they’re assessing and how students are progressing.

Teachers demonstrated growth in assessment literacy in all configurations by stating that the process had become easier over time and their ability to identify what needed to be changed and revised had improved. One teacher summed it up as “I’ve learned about what are good question techniques, watching for the biases and just knowing what a good assessment looks like.” Many districts were developing their own assessments in a collaboration or within their own district. Some districts were using assessments from a common bank or pool. However, data from this study indicated that teachers in districts using questions from a common bank or pool were involved in developing and selecting the individual questions for their assessments. On the other hand, some teachers did reveal that their curriculums sometimes did not match the assessment items being used to measure the standards.

Theme 3: Data

In the area of data, all configurations ranged from 2.36 to 4.58 on the five-point Likert scale. There were some statistically significant differences in perceptions regarding data, district scoring, and interpreting all STARS assessments. Individual districts’ perceptions were significantly stronger statistically than both collaboration and consortia districts. Consortia districts had a slightly stronger perception of timely return of data than either collaboration or individual districts.

In the area of data, all consortia ranged from 2.27 to 4.71 on the five-point Likert scale. The Online consortium perceptions were significantly stronger for scoring all STARS assessments. The Panhandle and “Other” consortia perceptions were significantly lower in this area.

Districts in all configurations were making data available to district educators in a variety of disaggregated formats. However, not all data was scored in the same way nor was it always provided to educators in a timely manner. Data was being interpreted in a wide variety of ways, but data was not always being used for instructional improvement. A principal shared the importance of involving staff in the use of data:

I can look at the data and say that this is where we’re going to go and maybe as an educational leader that’s just what we have to do. But I think when the teachers take ownership of it, and they see the data, they know where it’s going to go, and they know where they want it to go, and they’re able to implement it so much better when they believe it and buy into it than just the principal saying this is where we’re going.

Some school districts were in the early stages of using data and still needed assistance as noted by a superintendent,

I think we’ve only begun to understand how to read data. And data for the most part has scared us because we didn’t really know how to make as much out of it as we could. So I would say, all of us from our physical education teacher right through to the high school principal and probably the superintendent, better training in how to use data to make good, sound educational decisions.

Some districts were large enough to have the expertise needed to solve data problems while other schools lacked the needed expertise or resources and saw data management as a huge obstacle. One superintendent noted:

My role is to say, if we can't break it out by free and reduced lunch, figure out a way to get it broken out by free and reduced lunch. We are working real hard at taking the technical aspects of assessments out of the hands of teachers and allowing teachers to teach. Somebody else has got to do the technical parts of it and our teachers have got to do the instructional part of it.

Theme 4: Instructional Impact

In the area of instructional impact, all configurations ranged from 3.48 to 4.42 on the five-point Likert scale. There were some statistically significant differences in perceptions regarding instructional impact in the alignment of curriculum to state standards and modifying instruction in cases where students did not perform well on assessments. Individual districts had a stronger perception than both collaboration and consortium districts in curriculum alignment, establishing benchmarks for meeting state standards K-12, and aligning lesson planning to assessment data and state standards. Consortia districts had a stronger perception in teachers modifying instruction than did collaboration and individual districts. Consortia districts' perceptions were slightly less than collaboration districts when considering curriculum alignment to state standards.

In the establishment of benchmarks for meeting state standards K-12, collaboration districts were significantly lower than both consortium and individual districts. Additionally, lesson planning, alignment to assessment data, and state standards were also significantly lower for collaboration districts than consortia and individual districts.

In the area of data, all consortia ranged from 3.25 to 4.33 on the five-point Likert scale. "Other" consortia perceptions were significantly higher on aligning curriculum to state standards, establishing benchmarks for meeting state standards K-12, and modifying instruction in cases where students did not perform well. Panhandle consortium perceptions were slightly less in these same areas, while the Online consortium perceptions were significantly lower in all three areas.

Strong evidence from educators across the state for all configurations indicated that schools have aligned their curriculum to state standards. However, there was little evidence to support K-12 benchmarks within the state. Teachers spoke of aligning their lesson plans to assessment data, but their instruction was sometimes not modified as a result of student performance. Teachers were questioning what they had been taught in the past and its value to students. A secondary teacher stated:

I think everyone who has to give these assessments has really gone through and questioned every daily activity you do with your students as far as trying to plan for an entire year. You realize there's not very much time for anything that you did before that wasn't directly connected to the state standards. But, we are definitely more calendar driven in our plans, and in our activities, and in our daily assignments

for our kids. It's much tighter. You really question what you've done and if it's serving the purpose that it needs to get those state standards covered. A lot of soul searching, I guess, as far as a teacher.

Principals said that they believed that teachers were making the link between data, student achievement, and improving learning by sharing data and information:

Communication about student achievement and learning is at an all time high.

Linking the STARS process to school improvement was a significant step to the improvement for student learning.

Theme 5: External Support

In the area of external support, all configurations ranged from 2.73 to 4.23 on the five-point Likert scale. There were some statistically significant differences in perceptions regarding external support provided to districts. Collaboration districts' perception of the NDE providing assessment training for teachers was significantly lower than consortia and individual districts.

In the area of external support, all consortia ranged from 2.76 to 4.36 on the five-point Likert scale. "Other" consortia perceptions were significantly stronger than Online and Panhandle consortia districts for providing time, resources, and on-going assessment training for teachers. The Online consortium was slightly less in these three areas, with the Panhandle consortium being significantly lower.

Our research indicated strong evidence that external support was being provided by the ESUs and the NDE for all configurations. However, beginning teachers reported limited knowledge of assessment requirements and the STARS process. Additionally, schools reported that external support was provided by local workshops, the University of Nebraska-Lincoln, Buros Center for Testing, North Central Association Commission on Accreditation and School Improvement, Assessment Training Institute (Richard Stiggins), and Mid-Continent Research for Education and Learning (McREL).

External support for technology regarding assessment data collection and analysis was limited. Much of the data collection, scoring, analysis, and reporting was not centralized and lacked uniformity. Leadership from both the NDE and the ESUs is needed to make data management more efficient and accurate.

Theme 6: Accountability

Most language arts and math educators in grades 4, 8, and 11 demonstrated great ownership for the STARS process and were accountable for student learning. STARS has raised the bar for teachers and students while improving student learning at these levels. However, teachers at grades 4, 8, and 11 expressed the feeling that "other teachers, who have not had to participate, are uninformed of the process and the tremendous amount of work that goes into it." Educators emphasized that "all" teachers need to be involved in the STARS process. Most educators emphasized the growth that had occurred for both teachers and students due

to their involvement in the STARS process and meeting the accountability requirements established by the NDE and No Child Left Behind (NCLB).

Some concern was expressed by teachers and administrators that the portfolio process was cumbersome and took time away from teaching. Other educators indicated that the portfolio process did not truly measure the quality or rigor of the individual classroom assessments but instead measured the process that was put into place to develop the assessments. A superintendent emphasized a much broader picture of accountability by stating:

We have learned that we are not only accountable to our students, but I think it's made us more aware of accountability to a broader picture...the state of Nebraska and really, I think we're accountable to a certain extent to every student and every administrator and every teacher within the state of Nebraska.

Theme 7: Leadership

Teachers were feeling empowered and administrators were stepping up as instructional leaders in all configurations. New instructional leadership skills were emerging not from just principals and superintendents but from teachers in the field as illustrated by this statement,

And throughout the entire process we've made revisions on a yearly basis trying to see what we can do to make those assessments more successful, not only for teachers, but also for our students and their needs. Within my building, I've been a leader as far as helping other teachers understand the assessment process.

However, buy-in across the board remains an elusive goal as some teachers and administrators struggle with STARS and the implementation of standards, assessment, and accountability on a daily basis. One principal said:

But what I'm finding over the years, this whole conversation about standards and STARS is becoming really standard operating procedure, even from the veterans who are really resistant to it. I think it's one of those things that over time have pretty much taken over the school culture.

Theme 8: Professional Development

Professional development was evident for teachers of language arts and math in grades 4, 8, and 11, principals, superintendents, and assessment coordinators across all configurations. The Educational Service Unit (ESU) responses supported the evidence that districts were participating actively in professional development and having conversations at the district level concerning STARS. Staff developers also indicated that they had to always be one step ahead of schools as noted below,

I think you know in the world of staff development it's a constant learning curve, if it isn't, there's something wrong with what you're doing. Because you need to always be out in front, you need to be on the cutting edge of what's about to occur and understand it because the schools have definite expectations of you knowing how to lead them.

Staff developers shared that teachers have different conversations today than they had in the past as evidenced by this statement,

If the teachers could hear themselves today compared to five years ago, they would not even know they were the same teachers. Their vocabulary, their literacy in the assessment world, their ability to figure out how to get it all together, their reporting, their knowing about how to make things reliable, they just had no idea that they could stretch like that.

Teachers indicated that the opportunity to dialogue with other teachers was the greatest benefit of participating in the STARS process. There was a major concern voiced by all educators in the survey and interviews about the time and resources it takes to train educators in the process and maintain the most up-to-date assessment knowledge to better implement the process across all configurations. This contributed to the challenge of finding time for educators to do the work effectively.

Theme 9: Challenges and Obstacles

As with the implementation of any new initiative, there were challenges to be faced and obstacles to be overcome across all configurations. The primary obstacle facing the implementation of the STARS process was “time.” This is not a new issue but an ongoing issue identified as early as 2001 by Gallagher in the initial comprehensive evaluation report. Educators initially needed time to learn the new skills associated with building an assessment and accountability system within their own school districts. Once the new skills were learned and implementation began, many other challenges arose as the STARS process became increasingly embedded in the day-to-day work of the school.

Time was identified as an issue by teachers, principals, superintendents, assessment coordinators and staff developers in every interview conducted. There was a consistent call across all three assessment configurations for teachers to be compensated for time outside of the school day that they were putting into making the process successful. Changing the school calendar or length of the school day to give teachers additional time to provide quality assessment and reporting of student achievement was one suggestion for increasing time. This should only be considered if all teachers at all grade levels are involved in the STARS process.

Other challenges and obstacles included were: ownership and participation by all grade level teachers; coping with continuous change that invokes conflict concerning philosophical issues that impact student achievement; providing the opportunity for all students to learn before being tested as some students are unable to master advanced subject matter called for in some standards; reducing the number of standards being reported in reading and math as in science and social studies; providing tests that are appropriate for all learners, i.e., Special Education and English Language Learners; the appropriate use of summative and formative assessment; a reconsideration of the use of the STARS portfolio as the way that school districts validate their assessment systems; development of a statewide policy and/or decision on how many times a student can be re-tested before reporting final data to the state; assistance with technology and data management issues at the school and district levels; and finally, a feeling that a state test should be considered (as expressed by a limited number of research participants).

Theme 10: Next Steps for Success

As educators continued their implementation of the STARS process within their school districts, many of them shared thoughts about the next steps to make the process more successful and manageable. They shared ideas about the need for ensuring that all students had the opportunity to learn by guaranteeing that all districts in configurations align their curriculum to match standards and the assessments that measure those standards; changing school schedules to provide increased time for mastering content or remediation for students not showing mastery of standards; increasing credit hours for graduation in major content areas; offering courses more frequently or before the senior year; development of policies on re-testing of students prior to reporting; providing assessments more appropriate for all learners i.e., Special Education and English Language Learners; finding time and resources to support continued dialogue and rewriting of assessments by teachers; reducing the number of standards reported in reading and math; and finding alternatives to the portfolio system currently used by NDE for ensuring a quality assessment system.

Suggestions regarding data included creating a data management system that was more manageable and allowed for more timely return of data with the help of technology. The theme of professional development prompted suggestions that educators be provided with continued training in the use of data and the interventions needed to support all students needing additional assistance and that the responsibility and accountability for assessment literacy be shared by all grade level teachers providing benchmarks for meeting all standards.

SUPPLEMENTARY QUANTITATIVE STUDIES

Six supplementary quantitative studies were conducted in 2004-05 that generated additional recommendations for STARS. A summary of results is shared for each study. The full report for each study is available in the Appendices of this document.

Effective Leadership for Learning

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

Dr. Leon Dappen, Assistant Professor, University of Nebraska-Omaha

Purpose

The purpose of this study was to survey perceptions of the leaders involved, to examine the status of leadership efforts regarding the development and implementation of STARS.

Findings

A survey of Nebraska board of education presidents, superintendents, assessment coordinators, elementary principals and secondary principals indicated they all perceived assessment data as being used to impact achievement. Standards, assessment, and accountability efforts were having a strong impact on teaching and learning. All groups reported having had at least “some to extensive” related training in their career, particularly in recent years. They indicated a common vision of assessment with other district leaders.

On a survey examining areas most important to the support for these efforts, the groups rated “leadership for assessment reform” highest, followed by “use of achievement standards,” “development of assessment literacy,” “commitment to all users,” and “effective communication.” Supportive policies were seen as the least related to the impact upon achievement.

**An Examination of 2002 to 2004 Nebraska Criterion Referenced Tests,
Norm Referenced Tests, and District Portfolio Ratings
For Math at Grades 4, 8, and 11**

Dr. Leon Dappen, Assistant Professor, University of Nebraska-Omaha

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

Purpose

The purpose of this study was to examine the first achievement data available in math for the Nebraska STARS program from 2002 to 2004. Both locally developed CRT scores and NRT scores were examined. This included just over 94% of the public school students in Nebraska.

Findings

The average percent of students meeting the defined mastery level for their locally developed Math CRT measure increased over six percent at each grade level from 2002 to 2004. This lends strong support to continuing CRTs in the instructional activities for improvement of math achievement. Additionally, NRT scores increased three percent at the fourth and 0.5% at the eleventh grade, and declined 0.7% at the eighth grade. These scores did not appear to have been negatively affected by the implementation of the CRT program; indeed, there is some indication of slight increases.

Portfolio ratings from 2002 to 2004 increased strongly at each grade level. This would seem to reflect growing knowledge and skill in assessment practices by district staff. While this is initial information, the data are compelling to support the continuation of the Nebraska STARS approach.

**A First Look at Student Performance on Nebraska’s STARS Statewide Writing
Test at Grades 4, 8, and 11**

Dr. Leon Dappen, Assistant Professor, University of Nebraska-Omaha

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

Purpose

This study examined the first achievement data available for the Statewide Writing Assessment for STARS, comparing scores available in grades four, eight, and eleven. This included just over 94% of the public school students in Nebraska. The score reported is the average percentage of students in districts across the state that scored above the statewide cut score set for that year. Because of the primarily criterion referenced nature of the assessment, descriptive data is reported and discussed. However, the assessment was common across all

districts; therefore an equal interval scale and inferential statistics were used to examine statistical differences between pre/post scores.

Findings

Significant gains were made at the fourth and eighth grades where pre/post scores were available. At fourth grade, there was a 6.85% gain in the percentage of students demonstrating proficiency in writing, at the eighth grade a gain of 6.90%. Both are significant beyond the .001 level. Eleventh grade post scores were not available.

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

Sue Anderson, Coordinator of Statewide Writing Assessment, NDE.

Purpose

The purpose of the study was to ascertain the raters' perceptions of the effectiveness of the training they received during the scoring of the 2004 Nebraska Statewide Writing Assessment and determine the impact of participation in the scoring of the Statewide Writing Assessment on raters' classroom practices in the teaching of writing.

Findings

The results across all grade level scoring sessions showed that raters evaluated all components of the training they received as very positive and that they associated the scoring experience with benefits to their teaching. The findings supported 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.

Educators' Perceptions of STARS

UNL Graduate Students, Toby Boss, Staff Developer, ESU#6; Dan Endorf, Principal, York High School; Tammy Heflebower, Staff Development Director, ESU#6; and Phil Warrick, Principal, Waverly High School

Purpose

The purpose of this study was to determine the perceptions and practices of educators across the state of Nebraska as they implemented STARS.

Findings

The major finding of this study was that educators were generally positive in their perceptions of STARS. ESU staff developers gave the STARS model the most positive responses of the four educator groups surveyed; conversely, teachers gave STARS consistently lower marks relative to the other groups. Scores from assessment coordinators and principals were generally similar and usually placed between the scores of ESU staff developers and teachers.

All groups believed that public education in Nebraska improved due to STARS. Aligned with the generally positive impression of STARS mentioned above, 100% of ESU staff developers believed public education in Nebraska improved due to STARS. Seventy-six (76%) of assessment coordinators, 75% of principals, and 56% of teachers agreed that public education in Nebraska improved because of STARS.

Other findings that emerged from the data related to alignment of assessments with the Six Quality Criteria, the impact of STARS on school improvement, and the role of teachers as leaders of learning.

Nebraska Public School Superintendents' Perceptions of Nebraska's Assessment/Accountability System

Dr. Wanda M. Clark, Instructional Research Administrator, Omaha Public Schools

Purpose

The purpose of this study was to determine superintendents' perceptions of the Nebraska assessment/accountability reporting system (STARS) as they affected high schools. Data measuring the superintendents' perceptions of the STARS program included how they perceived the STARS program's impact on teachers, instructional practices, students, building principals, and resources.

Findings

Three main themes emerged from the data: 1) Overall, superintendents perceived STARS as having had no significant impact on high school teachers, instruction, students, or resource allocations; 2) Overall, superintendents perceived STARS to have had a positive impact on high school building principals' leadership practices ($M=3.78$, $SD=.602$), and; 3) Specifically, superintendents in districts with student populations that ranged between 100-1000 (81% of respondents) perceived that the impact of STARS on high school teachers and on instruction were less positive than did superintendents in school districts that ranged between 2000-5000 students (10% of respondents).

CONCLUSION

The future looks bright for continued growth and success for the STARS process if the challenges can be turned into successes and the obstacles can be minimized. A superintendent cautioned that we shouldn't be "afraid of the future, to pull back, to say, this didn't work, or make the modifications to put into practice what we expect of our educators. If, in fact, the strategy is not paying the dividends or the results, not to hesitate to pull that off the table, to say, let's regroup."

For the last four years, Nebraska educators have created new learning through professional development and dialogue. They have developed and refined assessments, collected and analyzed data, and implemented best instructional practices with the support of external agencies. Successful learning communities have been formed in school districts where

teachers dialogue on a regular basis, trying to wrestle with the problems that have arisen from the implementation of STARS. They have a wealth of experience to offer teachers across the nation about what true classroom assessment is all about and how student learning has been impacted positively by their own personal growth and the growth that has occurred for students.

The fourth year comprehensive study embodied six studies in addition to the primary study. All of the studies supported the primary research in special ways. Boss, Endorf, Heflebower, and Warrick's (2005) study noted that curriculum, instruction, and assessment practices were positively impacted by STARS as supported by an individual teacher's quote from the primary study, "If nothing else comes out of the whole process, we have good curriculums, people are using them, they're aligned to the standards, and there's a focus now in the learning process that wasn't there before." An individual principal in the primary study shared agreement by stating, "They (teachers) like what we're doing. They think it's good for kids. And STARS allows it. I'm eternally grateful for the leadership the Commissioner has given. I think he stuck his neck out to do the right thing."

The major finding of Boss et al. (2005) was that educators were generally positive in their perceptions of STARS. ESU staff developers gave the STARS model the most positive responses of the four educator groups surveyed; conversely, teachers gave STARS consistently lower marks relative to the other groups. Scores from assessment coordinators and principals were generally similar and usually placed between the scores of ESU staff developers and teachers. These findings supported the voices of educators in the primary study as they indicated strong support for STARS. Additionally, all groups in the Boss study believed that public education in Nebraska improved due to STARS. Aligned with the generally positive impression of STARS mentioned above, 100% of ESU staff developers believed public education in Nebraska improved due to STARS. Seventy-six percent (76%) of assessment coordinators, 75% of principals, and 56% of teachers agreed that public education in Nebraska improved because of STARS.

Boss et al. (2005) shared that educators from all groups strongly suggested that based on the six specific questions from the survey that STARS assessments are generally aligned to the six quality criteria. Dappen and Isernhagen (2005) found that portfolio ratings from 2002 to 2004 increased strongly at each grade level. This would seem to reflect growing knowledge and skill in assessment practices by district staff. In the primary study, alternatives to the assessment portfolio submission process have been recommended. Since the findings of Boss et al. (2005) indicated that all groups surveyed believed that STARS assessments were generally aligned to the six quality criteria, this may indeed be an appropriate consideration for the STARS process. An individual principal showed agreement by stating, "But we need to take a step in Nebraska. And that is to do that portfolio evaluation process and do it right."

Dappen and Isernhagen (2005) indicated that Nebraska's reading and math scores increased for both criterion referenced assessments and norm referenced assessments. The average percent of students meeting the defined mastery level for their locally developed Math CRT

measure increased over six percent at each grade level from 2002 to 2004. This lends strong support to continuing CRTs in the instructional activities for improvement of math achievement. In this same study, NRT scores increased three percent at the fourth and 0.5% at the eleventh grade, and declined 0.7% at the eighth grade. These scores did not appear to have been negatively affected by the implementation of the CRT program; indeed, there is some indication of slight increases. While this is initial achievement information, the data are compelling to support the continuation of the Nebraska STARS approach.

In a third study by Dappen and Isernhagen (2005), significant gains were made at the fourth and eighth grade where writing pre/post scores were available. At fourth grade, there was a 6.85% gain in the percentage of students demonstrating proficiency in writing and at the eighth grade, a gain of 6.90%. Both are significant beyond the .001 level. Anderson (2004) noted that the persons who rated the writing assessments of students perceived that their training was very positive and that they felt they were better teachers of writing because of the training opportunity suggesting that teaching of writing is improving thus student performance gains are noted.

The study conducted by Clark (2001) noted that superintendents believed that principal leadership was significantly impacted by STARS while high school teachers were not. Her conclusions concerning district size and the effect of STARS on high school teachers and instruction yielded questions for consideration in future research as to whether the effect is different at the elementary and middle school in comparison with high school. The results also showed that superintendents in large districts perceived STARS effect differently from superintendents in smaller populations. Additionally, districts receiving less money for implementation were also those who were less positive about the state system of assessment.

Isernhagen and Dappen (2005) administered a survey developed by Richard Stiggins to board presidents, superintendents, assessment coordinators, elementary principals and secondary principals that described areas of assessment practice. In growth of assessment knowledge, assessment coordinators indicated “extensive growth” (4.6) on the five-point Likert scale while elementary and secondary principals indicated “quite a bit of growth.” Superintendents were lower than principals (3.9) however not significantly. Board presidents were significantly lower than all groups (3.2) indicating “some growth.” Results of this study indicated that elementary principals were significantly stronger than all other groups in “use of achievement standards,” “effective communication,” “supportive assessment policies,” and “leadership for assessment reform.” This may be a result of elementary principals generally being more involved and affected in the early stages of the STARS process. Secondary principal ratings were the most varied, generally having the highest differences in standard deviation. The primary study found much evidence that Nebraska’s leadership is emerging with the implementation of STARS.

This year’s study reconfirmed that the language of educators in all configurations has changed. Educators consistently used their new language in appropriate ways in conversations about the STARS process. There is considerable evidence from the survey and

interviews that language arts and math teachers at grades 4, 8, and 11 in all configurations are increasingly more assessment literate and supportive of STARS each year. Without a doubt, teachers and administrators alike believe that the alignment of curriculum to standards has been one of the best things to happen to Nebraska schools in a long time. However, some resistance still remains across all configurations. In order to continue to build ownership in the STARS process, educators must be continually reminded of the big picture as shared by a STARS teacher,

We need a workshop, or a one-day or a half-day in-service on explaining how this all fits together. I'm just working with some pieces here in the corner of this puzzle and there's this master puzzle and I still don't see how all the pieces fit. I'd like to see the bigger picture and the people who work in it everyday. It's become mechanical to them.

A principal shared the need for next steps focusing on maintaining the big picture, "I think our next step is right where we're going. And that's connecting those assessments to learning, connecting...making that big picture just even stronger."

A superintendent emphasized the importance of maintaining the integrity but improving the process by stating, "I think we need to be on the lookout for those things that can be more efficient and can streamline the process. I guess I'm always looking for better ways to do things. What is the next step?"

There are both old and new challenges to be considered. New questions are being raised about retesting students after reteaching has occurred, examining standards-based grading practices, increasing requirements for graduation, and developing appropriate interventions to ensure students have multiple opportunities for standards mastery. More and more educators in all configurations have indicated that STARS is the right thing to do for Nebraska schools and children.

RECOMMENDATIONS

These recommendations remain from previous studies:

1. Continue to use multiple measures of student performance.
2. Disseminate step-by-step processes and methods for districts to improve their local assessments, especially on sufficiency.
3. Use local talent to model the building of data processes and the integration of data into school improvement.
4. Require all districts to have an assessment coordinator.
5. Continue to educate various constituencies about the nature and purposes of STARS. Focus special attention on local media and educators.
6. Continue to sponsor a "leaders of learning" academy for principals and superintendents.
7. Continue to work with the measurement community to develop appropriate metrics for the kind of data generated by STARS.

New recommendations based on 2004-05 study:

1. Continue to focus professional development offerings for all educators in working with (i.e., interpreting and using) data.
2. Research the assessment literacy knowledge and skills provided by higher education for new teachers entering the field.
3. Foster a partnership with school districts to experiment with creative uses of time and resources to support continued dialogue in development and revision of assessments for teachers.
4. Development of a statewide policy and/or policy decision on how many times a student can be retested before reporting annual assessment data to the state.
5. Consider reducing the number of standards reported in reading and math.
6. Continue to work with schools as partners in developing new technology support systems to handle data management and reporting.
7. Research the current use and success of district benchmarks statewide.
8. Research the knowledge and views of elementary, middle, and secondary teachers to identify similarities and differences about STARS.
9. Develop assessment literacy and participation in the STARS process for all teachers.
10. Work collaboratively with ESUs to encourage the development of benchmarks and assessment literacy at all grade levels.
11. Consider alternatives to the portfolio process currently used for district validation of their assessment system.
12. Research student perceptions of the impact of the implementations of standards, assessments and accountability on their achievement.
13. Research parent/community perceptions of the impact of STARS on the quality of education offered by their district.
14. Research the use of STARS data by teachers in designing and implementing differentiated instructional strategies for students.
15. Research achievement growth of sub-groups of students (Special Education students and English Language Learners, etc.).

CHARTING STARS: VOICES FROM THE FIELD

YEAR FOUR PRIMARY RESEARCH STUDY

“Leading ‘the charge for change’ takes extensive personal effort and unfailing passion for a vision that delivers results in student gains.”

National Study of School Evaluation (2004, p. 16)

Overview

With the emphasis to demonstrate increased achievement for all students, and all schools to be accountable to their constituents, many states have developed or adopted state mandated tests to assess student academic performance. Nebraska stands alone. Those in state leadership positions in Nebraska have made a conscious decision to lead the way in a new system of accountability that focuses upon building assessment literacy among educators and enhancing student performance through the use of a quality assessment system.

Within the year four study, the majority of Nebraska educators vigorously supported the vision of assessment and learning espoused by those in state leadership positions by their responses to interview questions. A secondary principal stated,

“I think that one thing I would say is that I’m a real supporter of Doug Christensen and the department in that they do allow some autonomy in the state regarding districts and the whole standards process. So they haven’t totally given in to just a state or national exam. I think that the approach we’ve been allowed to use as a district has really been more meaningful and easier to support. It’s something that makes sense for education and I’m pleased that we’ve been allowed to do it.”

A superintendent in another district showed similar support,

“They like what we’re doing. They think it’s good for kids. And STARS allows it. I’m eternally grateful for the leadership the Commissioner has given. I think he stuck his neck out to do the right thing.”

Purpose of the Study

The purpose of year four’s primary study was to determine the impact of the implementation of STARS on assessment literacy, use of data in classroom settings, leadership, and support

from external agencies on consortium districts, collaboration districts, and individual districts.

This year's primary study examined three configurations—collaborations, consortia, and individual districts—to determine the impact of STARS. Nebraska was comprised of 495 school districts when the Fourth Year study began. Of the 495 school districts, 72 were participating in collaborations, 289 were participating in consortia and, 134 remained as individual districts for the implementation of STARS.

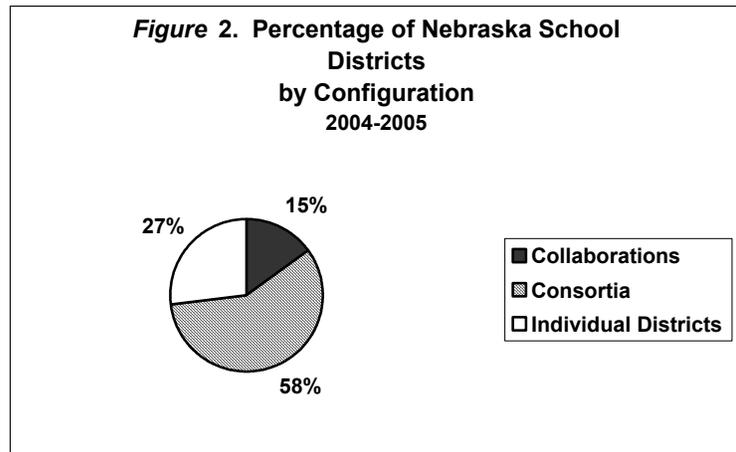
The study was approved by the Institutional Review Board (IRB) at the University of Nebraska-Lincoln in order to maintain confidentiality and protect the privacy of participants in the study (Appendix I).

Research Design

A mixed-methods design was selected for use to strengthen the study results. The study began with a broad survey in order to generalize results to a large statewide population of educators at all levels. As a second step, open-ended interviews collected detailed views from a purposeful sample of educators belonging to or serving districts in collaborations, consortia, and individual school districts. This selected group of participants was interviewed to help uncover, confirm, or qualify the basic findings from the survey.

Survey Sample

Of the 495 total school districts in the state, 72 (15%) of all school districts were participating in collaborations, 289 (58%) were in consortia, and 134 (27%) were individual districts. Selection of sample school districts for the survey was based on a stratified purposeful sample using district class (Classes I-VI), free and reduced district lunch rate (high, middle, low based on statewide average), and geographical areas (East, Central and West). The survey sample included districts in 22 collaborations, 130 consortia, and 78 individual school districts (four additional individual districts were added to balance class distribution) based on a proportionate representation of the total. Researchers doubled the number of school districts in each configuration to obtain an adequate number of surveys for statistical analysis.



Interview Sample

From the survey sample, we selected nine districts in the East, seven districts in the Central, and eight districts in the West. Sample districts that had been used previously and identified by purposeful sample in 2001, were used first and then other schools were selected based on sampling criteria. Interviews were also conducted in seven ESUs across the state.

Instruments

The STARS survey was designed by the researchers to address recommendations based on previous studies conducted about the STARS process to research the effects on schools of using assessments designed elsewhere, including consortia, ESUs, and districts; research the kinds and quality of support schools receive from various outside sources including ESUs, consortia, and the NDE; and target assessment literacy opportunities for teachers of grades 8 and 11. The STARS survey used for teachers, school and district administrators, assessment coordinators and educational service unit staff developers of districts in all configurations collected demographic information and responses on five major categories (Appendix J). These categories were (1) District Support, (2) Assessment Literacy, (3) Data, (4) Instructional Impact, and (5) External Support. Participants responded to the 46-item survey on a five-point Likert format 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 for all groups.

The STARS Research Interview Protocol (Appendix K) consisted of demographic information and six major questions. These questions targeted the participants’ role in the assessment process, new learnings from their involvement, type and use of data used to improve student performance, major obstacles faced in the implementation, next steps for better implementation, and the role of their configuration in the STARS 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

Survey

During Phase One of the Fourth Year study, 1722 participant surveys were distributed by mail to school districts across the state. Of the 1722 surveys mailed, 869 completed surveys were returned representing a 50% overall return rate. Of the 22 collaboration districts included in the study, 16 collaboration districts (73%) returned surveys; of the 130 consortium districts, 85 consortia districts (65%) returned surveys; and of the 78 individual districts, 47 individual districts (60%) returned surveys.

Table 2: STARS District Return Rate by Configuration 2004-2005

Configurations	No. of Districts Surveyed		No. of Districts Returning Surveys		District Return Rate
Survey Category	Collaborations		Consortia		Individual Districts
Collaboration Districts	22		16		73%
Consortium Districts	130	4.3	85	4.5	65%
Individual Districts	78	3.7	47	3.9	60%
Data		3.6		3.9	4.1
Instructional Impact		3.9		3.9	4.0
External Support		3.5		3.6	3.7

Results
from

the surveys indicated strong support for the STARS Process in all three configurations. The average mean on a one to five-point Likert scale for all items for each survey category by configuration was:

Table 3: Average Means for Survey Categories by Configuration 2004-2005

These results indicated strong support for assessment literacy, use of data, instructional impact, and external support for all three configurations.

Interviews

Interviews were conducted in 24 school districts. Five (5) districts were in collaborations, 12 districts were in consortia, and 7 were individual districts totaling 169 interviews. From the interviews conducted, 12 districts were selected for coding. The interviews supported the five themes examined on the survey and five additional themes emerged. The ten themes were: Total Support, Assessment Literacy, Data, Instructional Impact, External Support, Accountability, Leadership, Professional Development, Challenges and Obstacles to Success, and Next Steps for Success.

These themes helped to paint a picture of what a collaboration, consortium, and individual district looked like as they grew in their knowledge of assessment, collection and interpretation of data, and application of their new learnings in instruction.

CHARTING STARS: VOICES FROM THE FIELD

ASPIRING FOR EXCELLENCE "CONFIGURATIONS"

“As long as you believe what you’re doing is meaningful, you can cut through fear and exhaustion and take the next step.”

Kouzes & Posner, Leadership: The Challenge (2002, p. 184)

For the last five years, Nebraska educators have created new learning through professional development and dialogue, developed and refined assessments, collected and analyzed data, and implemented best instructional practices with the support of external agencies. During this process, educators in school districts began to share and dialogue with others in the state to identify better ways of delivering STARS. As a result of this dialogue, three configurations of delivery emerged—collaboration districts, consortium districts, and individual districts. School districts made a decision as to the type of configuration they wanted to align their district with to better implement the STARS process. Outlined below is each configuration as described by the “voices from the field.”

Collaboration

The Nebraska Department of Education defines a collaboration as “A group of districts that work together to develop and/or share individual assessments and submit individual assessment portfolios to the NDE. All portfolios for districts participating in a collaboration are sent to the same reviewer.” The researchers found that the voices from the field defined collaborations in a much more personal way. The following quotes from interviews across the state expressed teacher, principal, and ESU staff developers’ perceptions of a collaboration.

The value of belonging to a *collaboration* was described by *teachers* as:

- ❖ The ESU collaboration has given us that wonderful opportunity—dialogue between and among other language arts educators. This has been a tremendous benefit of this educational directive.
- ❖ The collaboration that our district has participated in has focused on the development of quality assessments. The curriculum work has been done at an individual district level. Our collaboration efforts have been very beneficial to us.

- ❖ The value that our staff and I gain from being in the collaboration is that we don't have to invent the wheel all over again.

As the STARS process evolved, *teachers* and ESU *staff developers* felt that the *collaboration* helped them to understand the process as evidenced by the following quotes:

- ❖ The STARS process is a success. Getting curriculum aligned and having teachers examine the kinds of topics and questions at every level are very beneficial.
- ❖ This whole process makes you look at your data and makes you do something with it.
- ❖ We work completely as a team. It has worked out very well because as we sit down to score these standards, we all collaborate on exactly how they should be scored so it is accurate and equal for each student within the building. Then, as we work with these standards, we can also then discuss where students are failing. . . if it's the test, if it's our lessons.

The learning that occurred and was shared at the school level as staff members continued to work together was characterized by a *collaboration teacher* as:

- ❖ Probably the best learning for our staff is that we are being able to communicate on the same level. We know that there is some essential education that has to be taught and our teachers have bought into that. I think it's allowed us to have more professional dialogue, not only what needs to be taught, but what are the best practices to teach that? And it's forced educators to have a common language.

The impact of the STARS process was shared by a *collaboration secondary teacher* as:

- ❖ I think it's made me a better teacher. And I think after teaching for a number of years, we tend to get in a rut and sometimes think we've got it all down, we know what we're doing. I think we need to be reminded that there's always room for us to learn and we have new groups of kids coming in and things change from time to time, and we need to be willing to change with those.

Administrators' roles in the STARS process have evolved along with the work in the *collaboration*. Instructional leadership has become an essential element in all configurations.

As a secondary *principal* stated:

- ❖ My role was to assist the teachers in the interpretation of what needs to be included, to follow the criteria for the development of the assessments. I think just to assist the teachers in the creation of the assessments and doing what was intended to do.

Consortium

The NDE defines consortium as “a group of districts that work the STARS process together, use all the same assessments and submit a single assessment portfolio to NDE.” In addition, each school district is required to submit information verifying how they have met Quality Indicator 2, providing an opportunity for students to learn the content. Belonging to a consortium allows a district to select from one of three models--Online consortium,

Panhandle consortium, and “Other” consortia. Within the description provided by teachers and administrators, all types of consortia are noted.

One *teacher* from a small school indicated that “Our *consortium* is a life saver. Our small school would be lost without it.” Several teachers indicated the value of belonging to a consortium:

- ❖ I think it’s a good system and it seems like it saves a lot of paperwork and a lot of time on the teachers’ part and I think that’s a benefit. I think it benefits the kids, too, because they’re being held a little more accountable for what they learn. And even the teachers, too, because we’re held accountable for what the students learn.
- ❖ I think just being around other teachers and finding out what they’re doing in the classroom and how they’re approaching the standards. I think that’s probably been the biggest value because everybody’s in the same boat.

Teachers from the Online *consortium* stated:

- ❖ One nice thing about our Online assessments is that students get immediate feedback, so they know right after the test how well they did.
- ❖ But with this consortium process, with the use of computers, we’ll get to where we want ten years from now, which is making sure that every child has every opportunity to become proficient.
- ❖ Well, the scoring part, I mean that’s all that ridiculous paperwork and all of that is basically done. I mean it’s just all kept track of in the online assessment program, which is what I really like about it.

A *consortia teacher* and *administrator* stated:

- ❖ The most difficult and frustrating aspect of this entire process of standards and portfolios for smaller schools has been (at least for ours) the statistical process required in Criteria 5 and 6. By joining a consortium, our school was provided a better avenue for reporting reliability.
- ❖ A lot of times it was having our teachers working together, so they could work together as a team, so that way they’re not out there alone which is a pretty lonely feeling when you’re trying to do this on your own, so it was more of a team effort to get together to share ideas.

Teachers in *consortia* indicated that belonging to a consortium helped them to understand the STARS process and they worked cooperatively as indicated by these quotes:

- ❖ We develop our assessments as a group. We determined their validity, bias, etc., as a group. We aligned our curriculums as a group. We still have our own local standards and curriculum. We make our own local decisions for instruction based on our consortium test results to modify our curriculum and instructional strategies.
- ❖ The consortium has been responsible for developing the assessments that we use to measure student progress toward the standards. They have been very beneficial in helping us with our assessment portfolio in the sense of development. They

wrote the portfolio for those parts that were consortium issues and that only left us with our individual things like, how do we monitor that this kid's had the opportunity to learn before the assessment.

- ❖ The assessments we use are given and scored by each teacher. We are placed in pods so the whole consortium does not use the same assessments. The teachers of each pod meet in the summers to revise the assessments.
- ❖ Teachers from our school go to a retreat in the summer, where we write and rewrite questions. Those questions are then put into a bank for our consortium to use. Do I feel that I have personally written each question? NO. Do I feel like I have control over the questions put into the bank? YES.

Teachers in consortia shared that consortia help with data collection and analysis:

- ❖ And now I feel like it's streamlined, it's what I want. I help make up the tests and all the data is gathered and kept track of. It's just great.
- ❖ They're helping us to record the data, know what to do with the data, fine tune the questions, giving us the test itself and making sure we have all the materials that go with it, a support system if we do need help and questions.
- ❖ They break it (data) all down and give it to us on a spreadsheet and it's however we want it. If we want it by gender, socioeconomic, whatever, it's broken down and really given to us in a pretty neat little package.

Individual Districts

“A district that develops its own assessment system and submits independently its District Assessment Portfolio to the NDE,” is the definition that the Department of Education uses when defining the individual district configuration. Teachers and administrators across the state who participate in an individual district gave voice to their perceptions.

The value of belonging to an **individual district** can be found in these statements from **teachers** and an **assessment coordinator**:

- ❖ I think the value in our own district doing this rather than a consortium is that we can assess what we actually teach. Otherwise, I know like with other configurations, they come up with some questions and you may have to add something to the curriculum to cover something that they have come up with or you may have to take some things out in order to make those adjustments. And I think that doing it district-wide gives us a little more local control and you're more able to teach, or to test on specifically what you teach.
- ❖ I think because its so teacher involved and teacher-driven, I think that's the value of it and it relies on teachers, not just individual teachers but teacher collaboration.

Teachers and administrators in individual districts indicated that this configuration provides a great deal of flexibility as evidenced by these comments:

- ❖ But having the time and the freedom to change and adapt those (assessments) as we needed to over the years is definitely a huge benefit, even though we had to work quickly at first.
- ❖ That's one thing in our district that we've always had is flexibility and that's one reason I think that our STARS are so successful is that they've really given control to the teachers. And I think that's because they believe in us and know that what we do is important.

Individual district teachers and **assessment coordinators** valued the opportunity to be involved in assessment development as evidenced by these statements:

- ❖ As a district we did set up assessments and made sure that our curriculum was aligned with the standards and developed the assessments for each grade level.
- ❖ I feel that I've had at least a say in what goes into the assessments. I also have been able to talk to the other math teachers so I make sure they have a voice in it, too, which I feel is very important.
- ❖ We learn every day. We learn from kids. We learn from each other. We learn from the STARS process. It's ongoing.

A **principal** in an **individual district** shared how he/she viewed data and the impact that it had on decision making as evidenced by this statement:

- ❖ Well, our data retreat is a prime example. We brought our school improvement teams together. We looked at our district data, our historical data from our assessments—that data was a primary source—and then from there identified our strengths, our weaknesses, what were our common themes? And then from there, our action plans were developed and our building goals, which then led to our classroom goals which also fed into our curriculum and best practices.

Teacher ownership is an important aspect of an **individual district** as evidenced by these statements made by a secondary **teacher** and an **administrator**:

- ❖ I feel like there's a lot more ownership to it so when I'm teaching whatever the concepts are with the kids, I already know ahead of time what the intent is of the test, where the focus is and not just reading a 20-word standard and then trying from there to figure out what's going to be on the test.
- ❖ The local control for the development of standards and assessments is very empowering. Teachers have strong ownership in the final product. The assessment results are very useful for local school improvement.

Teachers and **administrators** defined the work of leaders in **individual districts** in the STARS process as:

- ❖ I think if your district is going to be effective, then I think that you have to have the central office leading the charge in that because, otherwise, it falls back to each individual principal and what their knowledge and skills are.
- ❖ I feel as if my voice matters and that what I feel is best for students is being brought up in meetings and I'm able to discuss my opinions with other teachers

and we can bounce ideas off one another. I think it helps with communication throughout the district when you have a variety of leaders from various buildings and they're able to talk about what's happening in their building as opposed to my building.

CHARTING STARS: VOICES FROM THE FIELD

THEME 1: TOTAL SUPPORT FOR THE STARS PROCESS

I'm very proud of Nebraska and our assessment system. I'm glad to be a part of it. I talked to other teachers from other states. They're amazed at what we do. I think it's a good thing. I think more people need to know that and to understand how very lucky we are to be able to teach in Nebraska and the ways that we have the assessment system available to us.

Interview Quote from a Nebraska Teacher (2005)

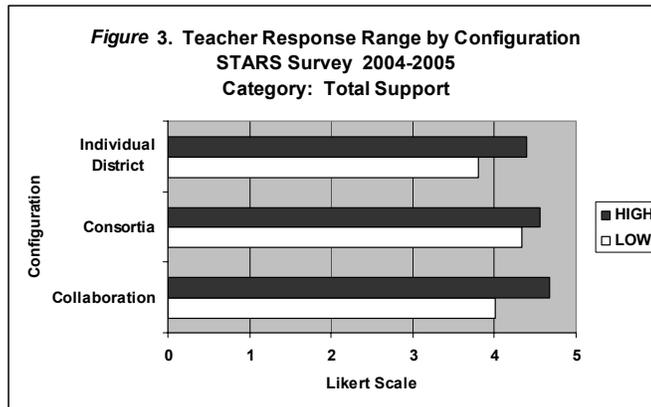
OVERVIEW

This section will highlight the statistically significant or noteworthy findings from the first section of the survey—Total Support for the STARS Process. Surveys were administered to language arts and math teachers in grades 4, 8, and 11, principals, superintendents, and assessment coordinators from districts in collaborations, consortia, and individual districts, and ESU staff developers serving those districts. The significant findings will be supported with quotes that were collected during the interviews for those items where quotes were available.

TEACHER SURVEY RESPONSES

Configuration Support (Questions 8-11)

Teacher response rate ranged from 3.80 to 4.67 for all configurations for total support as shown below:



In the area of total support, the survey responses of teachers in collaboration districts ranged from 4.01 to 4.67. There was a strong perception by teachers regarding collaboration support for providing on-going assessment training for teachers as noted in these quotes from the interviews.

A secondary *teacher* in a *collaboration* stated:

- ❖ I guess part of my new learning was the training I first took just to learn what a standard was, how to do benchmarks, how to write a good assessment, so I've had a lot of learning and just a lot of terminology that came along. It was kind of new for Nebraska teachers.

Elementary *teachers* in *collaborations* particularly liked the opportunity to dialogue with others as they worked on assessments. They stated:

- ❖ It's critical to bring people together to discuss, because I found out as an educator the best ideas you'll get comes from others.
- ❖ We are able to get together with more than just our school and we're working on assessments and those kind of things which gives you a broader base of teachers to work with, more ideas to work with. People aren't out there on their own. It's been very good because our people have been able to collaborate with other schools and other people.

The issue of teacher ownership of the STARS process was addressed by many *collaboration teachers* in the interviews as cited in the following teacher quote:

- ❖ I just think teachers need to buy into the process and get on board and use it and be more conscientious of it. And I don't know that the collaboration can do that for an individual.

Overall *teachers* in *collaborations* rated their own training primarily provided by ESUs as excellent. They indicated that the main value of their training was increased knowledge about the development of assessments and the opportunity to work with others.

In the area of total support, teacher responses in consortia districts ranged from 4.34 to 4.56. There was a strong perception of teachers regarding consortia support for providing on-going assessment training for teachers.

Teachers in consortia stated:

- ❖ I really like that they're experts that know about bias and can help me with that. And there are experts there talking to us about readability and helping us make up the tests so they're reliable. They really walk us through the process.
- ❖ Our move toward consortium-designed assessment diminishes the enormous burden of test design, enhances professional discussion among teachers, and improves comparability of scores.

The issue of teacher ownership of the STARS process was addressed by a *teacher* in a *consortium* during the interviews:

- ❖ I think that's where our consortium's having the most difficulties. People buying into this is the test we're going to use.

In the area of total support, teacher responses in individual districts ranged from 3.80 to 4.39. Teachers in individual districts' perception rated lower than consortia teachers but higher than collaboration teachers for providing on-going assessment training for school administrators. Teachers in individual districts rated on-going assessment training for teachers as the lowest of the three configurations.

Teachers in individual districts shared their thoughts:

- ❖ The district was involved by providing workshops and staff development for teachers and for our curriculum coordinator, to make sure that they had a strong understanding what was expected from the state department in establishing our own assessments for our curriculum.
- ❖ Well, I think that our district is doing a great job of taking STARS assessments and listening to the feedback of teachers and making changes that are needed.

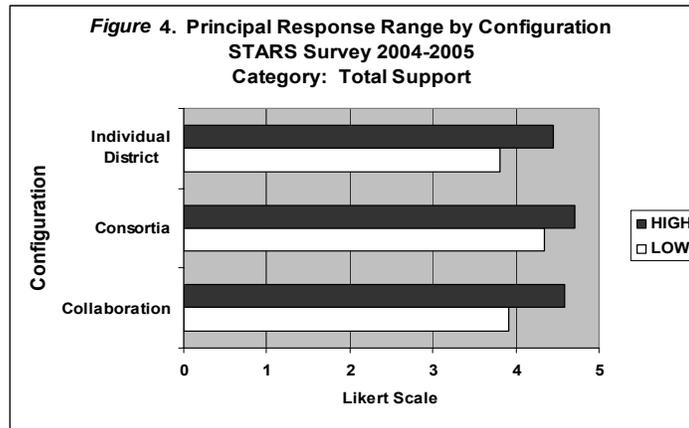
There is a need to adapt to change in order to implement the STARS process successfully. Two *teachers* in *individual districts* shared these viewpoints about change:

- ❖ Any time you change, people are resistant to change. Teachers were upset about 'Why do we have to do this?' It's just another thing from the state.
- ❖ Myself, I feel very comfortable with what we're doing. It's part of my classroom. It's part of the curriculum. It flows very easily.

PRINCIPAL SURVEY RESPONSES

Configuration Support (Questions 8-11)

Principal response rate ranged from 3.81 to 4.71 for total support for all configurations as shown below:



In the area of total support, principal responses in collaboration districts ranged from 3.91 to 4.58. Principal perceptions in collaborations rated higher than principal perceptions in individual districts, but lower than principals in consortia districts, regarding district support for “school-based teacher-led” assessment, providing services that schools need for the implementation of STARS, and providing on-going assessment training for school administrators.

Collaboration principals indicated that it was critical to involve teachers in all steps of the STARS process. One secondary principal stated:

- ❖ One is to have a good understanding of what the state is asking for and to be able to communicate that to our staff. I think another one is to work with the staff and encourage them, and try to explain and provide some direction in interpreting what the state wants. I think also to stress to the faculty that it’s important to have these discussions. I think that’s been the benefit of the STARS process.

School administrators indicated that professional assessment training was available and that their knowledge and skills about the STARS process was growing. Two elementary *principals* in *collaborations* indicated:

- ❖ I think we still need to continue to put that knowledge there so that people can gain from it. We need to have great leadership all the way through and expectations, high expectations, of everyone and not leave it on just a few people.
- ❖ I’m very pleased to have the knowledge base that I have because I think it helps me make better decisions for kids.

The building of teacher ownership is critical to leading the STARS Process. A *collaboration principal* expressed his/her role in facilitating that ownership with this statement:

- ❖ I think when the teachers take ownership of it, they see the data, they know where it’s going to go, and they know where they want it to go. They’re able to implement it so much better when they believe it and buy into it than just the principals saying this is where we’re going.

In the area of total support, principal responses in consortia districts ranged from 4.35 to 4.71. Principal perceptions in consortia rated the strongest regarding consortia support for “school-based teacher-led” assessment, providing services that schools need for the implementation of STARS, and providing on-going assessment training for school administrators.

Consortia support for “school-based teacher-led” assessment was evident in the following comment made by a secondary *principal*:

- ❖ A lot of it was getting the teachers together and collaborating a little bit in making the assessments, in making the standards, making sure that we had schedules set for the assessments. A lot of times it was having our teachers working together, so they could work together as a team. So that way they’re not out there alone, which is a pretty lonely feeling when you’re trying to do this on your own.

A *principal* indicated that *consortium* teachers have taken ownership for the implementation of the STARS process as noted by this statement:

- ❖ The best part of the STARS process was aligning the curriculum. Everyone is on the same page as to what is being taught and where.

In the area of total support, principal responses in individual districts ranged from 3.81 to 4.45. Principal perceptions in individual districts rated the lowest on district support for “school-based teacher-led” assessment, providing services that schools need for the implementation of STARS, and providing on-going assessment training for school administrators.

Principals in *individual districts* indicated that teachers are becoming more comfortable with the STARS process:

- ❖ They know what they have to do and they’re becoming more and more comfortable with it. Any time you are imposed to do something and have to be trained and implement change to do something that transition takes some time. They’re more at ease than what they were a couple years ago, which is good.
- ❖ I think at the beginning it was just moving across that giant hurdle of convincing people that our process was actually going to be meaningful and give us positive results as opposed to it just being a whole extra series and layer of work.

A *principal* indicated that services needed to implement STARS were provided by his/her *individual district*. The following quote emphasized these services:

- ❖ I think I have the support of the board and my superintendent to do what’s needed to get scores where they have to be. And if that is through staff development and training and implementing programs, I have not been turned down with a no to spend the necessary money to get that done. Or to have the teachers take the time off to get that done. So right now, I feel I’m receiving 100 percent support.

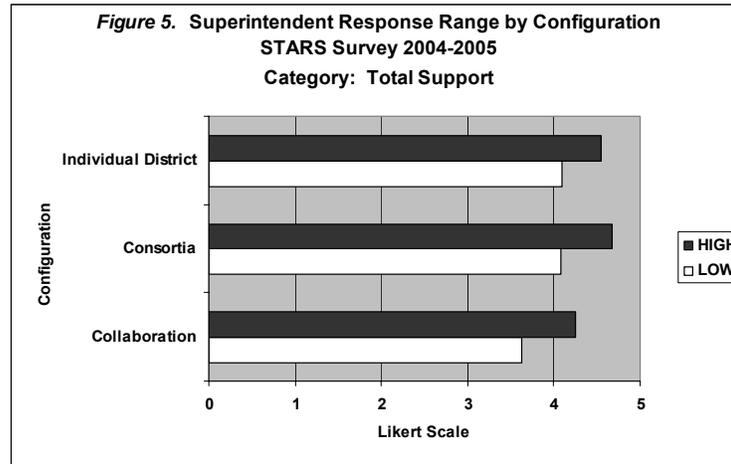
Principals addressed on-going assessment training for both teachers and administrators in *individual districts*:

- ❖ I just like it that they have continued to keep this whole issue on the front burner for us. And they've provided some excellent staff development and administrator development. So they're really training us to keep this whole system running smoothly. So there's a learning curve that goes on there at the beginning.
- ❖ I don't like to work in isolation and so it was a very collaborative experience. So for me to spend time with other professionals, teachers, administrators, specialists in the staff development area, all of us to pool our collective thoughts and expertise, you just learn and you grow.

SUPERINTENDENT SURVEY RESPONSES

Configuration Support (Questions 8-11)

Superintendent response rate ranged from 3.63 to 4.68 for total support for all configurations as indicated below:



In the area of total support, superintendent responses in collaboration districts ranged from 3.63 to 4.25. Superintendent perceptions in collaborations rated higher than superintendent perceptions in individual districts, but lower than superintendents in consortia districts, regarding district support for “school-based teacher-led” assessment and providing on-going assessment training for teachers. Superintendent perceptions in collaborations rated the lowest in providing services that schools need for the implementation of STARS and providing on-going assessment training for administrators.

A *collaboration superintendent* stated that his/her role was to provide what the schools needed for the implementation of STARS:

- ❖ I think right now my role is to support those people with time and money and to be able to do the assessments that they need.

A *superintendent* of a district in a *collaboration* said that on-going assessment training for administrators is critical to the success of the collaborations:

- ❖ (speaking of in-services) ... and it’s critical to bring people together to discuss because I found out as an educator the best ideas you’ll get from others. I haven’t come up with too many breakthroughs, but I’ve learned a lot from research and other things and I think it’s important to get together and visit.

In the area of total support, superintendent responses in consortia districts ranged from 4.08 to 4.68. Superintendent perceptions in consortia rated the strongest in district support for “school-based teacher-led” assessment and providing on-going assessment training for teachers. Consortia superintendent perceptions are stronger than

collaboration superintendents, but less than individual district superintendents, in providing services that schools need for the implementation of STARS and for on-going assessment training for administrators.

A *superintendent* of a *consortia* district stated that support for “school-based teacher-led” assessment is critical to the success of the consortium:

- ❖ I think the whole STARS process is very, very important. It’s probably been one of the best things that ever has happened to education. Even though the flak that we’re catching, it’s good because teachers are spending more time looking at the curriculum, being more concerned about what kids are learning in their classroom, and also empowering them.

In the area of total support, superintendent responses in individual districts ranged from 4.09 to 4.55. Superintendent perceptions in individual districts rated the strongest in providing services that schools need for the implementation of STARS and for on-going assessment training for administrators. Individual district superintendent perceptions rated the lowest on district support for “school-based teacher-led” assessment and providing on-going assessment training for teachers.

Superintendents of individual districts said that support for “school-based teacher-led” assessment is critical to the success of their districts:

- ❖ With the implementation of the STARS process, educators are incorporating data with instruction. Communication about student achievement and learning is at an all time high. Collaboration among teachers has increased. Assessments have greatly improved and their importance has increased. Linking the STARS process to school improvement was a significant step to the improvement for student learning. The time and effort is well worth the outcomes.
- ❖ I think as they’ve gotten into it, they realize that hey, this not only has held our students accountable, but it’s held us accountable to saying, okay, these are our course objectives, these are how these course objectives align with the STARS assessment, so we better be assessing our objective properly.

A *superintendent* in an *individual district* stated that ongoing assessment training for administrators is critical to the success of his/her district:

- ❖ We meet with our school improvement teams—every team in the district—once a month. And we go through a workshop with them. And during that workshop, we prepare materials so that they can go back and involve their entire staff. So every time a principal and school improvement team gets information and activities, every teacher in the district also gets the same thing. And it’s an involvement activity.

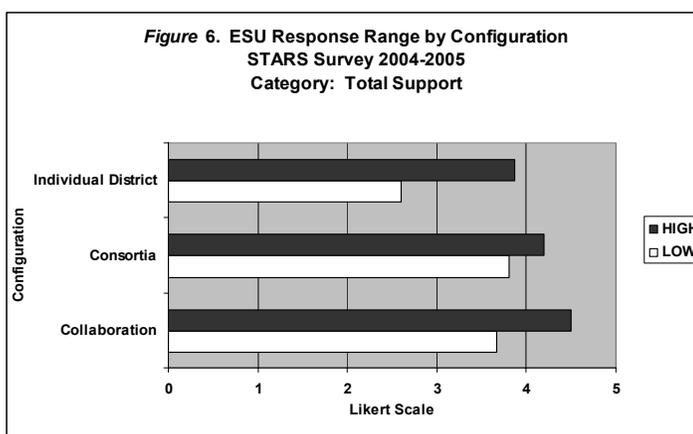
A *superintendent* in an *individual district* shared that his/her role was to provide what the schools needed for the implementation of STARS:

- ❖ We developed a process that fits with our beliefs about assessment. And the great thing about STARS is it allows you to carry out your own beliefs. I think the idea of involvement is motivating people and interesting people to do the right thing for kids.

ESU STAFF DEVELOPERS SURVEY RESPONSES

Configuration Support (Questions 8-11)

ESU response rate ranged from 2.60 to 4.50 for all configurations for total support as indicated below:



The ESU responses about collaboration districts ranged from 3.67 to 4.50. The range for consortia districts was from 3.81 to 4.19 and for individual districts from 2.60 to 3.87. There were statistically significant differences in the ESU perceptions for total support.

Collaboration districts rated strongest in providing services that schools need for the implementation of STARS and providing on-going assessment training for school administrators and teachers. Consortia districts and individual districts rated lowest in these three areas.

ESU staff developers' perceptions in the area of total support rated *collaboration* districts strongest in providing services that schools need for implementation of STARS:

- ❖ Our primary role has been a facilitator of the design and refinement process. And what that's entailed has been everything from ourselves being trained and getting a lot of information to bring back to our schools as well as to take all of that information, synthesize it, and create a process that we think we can take teachers through at all levels of their literacy. We really facilitate the process of both the development and the refinement of criterion referenced assessments.

Providing on-going assessment training for school administrators and teachers was rated the strongest for *collaboration* districts:

- ❖ Listening to the conversations that teachers have with other teachers from the other districts is, by far, the most exhilarating portion of the entire process. To hear them have conversations about real kids and what kids can really be able to do, and maybe—here’s a comment—maybe I’m not challenging mine enough, I didn’t think they could do that. It’s powerful for them to have those really professional conversations.
- ❖ We always have sessions. We usually have five sessions every spring for people who are writing portfolios. And they just come in here and work and we go through the things we’ve learned from Pat Roschewski and her group, and we’ve had Buors in every year.

ESU staff developers rated *consortia* and *individual districts* the lowest in all three areas of statistical significance:

- ❖ They have not asked for direct help with their assessments, but we work with their curriculum director a lot on process and procedures; she gets to some of our trainings that we go to as a group, and then one-on-one phone conversations with her.

ASSESSMENT COORDINATOR SURVEY RESPONSES

Assessment coordinators responding to the survey numbered less than 30 as a group. Therefore, they were not included in the responses for the total group and configuration reports for themes one through five.

CHARTING STARS: VOICES FROM THE FIELD

THEME 2: ASSESSMENT LITERACY

We've got to continue to convince teachers that designing your own assessments, even as hard as it is, is so much better than having somebody in Boston design an assessment that every 4th grader in Nebraska takes. That's easy and simple, takes the pressure off the teachers, but doesn't help kids; it doesn't help teachers (either).

Interview Quote from a Nebraska Superintendent (2005)

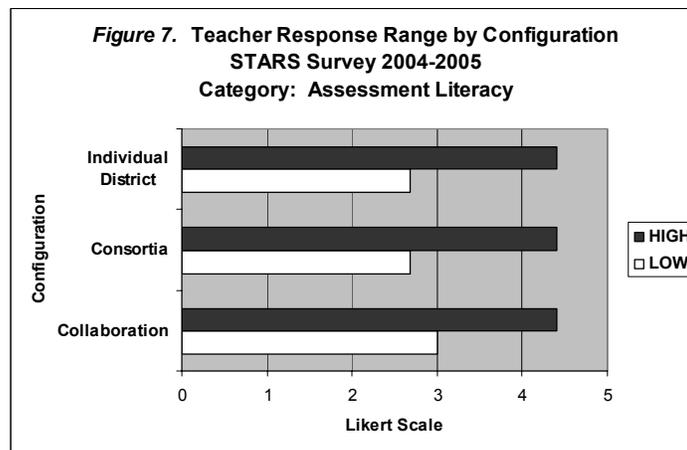
OVERVIEW

This section highlights the statistically significant or noteworthy findings from the second section of the survey, Assessment Literacy. The findings will be supported with quotes that were collected during the interviews for those items where quotes were available. Surveys were administered to 4th, 8th, and 11th grade language arts and math teachers, principals, superintendents, and assessment coordinators from districts in collaborations, consortia, and individual districts, and ESU staff developers serving those districts.

TEACHER SURVEY RESPONSES

Assessment Literacy (Questions 1-16)

Teacher response rate ranged from 2.68 to 4.41 for all configurations for assessment literacy as indicated below:



In the area of assessment literacy, teacher responses in collaboration districts ranged from 3.00 to 4.40. Perceptions of collaboration teachers rated higher than individual teachers and lower than consortia teachers on selecting assessment items from a common bank/pool provided by the district. Teachers in collaborations rated the lowest in developing the STARS assessment portfolio sent to the NDE and using the same grade level assessments to meet STARS requirements. Teachers in collaborations rated lower than individual teachers for writing their own STARS assessment at the classroom level and higher than consortia teachers.

Elementary and secondary *teachers* in *collaborations* indicated that they were writing their own assessments with the assistance of teachers from other school districts and their ESU rather than using assessments from a common bank/pool provided by their collaboration by stating:

- ❖ I've gone to workshops to learn the process of creating assessments, from unpacking (the standards) to writing to revising and revising again. And then to allow the students opportunities to learn. And then administering the tests back here at the school.
- ❖ A classroom teacher, especially an elementary classroom teacher who's responsible for all content areas, needs to be able to multi-task their assessments and get multiple scores out of one assessment that incorporates all of those standards. And that would require some rewriting of assessments, which takes a lot of time. . .having all your materials together. But it makes the most sense because you could get more done in one place. There are lots of natural connections. You just have to be able to put that all together in writing and have it all part of the process.

Teachers in *collaborations* stated that they use the same grade level assessments as indicated in these teacher responses:

- ❖ . . . our school worked with the ESU and other schools in our area to work together to form the same assessment. We've tested those over the last year, come back and refined those, and I think the goal was that the language arts assessments would be pretty well done after this collaboration period a year ago.
- ❖ All of the schools in the collaboration have sent representatives, differing in number or whatever, but sending the representatives to the workshops every summer when we wrote them, and then when we've. . .met through the year at different times improving them, or we work on making sure that they meet all of the criteria that assessments are supposed to have.
- ❖ We've developed a pod through our ESU. And through that group, there's five or six schools that are collaborating together to straighten up our assessments and revise them, and make them work for our students for all the schools in our group.

A *teacher* in a *collaboration* provided the following thought about writing the STARS Portfolio:

- ❖ I do not understand the necessity for all that goes into the portfolio. Aren't we supposed to be teaching and assessing students on standards? Too many teachers are concerned about the portfolio being accepted, it takes away time that could be better spent teaching.

In the area of assessment literacy, teacher responses in consortia districts ranged from 2.68 to 4.41. Consortia teacher perceptions rated the highest in selecting assessment items from a common bank/pool provided by the district, developing the STARS assessment portfolio sent to the NDE, and using the same grade level assessments to meet STARS requirements. Teachers in consortia rated the lowest on writing their own STARS assessments at the classroom level.

Some *teachers* from *consortia* said that they used assessment items selected from a common bank/pool while others stated that they develop their assessments with the ESU but use the same assessments:

- ❖ We participated in the writing of some of the questions on the STARS test through the education service unit and we have given the STARS test for as many years as it's been available.
- ❖ Well, I'm very involved and I really feel a part of the whole assessment process. We get together as teachers from the districts in our area and make up tests and, you know, figure the cut scores and I just feel involved and really comfortable with the whole process that we go through.

One *teacher* from a *consortium* stated that they decided to write their own assessments so that the assessing could be spread across grade levels at the high school level by stating:

- ❖ However, we did, on our own time in the next summer, develop assessments. We decided it was better to spread out the tests, etc., in the required classes of the 9th, 10th, and 11th grade rather than give them all in one year. All of my assessments require rubrics and are much more difficult to administer and evaluate. I think they still need refinement so that anyone could walk into my job and utilize them.

Teachers in *consortia* shared the following thoughts about developing the STARS assessment portfolio sent to the NDE:

- ❖ While stressful to the extreme, we have done well on our portfolios due to the dedication of our teachers!
- ❖ The portfolio is good because it holds you accountable and you have a record of what you're doing. But I think we could have used our time a little more wisely as far as actually teaching the kids.

In the area of assessment literacy, teacher responses in individual districts ranged from 2.69 to 4.41 on the five-point Likert scale. Individual teacher perceptions rated highest for writing their own STARS assessments at the classroom level. They rated the lowest for selecting assessment items from a common bank/pool provided by the district.

Teachers in individual districts rated lower than consortia, but higher than collaboration teachers, on developing the STARS assessment portfolio sent to the NDE and using the same grade level assessments to meet STARS requirements.

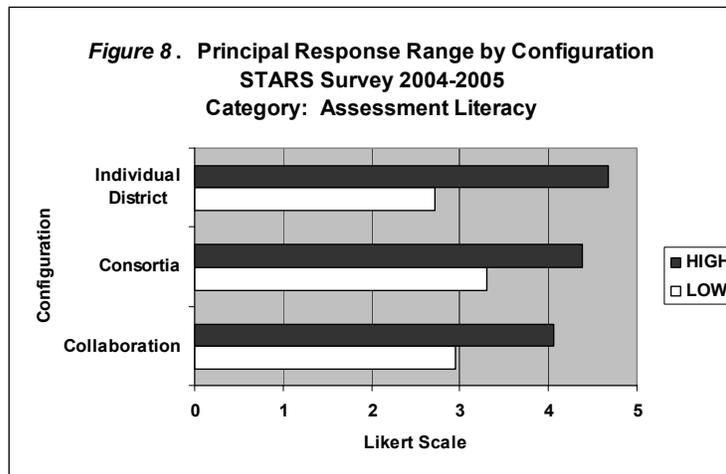
Individual teachers stated that they write their own assessments rather than select them from a common pool as noted by these comments:

- ❖ Well, one thing that we tried to do when they decided to set up the assessments was to see what we were already doing and what assessments we had in place or what activities we had in place that could act as assessments.
- ❖ I said over time it's gotten easier and we've gotten better at looking at what needs to be changed and revised. I'm probably sounding redundant but I've learned about what are good question techniques, watching for the biases and just knowing what a good assessment looks like.

PRINCIPAL SURVEY RESPONSES

Assessment Literacy (Questions 1-16)

Principal response rate ranged from 2.72 to 4.67 for all configurations for assessment literacy as shown below:



In the area of assessment literacy, principal responses in collaboration districts ranged from 2.94 to 4.06. Perceptions of collaboration principals rated lowest for students in schools being involved in understanding their own progress and achievement status; selecting assessment items from a common bank/pool provided by the district; for district development of standards based assessments for the schools; for districts developing the STARS assessment portfolio; and for all schools using the same grade level assessments to meet STARS requirements.

Principals in collaborations said that they are working on the development of assessment literacy and writing their own assessments:

- ❖ I would say, bringing everybody on board with that assessment literacy piece, that it's not a state requirement; it's what our school is going to do for kids. And kind

of that mindset or that philosophical piece that assessments are important and that we're doing this to make us better has been a challenge to just bring people along with that.

- ❖ We're working together as a group and that's what needs to be. We don't have staff that's available to work on this full-time so you need to have little groups that are working at a little bit here and a little bit there, and then you put it all together and you have a good assessment.

A *principal* in a *collaboration* district made this statement regarding developing the STARS assessment portfolio:

- ❖ I'm also concerned about the portfolio. Way too difficult. We are not trained to be testing proficient.

In the area of assessment literacy, principal responses in consortia districts ranged from 3.31 to 4.39. Consortia principal perceptions rated the strongest in students in schools being involved in understanding their own progress and achievement status; selecting assessment items from a common bank/pool provided by the district; and district development of standards based assessments for the schools. Principal perceptions rated higher than principals in collaboration, but lower than principals in individual districts, for developing the STARS assessment portfolio sent to the NDE and all schools using the same grade level assessments to meet STARS requirements.

Principals in *consortia* indicated that they like working with other schools on assessments and portfolios by stating:

- ❖ It's good to get together with them (principals) and talk about the problems we're having with the assessments—comparing our portfolios, comparing things, what we're doing. I think we've done a really good job on our portfolios and I think the main reason was just the communication between the different schools that were doing it.
- ❖ We develop our assessments as a group. We determined their validity, bias, etc., as a group. We aligned our curriculums as a group. We still have our own local standards and curriculum. We make our own local decisions for instruction based on our consortium test results to modify our curriculum and instructional strategies.
- ❖ In my role belonging to the consortium, probably the biggest part is the assessment development and the portfolio, helping write the portfolio. Since we did it as a consortium, they wrote the portfolio for those parts that were consortium issues and that only left us with our individual things like, how do we monitor that this kid's had the opportunity to learn before the assessment.

In the area of assessment literacy, principal responses in individual districts ranged from 2.72 to 4.67. Individual principal perceptions rated the strongest for developing the STARS assessment portfolio sent to the NDE and all schools using the same grade level assessments to meet STARS requirements. Individual principal perceptions rated

higher than collaborations, but lower than consortia principals, on students in schools being involved in understanding their own progress and achievement status; selecting assessment items from a common bank/pool provided by the district; and district development of standards based assessments for the schools.

Individual district principals said that students in schools were being involved in understanding their own progress by stating:

- ❖ I think the biggest thing is teachers who don't understand assessment. They don't understand the connection to assessment for learning, and don't understand the importance of giving kids feedback about their learning.
- ❖ Right now I feel that our criterion referenced test does a good job in addressing kids' abilities because kids take the test based on their own understanding of the content standards that are being used.

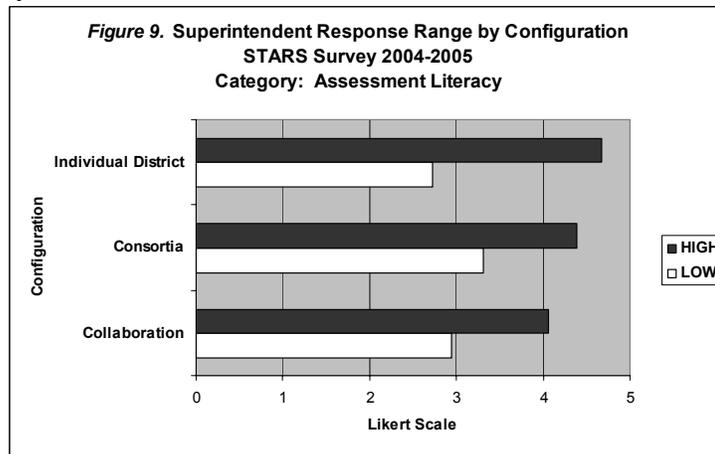
A *principal* indicated that his/her *individual district* used the same grade level assessments:

- ❖ Those common assessments were so critical because teachers had such autonomy before. Teachers go into their classrooms and do what they wanted and they all used different assessments. They used all different grading systems and assessment styles. And so nobody really knew how all the kids were doing. And so having those common assessments in place really gave us some key information to say, how are our kids doing, and what's happening to those kids that aren't achieving at that standard?

SUPERINTENDENT SURVEY RESPONSES

Assessment Literacy (Questions 1-16)

Superintendent response rate ranged from 2.53 to 4.62 for all configurations for assessment literacy as shown below:



In the area of assessment literacy, superintendent responses in all collaboration districts ranged from 2.53 to 4.13 as seen below. Perceptions of collaboration superintendents rated stronger than consortia superintendents, but less than individual district superintendents, on districts supporting the growth of administrators' assessment

literacy. Superintendent perceptions in collaborations rated the lowest in teachers designing assessment items for the district and all schools in the district using the same grade-level assessments to meet STARS requirements.

A *superintendent* from a *collaboration* district indicated stronger understanding of assessment literacy by this statement:

- ❖ I think initially there was a lot of help in how to design assessments. They brought in people from the Buros Institute. They talked about how to design. I think we've gone beyond that, that we know how to design assessments now.

A *superintendent* in a *collaboration* district stated that he/she supported teachers designing assessments and using the same grade level assessments by this statement:

- ❖ 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 students to that? I think they're more concerned with what they're assessing and how students are progressing.

In the area of assessment literacy, superintendent responses in consortia districts ranged from 3.05 to 4.43. Consortia superintendent perceptions rated the strongest in teachers designing assessment items for the district and all schools in the district using the same grade-level assessments to meet STARS requirements. Consortia superintendent perceptions rated the lowest in districts supporting the growth of administrators' assessment literacy.

A *superintendent* of a district in a *consortium* provided the following statement of support for teachers designing assessment items for the district and using the same grade level assessments:

- ❖ I think we've learned that the consortium can help us with the technical aspects of it but that we need to really write assessments that fit with our curriculum.

A *consortia superintendent* made the following statement about his/her own assessment literacy growth:

- ❖ Well, I went to the STARS academy last summer and, from beginning to end, that was a new learning. The thought of educators sitting down and going through a validation and reliability process over assessment issues is something I've never experienced before. Usually you, the night before a test, you sit down and you make out a test. Is it valid and reliable? That wasn't even an issue. But the whole thing was revolutionary, and arguing or discussing the merits of a particular question or the reason that this particular assessment was developed.

In the area of assessment literacy, superintendent responses in individual districts ranged from 2.55 to 4.62. Individual district superintendent perceptions rated the highest in districts supporting the growth of administrators' assessment literacy.

Individual district superintendent perceptions rated lower than consortia superintendents, but stronger than collaboration superintendents, in teachers designing assessment items for the district, and all schools in the district using the same grade-level assessments to meet STARS requirements.

Individual district superintendents indicated growth in their own assessment literacy by stating:

- ❖ The legitimate use of assessment has always been an integral part of our instructional process. It prompts the re-teaching, it moves us toward mastery based on the individual needs and, in some cases, it's as simple as more time allotted to a particular concept.
- ❖ We don't differentiate between the assessments we use for STARS and the assessments that any teachers would use. We want them to think about it the same way, that assessment is used to give feedback to kids and to change instruction so that all kids learn.

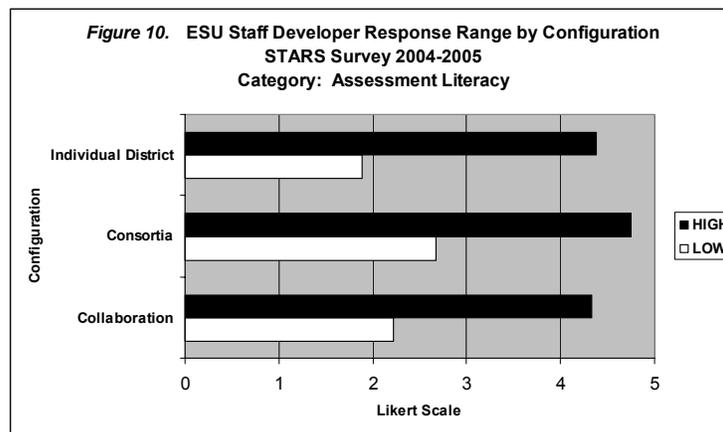
An *individual district superintendent* indicated that teachers designed assessments and all schools used the same grade level assessments:

- ❖ I heard them say on more than one occasion, I am glad I'm retiring so I don't have to do the assessments that you're going to be working on. And that's sad, but I think it tells you the initial belief that our staff had towards STARS assessment. I think as they've gotten into it, they realize that hey, this not only has held our students accountable, but it's held us accountable.

ESU STAFF DEVELOPERS SURVEY RESPONSES

Assessment Literacy (Questions 1-16)

In the area of assessment literacy ESU response range for all configurations ranged from 1.88 to 4.75.



The ESU responses regarding collaboration districts ranged from 2.22 to 4.33. The range for consortia districts was from 2.67 to 4.75 and for individual districts from 1.88 to 4.38. There were no statistically significant differences in the ESU perceptions in assessment literacy for collaborations, consortia, and individual districts.

CHARTING STARS: VOICES FROM THE FIELD

THEME 3: Data

We continually look at data after each assessment. It's something that we're getting better at. It used to be that we printed off the report, sent them in, and that was it. But I think we're starting to move toward the data-driven use of assessments to drive some of our decisions and what interventions kids need.

Interview Quote from a Nebraska Principal (2005)

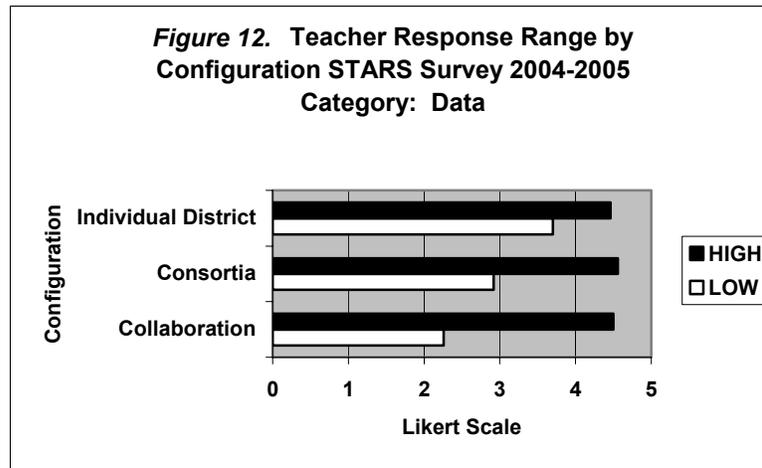
OVERVIEW

This section highlights the statistically significant or noteworthy findings from the third section of the survey, Data. The findings are supported with quotes that were collected during the interviews for those items where quotes were available. Surveys were administered to 4th, 8th, and 11th grade language arts and math teachers, principals, superintendents, and assessment coordinators of districts in collaborations, consortia, and individual districts, and ESU staff developers serving those districts.

TEACHER SURVEY RESPONSES

Data (Questions 17-22)

Teacher response rate ranged from 2.26 to 4.56 for all configurations as indicated below:



In the area of data, teacher responses in collaboration districts ranged from 2.26 to 4.50. Collaboration teacher perceptions rated lowest in receiving assessment data in a format that allows for disaggregation by student results; scoring all STARS assessments for schools; providing timely return of data; and interpreting the assessment results.

Collaboration teachers provided these thoughts about assessment data and data disaggregation:

- ❖ Data is that four-letter word I'm scared of.
- ❖ All of the scoring and reporting is done directly by the classroom teacher. We check the papers, we assign the raw scores, and we assign the mastery levels.

Secondary *teachers* in *collaborations* stated the following about the timely return of data and interpreting data results:

- ❖ Well, if the kids are not proficient or advanced, we try to reteach that material. And that's the biggest way I use that data and give them another chance to pass the assessment.
- ❖ We take a look at past scores, goals that we want to attain in that area or areas, and then we sit down and brainstorm and look for ways that we can improve it.

In the area of data, teacher responses in consortia districts ranged from 2.92 to 4.56. Consortia teacher perceptions rated the highest for receiving assessment data in a format that allows for disaggregation by student results and for timely return of data. Consortia teachers were higher than collaboration teachers, but lower than individual district teachers, for scoring all STARS assessments for schools and interpreting assessment results.

Consortia teachers stated the following about the data that they received for disaggregation:

- ❖ We're always looking at data...always looking at it for so many reasons and I think that's why it's important. I think we need as much data as we can possibly collect. We're always looking at it to study a problem or to find a strength or a weakness.

- ❖ So I think then you can start seeing some good trends and see what patterns are out there.

Consortia teachers commented about the timeliness of data:

- ❖ But now, I'm using the data more because it's done for me and it's not. . .time, more time I have to take out of my day. So this way, I think I can really use the data.
- ❖ It would be nice to have somewhere where we could get some of that information back to us. I think. . .we turn that in the spring and then when school starts, you get busy and you're all hyped and ready for school that we don't have. . .we don't have that data come back.

Teachers in *consortia* provided these thoughts about the scoring of assessments:

- ❖ I do all my own scoring. I have to grade them all and have to type all their answers into the computer.
- ❖ We score and report our own. We have a big database that's been created for our district and each of my 4th grade students is listed on a big paper that has each assessment for language arts and math across the top and then the levels—beginning progressing, proficient, advanced. And after we've given the assessment and scored the assessment, then we mark it on there.

Teachers in *consortia* provided these thoughts about interpreting results:

- ❖ I try to use that data to drive my instruction. If I see an area that we're weak in, then I've got to change my strategies for instruction or I've got to change my materials or find new materials to help students improve in that area.
- ❖ I really like that that's all done for me. And I can print off reports. . .it's like the whole system keeps track and I don't have to do all of that paperwork business that I was just buried in. I can print out a report and see which students have taken it and how well they did and which students still need to take it.

In the area of data, teacher responses in individual districts ranged from 3.70 to 4.46. Individual district teacher perceptions rated the highest on scoring and interpreting STARS assessments results. Teachers in individual districts rated higher than teachers in collaborations, but lower than teachers in consortia, on receiving assessment data in a format that allows for disaggregation by student results and providing timely return of data.

Individual district teachers shared the following thoughts about the data that they received for disaggregation:

- ❖ Our principal hands out a chart that shows the percentages that the state has required or as in No Child Left Behind, what is required for us to have. And it's broke down by free and reduced lunch, Hispanics, white, etc. So we get a nice chart with students disaggregated.

- ❖ My school happens to have a very high mobility rate. We looked at boys and how they were learning things, girls, also the SPED population in our building, and Title I.

A *teacher* in an *individual district* provided this thought about the scoring of assessments:

- ❖ Well, we collect data in different ways. It's not just your paper pencil. We've got performance-based data that we can collect. We've got rubrics that we use. We use data from those.

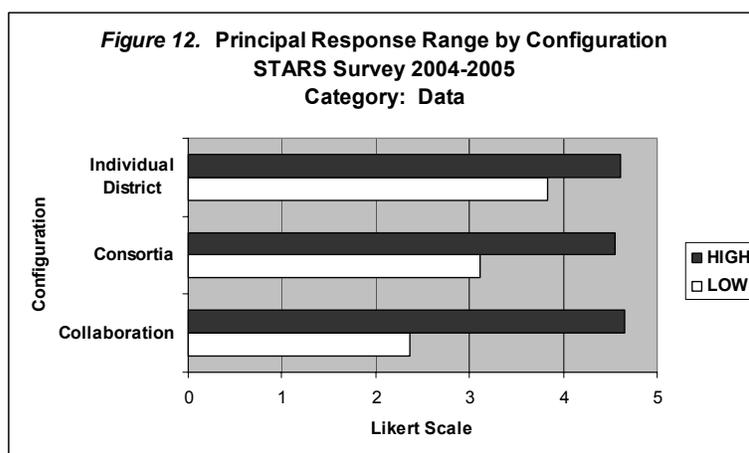
Teachers in *individual districts* shared these thoughts about interpreting results:

- ❖ I think we're only just beginning to try to find ways to use our data and, like I said, in the baby steps we've made I think it's pointed out to us certain learnings that need to be stressed more. I believe that it can help us because we've seen it just in very small ways.
- ❖ Well, it makes it much clearer on what I need to teach and how I need to teach and whether I taught it correctly or whether I taught it and it got through to the kids. It's given me a measurement so that I know that what I'm doing is successful. And if it isn't successful, then it's very clear that I need to go back and reteach.

PRINCIPAL SURVEY RESPONSES

Data (Questions 17-22)

Principal response rate ranged from 2.37 to 4.66 for all configurations as indicated below:



In the area of data, principal responses in all collaboration districts ranged from 2.37 to 4.66. Collaboration principal perceptions rated lowest in districts scoring all STARS assessments for schools, interpreting the assessment results, and districts providing recommendations for responding to assessment results.

A *collaboration principal* shared the following thought regarding scoring all STARS assessments for schools:

- ❖ We have a student information system that stores all the data, but also pulls it back out for us in forms of charts and graphs and we try to break it down by grade and by classroom.

Principals in collaborations made the following statements regarding the interpretation of assessment results:

- ❖ I would like us to use this time a little bit for the teachers to look at the data, for the teachers to dive into the data, and for them to come up with the solutions on how we're going to improve.
- ❖ We need to be better connoisseurs of the data itself. And again, I think right now there's a few key people in the district that are getting there and are starting to really push that data, but we aren't all there yet. So bringing everybody on board with how do you use data, what the data really means, interpreting it, and then actually taking those results and plugging them into instruction.

Collaboration principals' perceptions regarding recommendations for responding to assessment results:

- ❖ We've done several things with it (assessment results) as teacher teams, or I try to put it into. . . my teachers' hands as quickly as I can and sit down and we break it apart. Like I said, we do some different things with it as far as looking at classroom-to-classroom or breaking it into the subgroups if we need to. But we try to make it visual, have teachers look at it, have them talk about it, try to have a lot of discussion over what it tells us.
- ❖ As an administrator I have access to the cumulative data from our students and different years, and we've been able to utilize that data in determining target areas for our school improvement.

In the area of data, principal responses in consortia districts ranged from 3.11 to 4.55. Consortia principal perceptions rated higher than collaboration principals, but lower than individual district principals, in districts scoring all STARS assessments for schools, interpreting the assessment results, and districts providing recommendations for responding to assessment results.

Consortia principals provided the following thoughts regarding scoring all STARS assessments for schools:

- ❖ The consortium breaks it all down and gives it to us on a spreadsheet. And it's however we want it.
- ❖ It's built into our school improvement process that we do data review. We break the staff into groups, and they each take a grade level and a subject area like fourth grade math or eighth grade language arts. And then I have a spreadsheet that I developed that they look at the number of students in each quartile and enter that number. And then it prints out a chart that shows us where our students are at on each quartile in the different areas of the test.

Principals in *consortia* shared the following thoughts regarding the interpretation of assessment results:

- ❖ The data that we get, we use it to see what we can do to improve our programs that we have now. And this year we went and visited some other schools and their reading programs. So our data gives us the black-and-white in front of us score-wise. It tells us what we need to celebrate, what we've done right.
- ❖ I look at data and I think my main interpretation is it good or bad? Are the students progressing like they should? And if they're not, what do we need to change?

Consortia principals provided the following perceptions about responding to assessment results:

- ❖ I have an appreciation of how important that data is and how important it is to actually be accountable and able to verify just exactly what kids are doing. So we have a better idea of knowing where to start to help.
- ❖ For a long time everybody's been giving standardized tests. And then you put them in a file and you don't do anything with them. This whole process makes you look at your data and makes you do something with it. I think, from that point of view, it's been a really good thing.

In the area of data, principal responses in individual districts ranged from 3.84 to 4.61. Individual district principal perceptions rated the highest on districts scoring all STARS assessments for schools, interpreting the assessment results, and districts providing recommendations for responding to assessment results.

Individual principals provided the following thoughts regarding scoring all STARS assessments for schools:

- ❖ As far as the scoring right now, I'm not directly involved in the scoring, but I wanted to know how it was done at that time so I'd have an understanding. So I did spend time. And what we do now is we actually hire some substitute teachers that come in and help score.
- ❖ Now one way it does come back is broken out by those subgroups. So then you can talk about subgroups that have not performed up to the expectation that you would have. I think that our district is ready to move to the next level which is we are now being shown how to request information through our technology.

Principals in *individual districts* indicated the following comments regarding the interpretation of assessment results:

- ❖ We gather data and the way that we use data to assess students, and then look at that data to make changes in those assessments and make students more successful. In the past we've gone through the evaluation process with staff members like many other districts do. But we've never ever had a process lined up to where we could gather data or we could assess data, where we could look as a group toward improving student achievement.

- ❖ So one thing is just how students overall performed at a grade level, in a subject area. But also, we're able now to look at data by class, by teacher and by student. So when we get that information back now, we're actually able to go back to the teachers or the department, and we're able to talk with them about how their students performed in relation to other students.

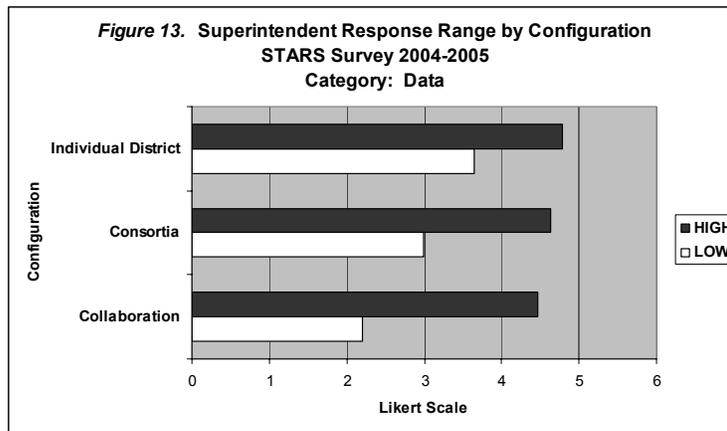
Individual principals provided their thoughts regarding recommendations for responding to assessment results:

- ❖ We've looked at it very strongly, especially at our primary grades in the area of reading. And the data has spoken to us in such a way that we have developed an intervention program in reading at our kindergarten and first grade level. So we have used that 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. It has been the key to what we do for staff development in this district.
- ❖ The plus is it gives you an opportunity to look at data and make decisions that are going to affect kids and outcomes on kids.

SUPERINTENDENT SURVEY RESPONSES

Data (Questions 17a-22)

Superintendent responses ranged from 2.20 to 4.79 for all configurations as shown below:



In the area of data, superintendent responses in collaboration districts ranged from 2.20 to 4.47. Collaboration superintendent perceptions rated strongest in schools receiving assessment data in a format that allows for disaggregation by classroom results. Collaboration superintendent perceptions rated lowest in schools receiving assessment data in a format that allows for disaggregation by school results and receiving data in a format that allows for disaggregation by relevant populations for individual students. Collaboration superintendent perceptions also rated lowest in scoring all STARS assessments for schools, providing timely return of data, and interpreting the assessment results.

A *superintendent* in a *collaboration* district offered this statement regarding schools receiving data in a format that allows for disaggregation:

- ❖ My role is to say if we can't break it out by free and reduced lunch, figure out a way to get it broken out by free and reduced lunch. We are working real hard at taking the technical aspects of assessments out of the hands of teachers and allowing teachers to teach.

Collaboration superintendents shared these thoughts about interpreting assessment results:

- ❖ And I just think it's too early to interpret the kind of longitudinal data. I think we have separate pieces of information that we don't really know how they all fit together and won't for a couple of years.
- ❖ Students who are not college-bound, but are life-bound, need the tools to go out and look for their first job. We have used the data to pull out some kids for extra help. The pull-out programs are tough. They're tough. They're good for kids, but they're not what the kids want.

In the area of data, superintendent responses in consortia districts ranged from 2.99 to 4.63. Consortia superintendent perceptions rated strongest in schools receiving assessment data in a format that allows for disaggregation by school results and timely return of data. Consortia superintendent perceptions rated lowest in schools receiving assessment data in a format that allows for disaggregation by classroom results. Consortia superintendents rated stronger than collaboration superintendents, but lower than individual superintendents, in receiving data in a format that allows for disaggregation by relevant populations for individual students, scoring all STARS assessments for schools, and interpreting the assessment results.

A *consortium superintendent* made the following statement about the scoring of assessments:

- ❖ A lion's share of that is record keeping and done by the consortium using the electronic technologies available to make that job as easy as possible.

A *superintendent* in a *consortium* district made this statement regarding the timely return of data:

- ❖ . . . actually, teachers know the data—the immediate data—right away. So, there's no excuse for not reteaching to the objectives and goals that may have been missed by the students. Teachers as well as students know immediately how they've done. So teachers can make a professional decision to either teach or reteach those particular areas that need to be addressed.

A *consortium superintendent* shared this thought about interpreting assessment results:

- ❖ In looking at when we sat down to begin work on our math improvement plan, the teachers brought to the meeting the reports generated from the various testing to look and see where we were scoring low and what we might need

improvement on. And then we discussed what we might be able to do to bring about a change.

In the area of data, superintendent responses in individual districts ranged from 3.65 to 4.79. Individual superintendent perceptions rated strongest in schools receiving assessment data by relevant populations for individual students, scoring all STARS assessments for schools, and interpreting assessment results. Individual superintendent perceptions rated stronger than consortia superintendents and lower than collaboration superintendents in receiving assessment data in a format that allows for disaggregation by classroom results. Individual superintendents rated stronger than collaboration superintendents, but lower than consortia superintendents, in receiving data in a format that allows for disaggregation by school results and providing timely return of data.

In *individual districts, superintendents* provided these statements regarding schools receiving data in a format that allows for disaggregation:

- ❖ At the school improvement level, when we're analyzing data, we use our student information management system. We look at every demographic factor you can think of. And it's just limitless what we look at. I mean we look at the main reporting categories for Nebraska and for No Child Left Behind. But what we're finding is that principals and school improvement teams now are asking for analysis in other ways.
- ❖ I think we try to collect our information and our data so that at some point it can be disaggregated and used as needed in any particular area or subgroup. Obviously, we want to use it to improve student learning.

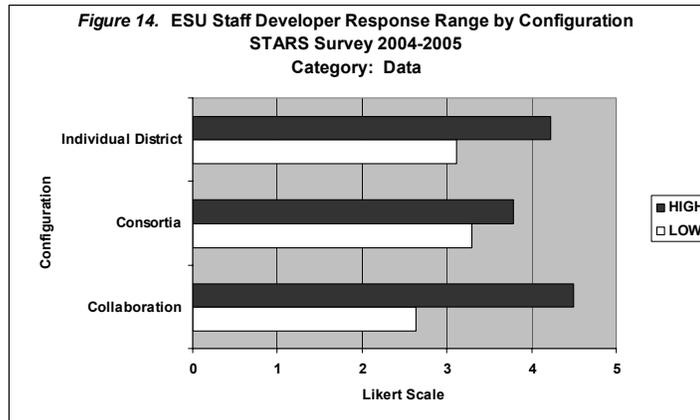
Individual district superintendents shared these thoughts about interpreting assessment results:

- ❖ I think we've only begun to understand how to read data. And data, for the most part, has scared us because we didn't really know how to make as much out of it as we could. So I would say all of us, from our physical education teacher right through to the high school principal and probably the superintendent, need better training in how to use data to make good, sound educational decisions.
- ❖ For every building and every principal and school improvement team, analyze that data. And we support them in that we walk through how you look at the data. We first started doing that three years ago with the initial set of data. So you have a principal and their school improvement team, usually four to five teachers. We walk them through how to analyze the data.

ESU STAFF DEVELOPER SURVEY RESPONSES

Data (Questions 17a-22)

In the area of data all ESU responses for all configurations ranged from 2.63 to 4.50 as indicated below:



The ESU responses about collaboration districts ranged from 2.63 to 4.50. The range for consortia districts was from 3.29 to 3.79 and for individual districts from 3.11 to 4.22. There were no statistically significant differences in the ESU perceptions about data for collaborations, consortia, and individual districts.

CHARTING STARS: VOICES FROM THE FIELD

THEME 4: INSTRUCTIONAL IMPACT

It makes you aware that every child does learn differently. If they don't understand it the first time, find out how you can get them to learn it. That is your job.

Interview Quote from a Nebraska Teacher (2005)

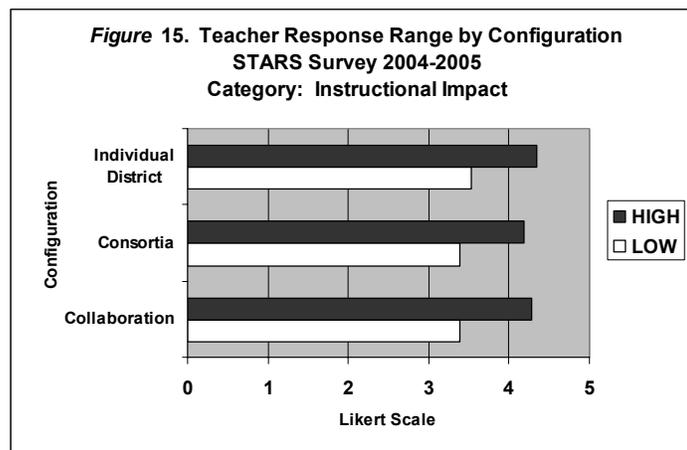
OVERVIEW

This section highlights the statistically significant or noteworthy findings from the fourth section of the survey, Instructional Impact. The findings are supported with quotes that were collected during the interviews for those items where quotes were available. Surveys were administered to 4th, 8th, and 11th grade language arts and math teachers, principals, superintendents, and assessment coordinators from districts in collaborations, consortia, and individual districts, and ESU staff developers serving those districts.

TEACHER SURVEY RESPONSES

Instructional Impact (Questions 23-35)

Teacher response rate ranged from 3.39 to 4.50 for all configurations as indicated below:



In the area of instructional impact, teacher responses in collaboration districts ranged from 3.39 to 4.28. The strongest collaboration teacher perceptions were in helping their

schools align curriculum to state standards. They rated lowest in helping schools establish benchmarks for meeting state standards K-12.

Collaboration secondary *teachers'* perceptions regarding helping their schools align curriculum to state standards rated the strongest:

- ❖ The biggest thing there was when we adopted the state standards; it kind of gave some direction, maybe, to my instruction on what topics I need to cover and some extra that I didn't need to cover. And there were some other important ones that we had to adopt from those state standards that directed our teaching to focus our instruction a little better.
- ❖ The curriculum and the STARS are intertwined at our school.

Collaboration teachers are concerned about all students having the opportunity to learn:

- ❖ One of our students at the high school level never got to the geometry course itself. A lot of geometry is covered in the other curriculum. But the 9th graders that come into pre-algebra and do algebra A and algebra, never take the geometry course. So, we have to make sure the geometry standards are covered somewhere else in their curriculum.
- ❖ We looked at the special education scores and we need to, as a district, really focus on the Individual Education Plans (IEPs) and what modifications are allowed for those students.
- ❖ Well, first of all, we'll have to go through the board of education if we're going to change some graduation requirements. We require three years of math. But it doesn't specify what three years, so a student may skip a year somewhere in there. Not being in a math class until they're a senior and the way the assessment is set up for 11th grade, that's kind of late to assess a student then because you already had to report on them.

Collaboration teachers' perceptions rated the lowest in helping their school establish benchmarks for meeting state standards K-12:

- ❖ But at the benchmark grade levels, the instruction has greatly improved. It's been more focused. We know where we need to focus because the standards are in place. They're basically telling us this is what needs to be completed before the end of this grade level.

In the area of instructional impact, teacher responses in consortia districts ranged from 3.39 to 4.19. The lowest teacher perceptions in consortia were in helping their schools align curriculum to state standards. They rated lower than individual teachers, but higher than collaboration teachers, in helping schools establish benchmarks for meeting state standards K-12.

Consortia teachers' perceptions regarding helping their school align curriculum to state standards rated the lowest of the configurations:

- ❖ I think, for me, it's made what I do in the classroom more meaningful. Before, our curriculums here in our district were based on the scope and sequence of the textbook series. And now what we're doing is we're working exactly backwards of that. We're writing a curriculum based on standards and then choosing a textbook that fits the curriculum that we've written.
- ❖ My gut feeling is that I'd like to see the standards, assessment, curriculum, and report cards aligned. They aren't aligned yet. I'd like to see our report cards and the way that we show the community, the patrons and the parents, that we're assessing their children.
- ❖ I know that if students come in to me from any other place in Nebraska or if my students leave and move to another school in Nebraska, hopefully, we're all close to being on the same page in terms of what we expect students at the 4th grade level to know and be able to do. I think it just gives us a lot more continuity across the state. What I see as most important locally is that we now write our curriculums based on the standards and then choose a textbook that best meets our needs instead of choose a textbook and then develop our curriculum.

The perceptions of *teachers* in *consortia* districts rated lower than those of teachers in individual districts and higher than those of teachers in collaboration districts in helping schools establish benchmarks for meeting state standards K-12:

- ❖ I wasn't actively writing the assessments, but I was working with teams as they wrote their assessments. When they did the unpacking of a standard, we were to guide them through that process. Then, when they started looking at what kind of questions or what type of an assessment they wanted to create, we were to provide input in terms of do you have enough questions at the progressive level, the proficient level, the advanced level. . .keep the big picture in mind.

In the area of instructional impact, teacher responses in individual districts ranged from 3.53 to 4.34. The strongest perceptions for teachers in individual districts were in helping schools establish benchmarks for meeting state standards K-12. Teachers in individual districts rated higher than consortia teachers, but lower than collaboration teachers, in helping their schools align curriculum to state standards.

Individual district teachers' perceptions of helping their district establish benchmarks for meeting state standards K-12 was rated the strongest of all three configurations:

- ❖ I think originally, when we first started out, when it was at 4, 8 and 11, a lot of us thought, 'Yahoo, I don't teach those grade levels so I'm safe.' And then once we made it K-12, I think just a lot of teachers felt as if they were being graded according to how their students did on the assessments and they were very wary and nervous. And I think over time we've really done a good job of making that feeling go away.

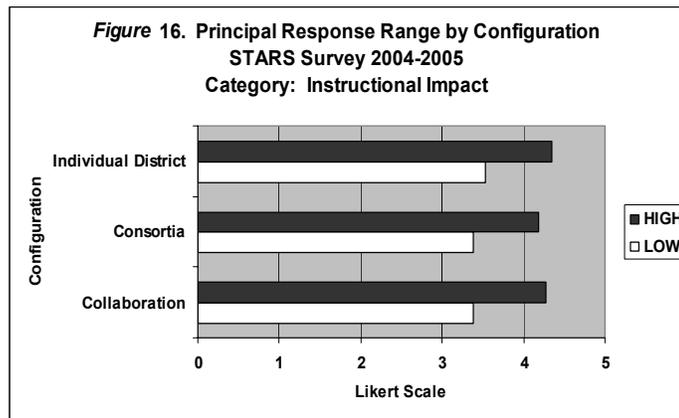
Teachers in individual districts rated higher than consortia teachers, but lower than collaboration teachers, in helping their school align curriculum to state standards:

- ❖ I think it makes me, as a teacher, more reflective about how I assess students, how I structure questions, and does my target match my test basically. I am much better at aligning that, I think.
- ❖ I actually feel that our assessments are hand-in-hand with our curriculum. We wrote them so that they could be used as a chapter test if the teacher wanted to. They could be used as supplemental assessing. So I don't feel that it's extra work, necessarily, for the teacher. It's something that they would do anyway whether they give the performance assessment.
- ❖ The STARS process in Nebraska is one of the best when it comes to allowing the district to decide what is important with the curriculum in the district.

PRINCIPAL SURVEY RESPONSES

Instructional Impact (Questions 23-35)

Principal response rate ranged from 3.61 to 4.47 for all configurations as indicated below:



In the area of instructional impact, principal responses in collaboration districts ranged from 3.61 to 4.31. The collaboration principal perceptions rated the highest in districts helping their schools align curriculum to state standards and helping schools establish benchmarks for meeting state standards K-12. Collaboration principals scored the lowest in schools aligning lesson plans to assessment data and state standards and teachers modifying their instruction in cases where students did not perform well on assessments.

Collaboration principals' perceptions rated the strongest in districts helping their schools align curriculum to state standards:

- ❖ We're talking curriculum now, where in the past districts had their curriculum guides which were to provide guidance to the teachers. But you never had a continual discussion about curriculum and instructional methodology and good assessments. At times that was piecemealed through various district in-services that you did or building-level discussions, but not something on a consistent year-to-year basis. I see that as a benefit of STARS. Now, curriculum is a major discussion item.

- ❖ I do have a problem with sometimes we are eliminating some general information that students might find exciting, might find motivating at a later point in their life. We've kind of narrowed things down in order to accomplish our assessment goals. And that's my observation, my personal observation of what our education in Nebraska has done in the last ten years.
- ❖ One of the things that I currently do is continue to challenge our staff to see if they are bringing the STARS into our curriculum, and it's part of our curriculum.

The perceptions of *collaboration principals* rated the strongest among all three configurations in helping schools establish benchmarks for meeting state standards:

- ❖ If I'm teaching in the classroom, if I'm teaching certain assessments that are being covered also by the math teacher, as an example, we right away set forth to try to determine and establish where the cross-overs are and where we might be covering assessments in different areas.
- ❖ We've had a lot of background information on proper development of assessments, working with the Buros Institute. And I have attended, personally, several workshops to help determine what valid assessments are as compared to, I guess, invalid assessments.

Collaboration principals rated the lowest in configurations in aligning lesson plans to assessment data and state standards:

- ❖ I'm very pleased to have the knowledge base that I have because I think it makes me make better decisions for kids.
- ❖ What can we do to improve reading scores? So we've taken that data and then implemented some ways that we think will help improve them.

Collaboration principals rated the lowest of configurations in teachers modifying their instruction in cases where students did not perform well on assessments:

- ❖ We look at what kids really need intense intervention and those kids, then, are the kids that we try to pull out separately and get that intervention.
- ❖ I think the greatest value is we have always scored well, let it be state or national test. This made us open our eyes a little bit to make us a little bit more accountable. Maybe make us look a little deeper at all the areas that, across the state, people are feeling are the essential outcomes that we want. And we had diagnosed a little bit better than what we've been doing before on some of the weak areas-some of the weak links-and what can we do to meet the needs of some of these students or some of these groups, or maybe a concept. I think it forces us to continue to see how we can improve.

In the area of instructional impact, principal responses in consortia districts ranged from 3.68 to 4.35. The strongest principal perceptions in consortia were in schools aligning lesson plans to assessment data and state standards and teachers modifying their instruction in cases where students did not perform well on assessments. The consortia principal perceptions rated the lowest in districts helping their schools align

curriculum to state standards and helping schools establish benchmarks for meeting state standards K-12.

Consortia principals rated the strongest for districts aligning lesson plans to assessment data and state standards:

- ❖ The standards have really done a nice job of earmarking what kids should know. Then we've aligned assessments to those standards to see if our kids are really learning those. So I guess what I learned for now. . . I learned if our kids aren't learning a particular content that they should be, let's figure out what we can do to get it; whether it's changing how we're teaching them, changing a textbook, it might be a lot of different things.

Consortia principals' perceptions rated the strongest in teachers modifying their instruction in cases where students did not perform well on assessments:

- ❖ If we're teaching and they're not learning anything, these assessments will tell us that. I know that it has brought about some different ways for our staff to look at how they teach materials. I don't know how much it's done for me but I know I've noticed the staff that they have looked at how they do things in a little different light because of it. And so I guess I'm looking at that with them.
- ❖ We find the weak areas and, as a staff, brainstorm ways to use different interventions. Like on our reading, reading was the first issue that we identified in our North Central Plan. Then the staff developed what they thought would be interventions that would be effective and we obtained training in those interventions.

Consortia principals' perceptions rated the lowest in districts helping their schools align curriculum to state standards:

- ❖ I've had to take a deeper look at what the school district curriculum really wants us to teach. Also, a deeper look at what the state standards are. No Child Left Behind also affects it. But it's just probably been a bigger study of the curriculum as a whole.
- ❖ It's gotten me involved with the teachers in the planning and the math curriculum and social studies curriculum. It gives me more of an idea of what we're offering, what levels the kids are taking. I just have a better understanding of the curriculum and a better understanding of what our teachers are teaching in the classroom.

In the area of instructional impact, principal responses in individual districts ranged from 3.67 to 4.47. Principal perceptions in individual districts were higher than consortia principals, and slightly lower than collaboration principals, for districts helping schools align curriculum to state standards and helping schools establish benchmarks for meeting state standards K-12. Principal perceptions in individual districts rated slightly higher than collaboration principals, but lower than consortia principals, on lesson planning aligned to assessment data and state standards and

teachers modifying their instruction in cases where students did not perform well on assessments.

Principal perceptions in *individual districts* rated higher than consortia principals and slightly lower than collaboration principals for districts helping schools align curriculum to state standards:

- ❖ We've looked at our curriculum; had to pare down some things that we've done in our curriculum to make sure that we are getting the instruction accomplished that needs to be accomplished so that they are successful with the standards at each grade level.
- ❖ It's a K-12 venture. And so there's a sharing of information. And I think the benefit with that—there's better articulation. Teachers at the high school can see what the middle school teaches and the grade schools and with that, I think there's a better concept of what we do district-wide in a K-12 subject.
- ❖ We are taking a serious look at our curriculum, making sure we're not having large gaps so that we're not teaching the same thing at two or three grade levels.

Principal perceptions in *individual districts* rated higher than consortia principals, and slightly lower than collaboration principals, in helping schools establish benchmarks for meeting state standards K-12:

- ❖ Well, obviously, the state department has mandated state standards in language arts and math, and continues on with science and social studies in the coming years. But I guess it allows us to revisit our curriculum guides, align our curriculum to our state standards, and provide instruction to meet benchmarks that the state department has laid out for us. I like the accountability aspect of state standards.

Principal perceptions in *individual districts* rated slightly higher than collaboration principals, but lower than consortia principals, on lesson planning aligned to assessment data and state standards:

- ❖ And so now I'm seeing a much more organized, coordinated approach to using assessments to help drive instruction. So to me, it's shown me that a process can be put in place that can make a difference, and you can use objective data to make decisions that will affect learning.

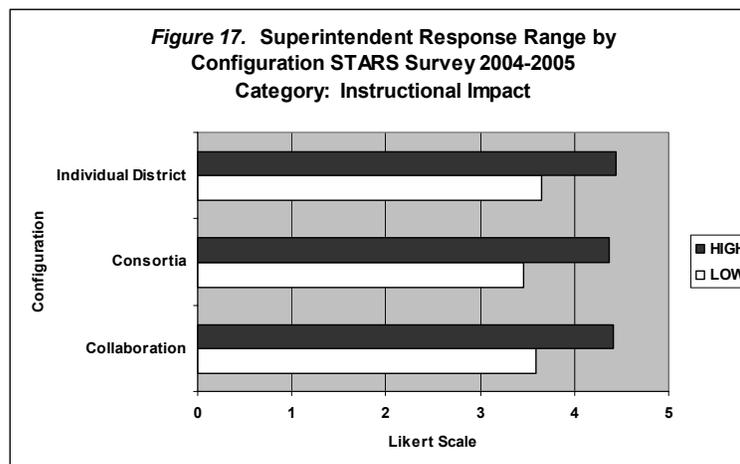
Principal perceptions in *individual districts* rated slightly higher than collaboration principals, but lower than consortia principals, on teachers modifying their instruction in cases where students did not perform well on assessments:

- ❖ Our central office has said that the assessments should really be a part of your grading process, not just something that's done to kids from the outside. You should use that as your classroom assessments because it's connected to your curriculum. Because then, that's one of your tests that's done and you're using it on a day-to-day basis to inform you about your kids.
- ❖ So we definitely use our assessment results to help generate our goal and strategies to improve reading and writing in this district.

SUPERINTENDENT SURVEY RESPONSES

Instructional Impact (Questions 23-35)

Superintendent response rate ranged from 3.46 to 4.45 for all configurations as indicated below:



In the area of instructional impact, superintendent responses in collaboration districts ranged from 3.59 to 4.41. The collaboration superintendent perceptions rated the strongest in teachers use of rubrics in instruction since the inception of STARS. The collaboration superintendent perceptions rated the lowest in lesson planning aligned to assessment data and state standards and teachers modifying their instruction in cases where students did not perform well on assessments.

The *collaboration superintendent* perceptions rated the lowest in lesson planning aligned to assessment data and state standards:

- ❖ I think a lot of times in the past; teaching and assessing have been two separate things. And I don't think that's true now. I think we're assessing what we are teaching.

The *collaboration superintendent* perceptions rated the lowest in teachers modifying their instruction in cases where students did not perform well on assessments:

- ❖ There should be no problem with a small school having vertical alignment of curriculum. We're small. We should know from first what's going on to second, second to third. That should never be an issue. So we should know where we're at and get to the goals.
- ❖ But most importantly, it looks at how we teach. What are things we need to work on and I said, 'How are we showing growth with our school improvement?'

In the area of instructional impact, superintendent responses in consortia districts ranged from 3.46 to 4.37. The consortia superintendent perceptions rated the strongest

in lesson planning aligned to assessment data and state standards and teachers modifying their instruction in cases where students did not perform well on assessments. The consortia superintendent perceptions rated the lowest in teachers use of rubrics in instruction since the inception of STARS.

The *consortia superintendent* perceptions rated the strongest in lesson planning aligned to assessment data and state standards:

- ❖ It's going to take time. We're going to have to examine the results of our students as well as our classes to see, number one, are we doing what we said we are doing? Is our curriculum aligned where it needs to be? Do we need to make some adjustments? And, if so, in which areas? That's the area that our school system is in need of spending more time and working more diligently toward.
- ❖ It's good because teachers are spending more time looking at the curriculum, being more concerned about what kids are learning in their classroom, and also empowering them.
- ❖ I think the process of establishing the goals and aligning our curriculum was a meaningful, practical way to accomplish what we always wanted to do in education.

Consortia superintendent perceptions rated the strongest in teachers modifying their instruction in cases where students did not perform well on assessments:

- ❖ Teachers as well as students know immediately how they've done. So teachers can make a professional decision to either teach or reteach those particular areas that need to be addressed.

In the area of instructional impact, superintendent responses for individual districts ranged from 3.65 to 4.45. The individual district superintendent perceptions rated stronger than collaboration superintendents, but lower than consortia superintendents, in lesson planning aligned to assessment data and state standards and teachers modifying their instruction in cases where students did not perform well on assessments. The individual superintendent perceptions rated stronger than consortia superintendents, but lower than collaboration superintendents, for teachers use of rubrics in instruction since the inception of STARS.

The *individual district superintendent* perceptions rated stronger than collaboration superintendents, but lower than consortia superintendents, in lesson planning aligned to assessment data and state standards:

- ❖ The curriculum groups look at not only what our standards are, but what are the best practices and how should those be taught. That translates into the assessments that very much match those best practices.
- ❖ We've put in a new curriculum development process, which includes the STARS and the assessments. So as new curriculum is written and put together, it's aligned with the STARS materials and the standards, and the assessments are developed at the same time. It's an ongoing process. Frequently the committees

are meeting, redoing the assessments, taking a look at the standards, and rewriting the curriculum.

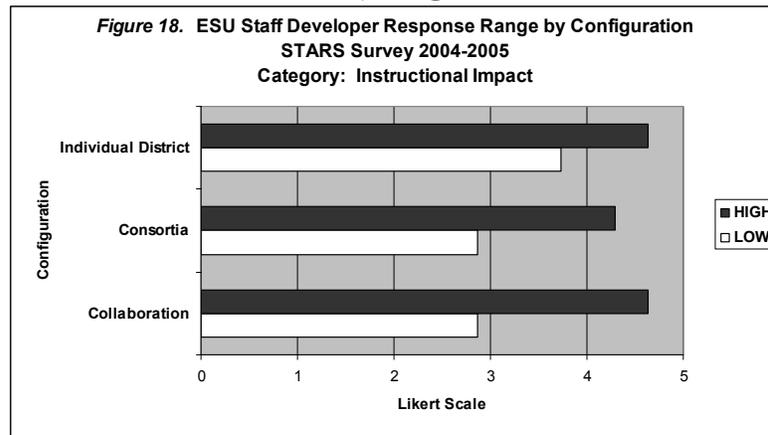
The *individual district superintendent* perceptions rated stronger than collaboration superintendents, but lower than consortia superintendents, in teachers modifying their instruction in cases where students did not perform well on assessments:

- ❖ The point is we use those assessments to give feedback to kids and to monitor learning and to adjust our instruction. So, we very much want that thinking district-wide that whatever assessment you give, whether it's for STARS or whether it's just something you do on a daily basis, you have to give feedback to kids.
- ❖ The legitimate use of assessment has always been an integral part of our instructional process. It prompts the reteaching, it moves us toward mastery based on the individual needs and, in some cases, it's as simple as more time allotted to a particular concept.

ESU SURVEY RESPONSES

Instructional Impact (Questions 23-35)

In the area of instructional impact, all ESU responses for all configurations (collaborations, consortia and individuals) ranged from 2.86 to 4.64.



The ESU responses about collaboration districts ranged from 2.86 to 4.29. The range for consortia districts was from 3.73 to 4.64 and for individual districts from 3.00 to 4.00. There were no statistically significant differences in the ESU perceptions about instructional impact for collaborations, consortia, and individual districts.

CHARTING STARS: VOICES FROM THE FIELD

THEME 5: EXTERNAL SUPPORT

Instead of saying you've got to figure all this out by yourself. There are all these experts on my side.

Interview Quote by a Nebraska Teacher (2005)

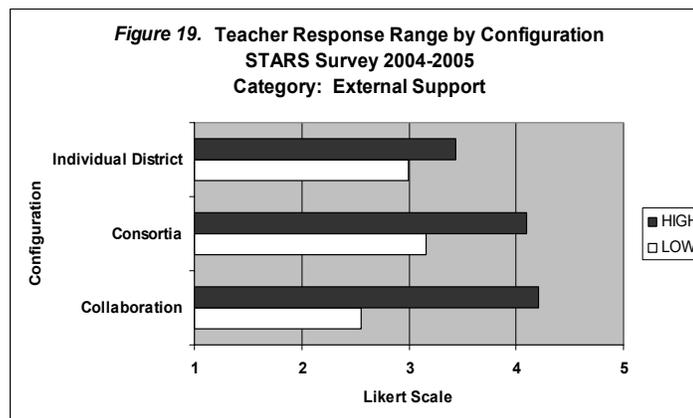
OVERVIEW

This section highlights the statistically significant or noteworthy findings from the fifth section of the survey, External Support. The findings are supported with quotes that were collected during the interviews for those items where quotes were available. Surveys were administered to 4th, 8th, and 11th grade language arts and math teachers, principals, superintendents, and assessment coordinators from districts in collaborations, consortia, and individual districts, and ESU staff developers serving those districts.

TEACHER SURVEY RESPONSES

External Support (Questions 36-42)

Teacher response rate ranged from 2.55 to 4.21 for all configurations as indicated below:



In the area of external support, teacher responses in collaboration districts ranged from 2.55 to 4.21. Teacher perceptions in collaborations were strongest in Educational Service Units (ESU) providing software for scoring, analyzing and reporting data to

support their school. Teachers in collaborations rated higher than teachers in individual districts, but lower than teachers in consortia districts, on the ESU providing on-going leadership training for school administrators and assessment training for teachers for implementing the STARS process, and the NDE providing leadership training for all administrators in schools. Teachers in collaborations rated the lowest in the NDE providing assessment training for teachers in schools.

Teachers in collaborations rated higher than teachers in individual districts, but lower than teachers in consortia districts, on the ESU providing on-going leadership training for school administrators and assessment training for teachers for implementing the STARS process:

- ❖ The ESU has really worked hard to get their staff out to work with school systems on an individual basis.
- ❖ I think they (ESU) do a super job. It's very organized and we have people on the local level that are leaders in each area that we can call if we need help. We can always call the ESU if we need help.

In the area of external support, teacher responses in consortia districts ranged from 3.16 to 4.10. Teacher perceptions in consortia were strongest in the ESU and the NDE providing on-going leadership training for school administrators and assessment training for teachers. Teachers in consortia rated the lowest on the ESU providing software for scoring, analyzing, and reporting data for schools.

Teacher perceptions in *consortia* were strongest for the ESU and the NDE providing on-going leadership training for school administrators and assessment training for teachers:

- ❖ I think our service unit is doing a fine job. They contact us to find out what we would like to have as far as teachers that would help us in our classrooms. And they usually provide very good programs.
- ❖ It seems like the relationship between our ESU. . .and our school is not very good because they don't like the consortium thing. They don't like the online testing and that's frustrating to me.

In the area of external support, teacher responses in individual districts ranged from 2.99 to 3.44. Teacher perceptions in individual districts were rated the lowest in the ESU providing on-going leadership training for administrators and assessment training for teachers. They were also rated lowest in the NDE providing training for administrators for implementing the STARS process. Teachers in individual districts were rated higher than consortia teachers, but lower than collaboration teachers, for the ESU providing software for scoring, analyzing, and reporting data to support the school. Teachers in individual districts were rated higher than collaboration teachers, but lower than consortia teachers, for the NDE providing assessment training for teachers in schools.

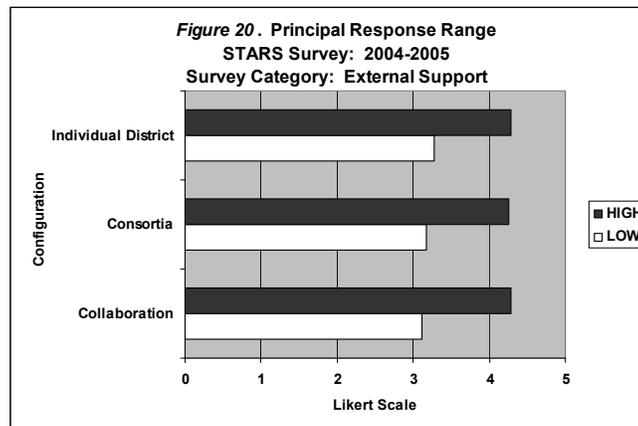
Teachers' perceptions in *individual districts* were rated the lowest in the ESU providing on-going leadership training for administrators and assessment training for teachers:

- ❖ I would say that our service unit really needs some kudos because they have been doing a lot of work to help teachers in our area understand and know what's going on. And they're always trying to keep on top of the schools and making sure that we know what's going on.
- ❖ As a FRESHMAN teacher, I have had no training about the STARS program.
- ❖ We have a monthly meeting with our ESU person and she's helping us with various issues and things that we need to do to get through the process.

PRINCIPAL SURVEY RESPONSES

External Support (Questions 36-42)

Principal response rate ranged from 3.12 to 4.28 for all configurations as indicated below:

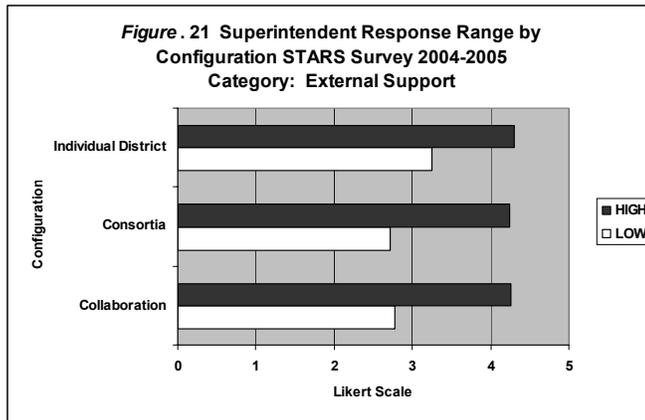


In the area of external support, principal responses in collaboration districts ranged from 3.12 to 4.28. In the area of external support, principal responses in consortia districts ranged from 3.17 to 4.25. In the area of external support, principal responses in individual districts ranged from 3.38 to 4.28. There were no statistical differences in principal perceptions of the external support provided.

SUPERINTENDENT SURVEY RESPONSES

External Support (Questions 36-42)

Superintendent response rate ranged from 2.71 to 4.30 for all configurations as indicated below:

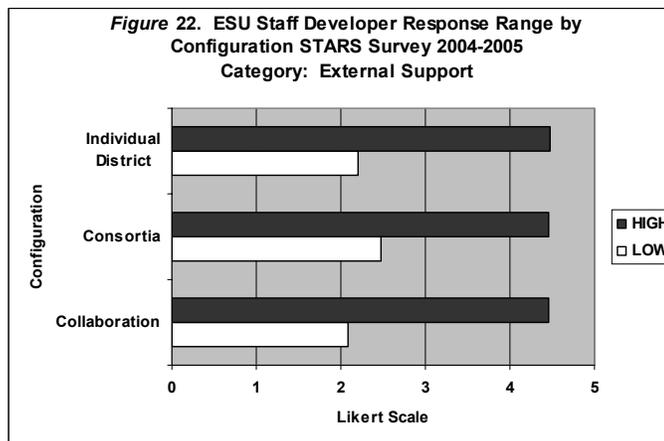


In the area of external support, superintendent responses in collaboration districts ranged from 2.78 to 4.26. In the area of external support, superintendent responses in consortia districts ranged from 2.71 to 4.25. In the area of external support, superintendent responses in individual districts ranged from 3.25 to 4.30. There were no statistically significant differences in superintendent perceptions of the external support provided.

ESU SURVEY RESPONSES

External Support (Questions 36-42)

In the area of external support, all ESU responses for all configurations ranged from 2.09 to 4.60 as indicated below:



The ESU responses about collaboration districts ranged from 2.09 to 4.45. The responses for consortia districts ranged from 2.47 to 4.60 and responses for individual districts ranged from 2.20 to 4.47. There were no statistically significant differences in the ESU perceptions about external support for collaborations, consortia, and individual districts.

CHARTING STARS: VOICES FROM THE FIELD

THEME 6: ACCOUNTABILITY

We have learned that we are not only accountable to our students, but I think it's made us more aware of accountability to a broader picture. I think we're accountable to a certain extent to every student and every administrator and every teacher within the state of Nebraska.

Interview Quote from Nebraska Superintendent (2005)

OVERVIEW

This section highlights noteworthy findings about Accountability, a theme that emerged when the interviews were coded for the purpose of identifying unanticipated themes. The following quotes were collected from grades 4, 8, and 11 language arts and math teachers, principals, superintendents, and assessment coordinators from districts in collaborations, consortia, and individual districts, and from ESU staff developers serving those districts.

TEACHER INTERVIEW COMMENTS

Teachers in collaborations shared a feeling of confidence about their own ability to write, implement, and score assessments to ensure that learning was occurring at the classroom level:

- ❖ To me as a professional, not only writing assessments for the state standards; but if I need to write an assessment for any area of the curriculum, I feel more confident in doing that and think I probably do a better job.

Teachers in both collaborations and consortia shared their feeling of accountability anxiety when stating:

- ❖ I've had these thoughts myself. . .you have to wonder. Do all teachers abide by rules and guidelines? It is a situation where you worry about how your school district is going to look because that immediately affects how you look as a teacher. And I guess I see us teaching. . .going back and teaching more to the test again. Not only doing that with assessments, but my district is starting to do that towards achievement tests again because we want good results.
- ❖ To me, this revolves around assessment and making sure the kids are learning and we're doing our job and they're doing theirs.

Teachers in *consortia* shared how students are provided feedback and held accountable:

- ❖ I think we have a good system and it seems like it saves a lot of paperwork and a lot of time on the teachers' part. And I think that's a benefit. I think it benefits the kids, too, because they're being held a little more accountable for what they learn. And even the teachers, too, because we're held accountable for what the students learn.

Teachers in *individual districts* shared their anxiety with federal mandates:

- ❖ You know, we want to have 100% of our kids be proficient in everything. Well, I think that's a little farfetched because I don't see that happening. Maybe some places far, far away, make believe, but not here.
- ❖ We start this process and our veteran teachers who are good teachers and have been here a long time say, 'You know what? It's about time to retire.' I just feel like it's looked upon as more of a burden as opposed to something that's going to improve achievement. I just think that No Child Left Behind has a lot of negative things with a lot of people and a lot of positive things with a lot of people.

PRINCIPAL INTERVIEW COMMENTS

A **principal** of a district in a *collaboration* shared the value of using trend data for school improvement:

- ❖ As an administrator I have access to the cumulative data from our students and different years. And we've been able to utilize that data in determining target areas for our school improvement.

A **principal** in a district in *collaboration* shared the accountability for teachers:

- ❖ I think it is holding the teachers responsible for the achievement of the students, which we've kind of always hoped has been there. But I think this is really holding some of their feet to the fire.

Principals in districts in *consortia* shared the value of accountability and the desire to improve student learning:

- ❖ We are more accountable and people are holding us accountable. And I think our school system here is working very, very hard to live up to that. And I think we're doing a good job.
- ❖ But it's given me the opportunity to look closer at student learning. And so, I think now we're more centered on what students know and should know.
- ❖ The administrators have taken ownership from the standpoint that we're going to make sure this is teacher centered and the teachers are involved.

Principals in *individual districts* shared the value of monitoring lesson plans and visiting classrooms to ensure students have the opportunity to learn:

- ❖ I guess the basic question would be how well are they following through. What I'm doing this year is tracking lesson plans. And what we're doing is having teachers write down their standard on the lesson plan and I keep track of the standards that have been addressed on the spreadsheet. So, more or less, a situation where we've established STARS and how well our teachers follow through.
- ❖ I believe my primary role has been to assist our curriculum coordinator and our classroom teachers by ensuring that before we do any STARS assessment, that instruction has occurred within the classroom. And in order to accomplish that, we do short classroom visitations, monitor lesson plans, and just make sure that the kids are instructed before they're ever tested on anything.

Principals in *individual districts* shared the support provided for being accountable to the STARS process:

- ❖ I think I have the support of the board and my superintendent to do what's needed to get scores where they have to be. I have not been turned down with a no to spend the necessary money to get that done. They know that they are state mandates and so we have to honor and respect state mandates.
- ❖ When I look at the district, I think . . . trying to find more time where we can get staff together and it can't be instructional time. So somehow, I think as a district we're going to have to step up and say we're going to pay a per diem or something like that to get together in the summer to help write the assessments and get the buy-in.

SUPERINTENDENT INTERVIEW COMMENTS

A superintendent of a district in *collaboration* shared his/her thought on the value of standards and assessments:

- ❖ The one thing about assessment that I think has been good is I think it's raised the level of what we're held accountable. You just don't walk out the door anymore; it's not a big deal. It is a big deal. The sad thing for me as a professional educator, it shouldn't take somebody looking over our back. But it has. It has improved the quality of education. It has.

Superintendents and *assessment coordinators* of districts in *collaborations* shared how STARS has helped them focus on what is best for the student:

- ❖ It comes down to what's best for a child. And we still have to make those decisions. I think you've got to show that kids have the opportunity to learn.
- ❖ I'll tell you what; it isn't about getting every child to college. It's about getting everyone to their success level.
- ❖ Sometimes I've gotten the feeling that a teacher would go into his or her classroom and, even though they had the goals or objectives, once they shut the door, they kind of went on their own. My involvement now, and the piece that I like, is that we can kind of tell our students and tell our parents, this is what we want all our kids to know. And we're going to assess to find out that we know it.

And then it's added that responsibility piece to the teachers that we are going to come back with this data to make sure kids are being taught what we want them to know.

Superintendents and *assessment coordinators* of *individual districts* shared some frustrations with state and federal mandates:

- ❖ It's causing difficulty for us and every year I think every school is going to realize exactly what kind of bind they're in. But we're trying to respond to it appropriately and meet the spirit, as well as the intent, of the law, which is a little frustrating.
- ❖ We're accountable not only to ourselves, but also to the entire state. So the entire state shows that, indeed, we're teaching kids, that we're doing a good job in seeing their improvement academically.
- ❖ They like what we're doing. They think it's good for kids. And STARS allows it. I'm eternally grateful for the leadership the commissioner has given. I think he stuck his neck out to do the right thing. But we need to take a step in Nebraska. And that is to do that portfolio evaluation process and do it right.

ESU INTERVIEW COMMENTS

ESU *staff developers* provided the following thoughts on accountability:

- ❖ When the NDE came out two years ago looking at sufficiency and looking at reliability, that's when all our schools, as a group of small schools, said we could no longer be individuals. We have to be a consortium to look at those specific criteria.
- ❖ If nothing else comes out of the whole process, we have good curriculums, people are using them, they're aligned to the standards, and there's a focus now in the learning process that wasn't there before.

CHARTING STARS: VOICES FROM THE FIELD

THEME 7: LEADERSHIP

One of the strong learnings for me is that leadership is everything.

Interview Quote from an ESU Staff Developer (2005)

OVERVIEW

This section highlights noteworthy findings about Leadership, a theme that emerged when the interviews were coded for the purpose of identifying unanticipated themes. The following quotes were collected from 4, 8, and 11 language arts and math teachers, principals, superintendents, and assessment coordinators of districts in collaborations, consortia, and individual districts, and from ESU staff developers serving those districts.

TEACHER INTERVIEW COMMENTS

Teachers in consortia districts shared how they have led the STARS Process:

- ❖ I don't see that we've had any major problems or obstacles. We just do it. As elementary teachers if they say you do it, you do it.
- ❖ I would even count the fact that I was absolutely forced to learn how to do this on a computer as an 'extreme-for me.' That was a learning experience, but it has also served as a confidence builder.

Teachers in individual districts recognized and displayed all of the qualities of leadership in the assessment process:

- ❖ 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 themselves that the teachers take and the ways in which the assessments are to be scored.
- ❖ Again, I think our principal does an awful lot of work there and he basically takes all the raw data that I give him and works with it to segregate it by population. He is great about reporting the results, sharing them with us, and discussing them with us. He plays a very strong leadership role there, I believe.

PRINCIPAL INTERVIEW COMMENTS

Principals in districts in *collaborations* shared increased knowledge and decision-making:

- ❖ I'm very pleased to have the knowledge base that I have because I think it makes me make better decisions for kids. It really allows us to look at what kids should know and are we there. Are we doing what we should be doing or are we doing things that are just fun and exciting?
- ❖ I think my knowledge base has increased tremendously. I've been a part of it and been in on the ground floor and really been in the trenches so to say. I led the process.

Principals in districts in *collaborations* shared their role in leading the STARS process:

- ❖ For me individually, I've always been a strong learner and believe that curriculum has always been a strong interest for me. And so, I think it has fed into what I consider one of my roles to be. . . a learner, to model that for our faculty, but also to communicate. That goes back to the instructional leadership.
- ❖ Again, probably I led the process and worked with the teachers closely as far as the assessment literacy and what makes a good assessment and the teachers actually wrote the assessments. But I was there with them every day that we wrote them, and checking them and going back through things, asking questions, trying to be kind of the person that reflected on things and looked at quality and those kinds of things—more the questioner I guess—to be sure we had high quality assessments.

Principals in districts in *consortia* shared their involvement in the STARS process:

- ❖ I just feel more familiar with our classrooms because now I know what's going to be taught. We may do it in a different avenue of some schools, but I guess, just as an administrator, it would give me a brighter picture of what we need to do with our kids and what our teachers are going to be doing.
- ❖ I would say that it is organizational things. You know getting the teachers to get involved in it, making sure we have teachers attending the different workshops, the sessions that we have, getting their input, their ideas and, somehow bringing all of that together so that we're on the same page and that we're headed in the direction that's going to be beneficial to the kids.

Principals of districts in *consortia* are leading the way in making decisions and taking ownership for the STARS process:

- ❖ But if there are any decisions that maybe, in the end, need to be made, it's made by the administrators in the consortium.
- ❖ The administrators have taken ownership from the standpoint that we're going to make sure this is teacher centered and the teachers are involved. And then we move along in the process.

Principals of ***individual districts*** shared the many leadership roles within the STARS process:

- ❖ I think leadership's a big key to that. Look at your curriculum director. Are they pushing standards? Are your principals and your superintendent?
- ❖ The benefit of belonging to a district in terms of STARS is it promotes leadership within the school.
- ❖ As an instructional leader, not only do you lead the instruction in the building, you also learn along with your teachers what is required by the district as well as the state to fulfill the requirements.

Principals in ***individual districts*** shared their views about the responsibility they have to students for ensuring that staff understands the expectations of STARS:

- ❖ Well, I think that one thing districts have to do is keep in mind what is being asked of their principals and if you want them to be more than just a manager. Do you want them to be an instructional leader? It takes a lot of time to be in the classrooms, to have these meetings, these discussions to coordinate those, to be the person who pushes and sells and says this is going to make a difference and keeps that whole thing going.
- ❖ It's my job as the building leader to get to the department heads, get to the teachers that are giving the assessments and making sure they're on target with the expectation, and meeting the timelines, and then getting that information to the state department.

SUPERINTENDENT SURVEY RESPONSES

Superintendents of ***individual districts*** and superintendents and assessment coordinators of districts in ***consortia*** shared their many leadership roles within the STARS process:

- ❖ I think my role was basically to provide the time and the impetus for people to develop those assessments; to provide the district funds to be able to pay those people, to provide the time; to be able to release them from class to develop assessments; and to promote that with our principals.
- ❖ My role was the organizational part of it, making sure assessments are given to the teachers, collecting data, collecting the scores and getting them into the database, the portfolio for the district, setting up if we need any in-service training, or we need a time to write, rewrite or work on assessments, or. . .you name it.

Superintendents of districts in ***consortia*** shared knowledge and responsibility with teachers:

- ❖ I think certainly I can be conversant with my staff instead of them talking about something that they're doing in their assessments and my eyes glazing over with a blank look. I have a basic understanding of what they're doing and it gives me an appreciation of the workload they have.
- ❖ They (teachers) feel more of the direct leadership in our school system and in the curriculum.

A *district administrator* for an *individual district* shared his/her view on the role of individual leaders to create ownership for the STARS process:

- ❖ For me in my role, it was just learning the power of getting people involved and getting people to buy into a project. And indirectly, that was an offshoot of STARS. I think I've really learned how to move a really large organization and to get buy-in to that process.

A *superintendent* of an *individual district* shared his/her view on the role of an individual district leader for improving schools:

- ❖ I would personally write back to every principal, having read their year-end report and their school improvement plan, with the communication highlighting what I perceived to be the strength of their work and encourage a goal that they may have identified in terms of future work—active and involved, but not intrusive.

Superintendents and *assessment coordinators* of *individual districts* shared their views on individual leaders sharing leadership with classroom assessment leaders:

- ❖ It's the leadership we've had in our district. We took good classroom-based assessments that our best teachers are already using and we standardized those. It's been a monster, an absolute monster because, I'll tell you, a one-time, district-wide test would have been easier, no different than a one-time statewide test. And it's the multiple assessments, it's the performance-based assessments that make it so difficult but yet make it meaningful for kids and teachers.

ESU INTERVIEW COMMENTS

ESU *staff developers* provided the following thoughts on leadership:

- ❖ When you take a leadership role, you learn a lot of things that you didn't know because you're trying to make sure the schools are meeting the requirements and also what's ethical, what's right, what makes sense, what's practical.
- ❖ I think it's a question of leadership. If they see, if the leaders—teacher or administrator—if they see that picture of how to make it ours, it fits our community and our district, then it will happen.
- ❖ You need to be out on the cutting edge of what's about to occur and understand it because the schools have definite expectations of you knowing how to lead them.
- ❖ Our biggest obstacle has been a major shift in leadership. We brought in administrators from out of state. While that could bring in new ideas, what it seems to be doing is actually saying, 'I don't necessarily support this system.'

CHARTING STARS: VOICES FROM THE FIELD

THEME 8: PROFESSIONAL DEVELOPMENT

I think the biggest value is that we work together. We support each other, that we're willing to give the time to provide the time for the teachers, either through in-service or time sometime throughout the year, to provide them an opportunity to learn ways to create the best assessments, or to learn new teaching strategies to instruct the kids.

Interview Quote from a Nebraska Principal (2005)

OVERVIEW

This section highlights noteworthy findings about Professional Development, a theme that emerged when the interviews were coded for the purpose of identifying unanticipated themes. The following quotes were collected from grades 4, 8, and 11 language arts and math teachers, principals, superintendents, and assessment coordinators of districts in collaborations, consortia, and individual districts, and from ESU staff developers serving those districts.

TEACHER INTERVIEW COMMENTS

A *teacher* of a district in *collaboration* assisted in the planning of professional development:

- ❖ I've had a stronger part in that this year than I have had other years because I am the language arts teacher and our goal has been reading. I've had input with our superintendent about bringing people in to do in-services. I've had a role in trying to help other teachers look at some intervention strategies that we've used.

Teachers in *individual districts* shared new learnings from professional development opportunities and dialogue:

- ❖ A lot of the things, of which I learned in these workshops and the staff development opportunities, I take back into my classroom. And I then work with my students with the new strategies I've learned or some of the new approaches. And then with our language arts meetings in my building, I'm able to share that information with other teachers. Sometimes I think that's a big benefit because we can sit down and talk about things that are working well.

- ❖ I feel as if my voice, it matters, that what I feel is best for students is being brought up in meetings and I'm able to discuss my opinions with other teachers and we can bounce ideas off one another.

Teachers in *individual districts* shared and built new knowledge:

- ❖ I think, in our district that teachers from all levels have been brought together to discuss not only what we thought students should be able to do, but also to understand what's done at kindergarten and then what's done at first grade so that you can come and see the continuation and how that expands as you carry through the grades. And we're working, at the moment, to get assessment in all of the elective courses, as well, so that all the way through K-12 we would have assessment in all areas.
- ❖ And then within buildings, we have learning communities so that people get a sense of what's happening and how they can get their voice heard. And I think through that process, you get more buy-in with teachers. And I think that they feel, then, that it's more valuable to them and their students, which I know I've already felt.

PRINCIPAL SURVEY RESPONSES

A **principal** in a *collaboration* district discussed professional dialogue and training:

- ❖ Probably the best learning for our staff is that we are being able to communicate on the same level. We know that there is some essential education that has to be taught and our teachers have bought into that. I think it's allowed us to have more professional dialogue, not only what needs to be taught, but what are the best practices to teach that? And it's forced educators to have a common language. We're hearing more and more professional dialogue that, in the educational community, I think is so important.

Principals in districts in *consortia* shared professional dialogue and training:

- ❖ I have just learned a lot by dealing with other administrators. It's good to get together with them and talk about the problems we're having with the assessments, comparing our portfolios, comparing what we're doing. I think we've done a really good job on our portfolios and I think the main reason was just the communication between the different schools that were doing it.
- ❖ The portfolio workshop in June was an excellent session—and having models really assisted in knowing what the reviewers wanted.

Principals in *individual districts* shared training opportunities for teachers:

- ❖ What we do with our curriculum, our school improvement drives what we do with our staff development. And we've had workshops available. We've tried to allow teachers time to attend those workshops and then, once they go, time to implement that information when they get back.
- ❖ When those teacher trainers come back and do presentations to department groups in our building, I have felt that the discussion was very positive, very professional, especially when they're talking about instructional strategies that

research has shown will be effective, such as pre-reading strategies or certain writing strategies that can be used across the curriculum.

- ❖ So I think we're really on the verge of having that kind of a meaningful discussion. And then I think the other thing about just continuing to keep the information flowing and get feedback from everybody so that everybody feels as though they have had input through their representative and they are staying abreast of something rather than it just being something that happens to them.

SUPERINTENDENT INTERVIEW COMMENTS

Superintendents of individual districts provided the resources for professional development:

- ❖ Professional development is critical, and we have a very strong belief of 'training the trainers' on those committees to take it back so it does get back to all the teachers.
- ❖ I do the professional development that goes along with the best practices needed for implementation of those assessments and strategies.

ESU INTERVIEW COMMENTS

ESU *staff developers* provided the following comments on staff development:

- ❖ We have a very articulated process here. We tweak it every year as we get new information and things change. But we really go through a flow of that process and I can give you a document that shares exactly what our process is. And now we have a process for how we refine. So not only is it in the development phase, but most all of them now in the full content areas are in the full refinement phase.
- ❖ There is absolutely something new every month. Whether it be a new national perspective with No Child Left Behind coming in that changed the path that we had gone on, we had to really see how we could merge the two systems as best we could.
- ❖ I think it's important to note that if the funding doesn't follow, it's really hard to keep all of those different content areas going.

CHARTING STARS: VOICES FROM THE FIELD

THEME 9: CHALLENGES AND OBSTACLES TO SUCCESS

What I wake up in the middle of the night worrying about is data and not losing it! It's just the storage of the data. And it's so complicated. We have to give a number and we have to go in and search through all these numbers and we have to enter them, and it's not really a user-friendly system. I kind of call it a data nightmare. I don't know if anyone else is talking about it; probably just me cause I have to worry about it.

Interview Quote from an Assessment Coordinator (2005)

OVERVIEW

This section highlights the obstacles that get in the way of the implementation of the STARS process. This theme emerged when the interviews were coded for the purpose of identifying unanticipated themes. The following quotes were collected from grades 4, 8, and 11 language arts and math teachers, principals, superintendents, and assessment coordinators of districts in collaborations, consortia, and individual districts, and from ESU staff developers serving those districts

TEACHER INTERVIEW COMMENTS

Teachers in districts in *collaborations* and *consortia* and teachers in *individual districts* all indicated that the amount of time that the process takes is a huge issue:

- ❖ Time, of course, is probably the biggest one. Teachers are stretched pretty thin anyway, and then more paperwork, and having a chance to have time to talk with your peers, and get a chance to collaborate with people that are involved in the process—it's just that time is such a short commodity in education. It always has been. I think it always will be.
- ❖ If I had the extra time, I'd get more feedback from my kids, write better tests, see what assessments other schools are using, use more technology, I think, reach those state standards.
- ❖ The amount of time and effort that Nebraska teachers are asked to undergo on this HUGE process. Why are we asked to work so much more without any more time or compensation?

- ❖ I'm glad that I'm a part of a consortium that does the Online to give us more of a voice. It's just a major time saver doing them online.
- ❖ My biggest concern about STARS testing is the amount of time involved to administer the tests and then score. I teach in a multi-grade classroom, administering the tests takes way too much time.

Teachers in districts in *collaborations* and *consortia*, and teachers in *individual districts* expressed concern about the effect of change:

- ❖ Well, the obvious one is any time you change, people are resistant to change. Teachers who are upset about, 'Why do we have to do this?'
- ❖ I guess one of the major obstacles would just be communicating the assessment system to all of the teachers in the district, trying to get everyone on board and understanding why we're doing it, and how it helps students, and a reason they can use the data to inform their instruction.

Teachers in districts in *collaborations* and *consortia*, and teachers in *individual districts* believed that they should be compensated appropriately for the amount of time spent in the STARS process:

- ❖ (We) should have a consistent payment to teachers for work in this area. Teachers are frustrated that their pay is not as much as other districts. This is an administrative issue, but could easily be addressed through STARS grant guidelines. Also, the pay is taxed so high that teachers don't really receive enough as an incentive.
- ❖ We realize this is a process in work, BUT teachers need to be reimbursed for their expertise at contract salary. Administrators are given 11-month contracts to do their work. Why aren't we?
- ❖ STARS funding has been well used and much needed in giving teachers incentives to do extra work compiling all the data and analyzing it.

Teachers in districts in *collaborations* and *individual districts* indicated that data collection, reporting, and management is difficult:

- ❖ I think another area that takes up a lot of time is the data management. It's trying to keep track of the scores, keep track of the papers, making sure that the scores are written down, that they're put into the database. It's a constant upkeep of data all the time.
- ❖ The STARS assessments record keeping is a nightmare. It seems as if that is all we got accomplished. It's also very difficult to implement the assessments when teaching multiple grades.
- ❖ You have to input all this data, somebody has to collect the data, and somebody has to interpret the data. I mean, we've had to add people to do this.

Teachers in districts in *consortia* are concerned about data collection, reporting, and management:

- ❖ The only thing I can say with fourth grade is we're tested beyond. We have too many tests. And that's an obstacle, not only with STARS, with whatever has to be reported. But I think maybe we test a little bit too much, and that might be an obstacle.

Teachers in a **collaboration district** and in an **individual district** believed that the issue of retesting students should be addressed:

- ❖ We have never retested. We took the score and that was the score that we reported. Now we're hearing that there are districts that retest and retest until the student has a proficient score. So as a district we need to make some decisions about that. How are we going to do that? Are we going to retest?
- ❖ There's always something it seems that interrupts an assessment. It could be an absent student or it could be another activity that disrupts us from being able to do something without being interrupted. And it's to be expected, I think, to more or less a greater extent. There's always going to be an interruption or somebody gone and, of course, we make up for that by allowing them to come in and do it later or even redoing an assessment if necessary.
- ❖ One district retests with their local product literally until students pass. . .to my knowledge they don't have more than one version.

Teachers in a district in a **collaboration** and in an **individual district** wanted all teachers to participate and become assessment literate:

- ❖ 4th, 8th, and high school English and math are very involved in the process, sometimes to the point of overload. But other teachers are involved very little. They seem to have the attitude that it doesn't pertain to them. I think that partly stems from frustration over not knowing exactly what they are supposed to be doing in this process.
- ❖ Individual district teachers want all teachers to participate and become assessment literate. I guess one of the major obstacles would just be communicating the assessment system to all of the teachers in the district—trying to get everyone on board, understanding why we're doing it and how it helps students, and why they need to use the data to inform their instruction.

A **teacher** in a district in a **consortium** questioned whether all teachers should participate and become assessment literate:

- ❖ Well, I think at the beginning the obstacle was getting everybody involved. It was hard to get all the schools, I think, to participate equally. Sometimes people would come from one school and not another maybe.

Teachers in districts in **collaborations** and in **individual districts** were overwhelmed by the STARS process as evidenced by the following comments:

- ❖ More and more, teachers are overworked and overwhelmed by the cumbersome process! The state has received and evaluated many assessments in portfolios

which were rated exemplary. It's time the state takes responsibility for creating a test which meets all quality criteria!

- ❖ This has GREATLY increased the stress of the teachers involved in the development and implementation of the assessments, especially in our small district where one teacher is responsible for multiple grades, multiple assessments, and multiple subjects.

Teachers in *collaborations*, *consortia*, and *individual districts* considered a state test the answer to accountability:

- ❖ As a 25-year veteran of teaching, I feel quite strongly that the time, money, and effort spent on STARS would be better used to develop a state-wide assessment used by ALL schools!!!
- ❖ I feel that we would be much better off having statewide tests, as is practiced in Colorado. These would, at least, be developed by professional test writers that know what they are doing and have the training to write valid instruments. We could give these tests in a shorter period of time, and have more time to teach our students.
- ❖ I hope Nebraska NEVER moves towards a standardized test.
- ❖ Nebraska is on the right track. One test does not accurately measure student achievement or provide information that drives instruction.

PRINCIPAL INTERVIEW COMMENTS

Principals in districts in *collaborations* and *consortia* indicated that the amount of time that the process takes is a huge issue:

- ❖ Time is always an issue. Building the assessments into your daily lesson plans, the time of scoring them, and recording all the data that is necessary. Those are some extra things that have been imposed upon certain teachers. But as they continue to get trained, and we continue to do it more than once or twice or three times, they get better at it—more efficient.
- ❖ I know one thing that we've talked about is the length of our school day. It seems like we're still going to school the same length of time that we did 20 years ago but we've added so much more to what we expect students to be able to learn and to do.

Principals in *collaborations*, *consortia*, and *individual districts* indicated that data collection, reporting, and management is difficult:

- ❖ Philosophically, I completely agree with the idea behind the STARS process. The . . . part that's been very difficult is the . . . bureaucratic record keeping.
- ❖ We have four years of data, but lack the management system to provide timely feedback to teachers so it can be used efficiently and effectively to improve student learning and teacher instruction.
- ❖ I think the biggest obstacle was finding a process and a direction that was going to fit for the whole district and getting that in place.

Principals in districts in *collaborations* and *consortia* wanted all teachers to participate and develop ownership for the STARS process:

- ❖ Those people (teachers) need the opportunity to get together so everyone understands this process starts in kindergarten. And this process continues to go for 13 years of these kids' lives. We're all working for the same goal. Once again, seeing the big picture.
- ❖ There's that lifelong obstacle of having certain staff members that don't necessarily buy into the process.

Principals in districts in *consortia* considered a state test the answer to accountability:

- ❖ I would like to see state tests so all schools could be compared on an equal basis.
- ❖ Assessments are good...especially the way we do it in Nebraska because we're measuring what we teach and what we're working on. If we ever go away from that, I'm afraid there will be a lot of teaching to the test and maybe not the knowledge that kids need to have to survive in society.
- ❖ AYP is very cumbersome, but the STARS process seems a better system than the one test approach.
- ❖ How valid are CRTs that are locally generated, locally scored, and locally reported? Where is the check and balance? I am new to Nebraska and find the whole process a waste of time. Why not use a standardized test?

A **principal** in a *collaboration* district was concerned about out-of-classroom teacher time:

- ❖ So we're trying to come up with some creative solutions, but we haven't been real creative because, as you know, if you take teachers out of the classroom, then you've got subs in there which may not be the best thing, either.

A **principal** in an *individual district* was concerned about English as a Second Language student issues:

- ❖ Probably the language barrier of our students is our biggest obstacle; trying to develop assessments that are not biased, but at the correct level of difficulty. If they're (students) coming from a different country or different parts of the United States, they may not understand. So that's been a challenge. . .to develop assessments that allow all of our students to be on an equal playing field.

SUPERINTENDENT INTERVIEW COMMENTS

Superintendents of districts in *collaborations* and *consortia* and superintendents and *assessment coordinators* in *individual districts* indicated that the amount of time and money that the process takes is a huge issue:

- ❖ I think many teachers are frustrated with the amount of work that they have to do and we've got to keep funding coming and we've got to keep support coming. Because otherwise, our teachers are going to throw up their arms and say just give me one test, but let me have it over with.
- ❖ Obstacles we have in school all the time: time and money.

- ❖ Financially, there is the cost to the district, although the STARS money has certainly helped. Finding time to get everybody together and work through some of the issues without taking away too much student time. If we do summertime, then we have to pay for it. So those are the issues.

Superintendents and *assessment coordinators* of districts in *collaborations*, superintendents of districts in *consortia*, and superintendents of *individual districts* indicated that data collection, reporting, and management was difficult:

- ❖ And the other part for me, in working with the other districts is finding out how they're organizing it and how they're keeping their data straight. Because that's becoming a big problem—to have so many assessments.
- ❖ It would be GREAT help if we just had to meet NCLB reporting and not double-up with everything the state requires.
- ❖ We hire a data entry operator through the consortium just to get all that data entered.

Superintendents of districts in *consortia* expressed concerns about technology:

- ❖ I think NDE dropped the ball. And I think it. . .it would have been very, very easy for them to do exactly what I'm getting through this particular consortium with the technology available.
- ❖ Hey, we can do it using technology and make it as easy and meaningful for the schools as possible. Now I think that would have been a role that I would have liked to have seen the Department take. Maybe they've listened and they're starting to do that right now.

Superintendents of *individual districts* wanted all teachers to participate and develop ownership for the STARS process:

- ❖ I think the obstacle is how do you get people involved? That's what we really focused on. How do you make it meaningful for everybody? And as long as we focus on kids, it seems to do that.
- ❖ Continue to get buy-in for some teachers is still a bit of an issue, especially for the teachers that have a little more tenure than some of the others.

A *superintendent* of an *individual district* was concerned about a state test:

- ❖ We continue to encourage leadership of the state to not fall victim, and so far so good. And I think the struggle is going to intensify and we look forward to being part of that.

A *superintendent* of a district in a *collaboration* was concerned about out-of-classroom teacher time:

- ❖ But when you take teachers out of the classroom, it impacts kids. When I covered the writing class yesterday, the superintendent didn't do as good a job as the writing teacher. And I'll be the first to admit that. Oh, we did instruction and I

did enjoy it, but I'm not the best one to learn writing from. When you take a teacher out of the classroom, it hurts kids.

A *superintendent* of a *collaboration* district showed concern about opportunity to learn:

- ❖ We have students who haven't had. . .let's say an advanced math class, and they're going to be assessed on that objective. I don't agree with that. It would be like asking me to fly the space shuttle and not giving me instruction in it. But that's not fair to kids. And the one thing that I really disagree with this approach is that we're trying to have a top end approach. Here's what the best will learn.

A *superintendent* of an *individual district* was concerned about testing years:

- ❖ I guess the only question I would ask. . .we have 12th-grade standards. We're asked to report them at 11th grade. We have to do the testing early in the 11th grade year, some midpoint in the year so we can get them done, get them aggregated and reported.

ESU INTERVIEW COMMENTS

ESU *staff developers* for districts in *collaborations*, *consortia*, and *individual districts* indicated that the amount of time that the process takes is a huge issue:

- ❖ Major obstacles . . . time . . . time for the schools to do a good job.
- ❖ So, when a teacher goes back and says, 'We need more time to be able to do this.' They're (administrators) going to say, 'Oh sure. You always want more time to do whatever it is you have to do.'
- ❖ They've got to have time. They can't always do it at night, after school, because that's not your best time to do good thinking.

ESU *staff developers* for districts in *collaborations* believed that educators should be compensated appropriately for the amount of time spent in the STARS Process:

- ❖ It's a lot of work on their own time. That's why we try to have sessions here so we can pay them from the STARS money, at least for something.

ESU *staff developers* for districts in *collaborations*, *consortia*, and *individual districts* wanted all teachers to participate and develop ownership for the STARS process:

- ❖ I really think that those who've worked collaboratively together have learned together, have unlearned together, have struggled together, really have a neat relationship, and rely on and trust each other more.
- ❖ They take the process into their school and make it their own. They align it carefully with their curriculum. They have their testing map and they have gone through the procedure so that when they change assessments you know they have gone through the reviews by another panel of people who weren't involved in creating them. They make them truly their own.

ESU *staff developers* for districts in *consortia* provided these comments on data collection, reporting, and management:

- ❖ As soon as the kids are done, the teacher knows and the kids know exactly what questions they got right and what questions they got wrong in regard to that specific standard. So all of our schools now are using that information because the next day or the next 20 minutes the teachers can look and see a student is not doing too well in an area.
- ❖ If we would have had this technology when we started five years ago, instruction would be stronger. But we didn't have it five years ago. You know, we were just all kind of learning together and figuring this out. So the technology part for consortium assessment has improved instruction in the classroom.

ESU *staff developers* for districts in *consortia* believed that the issue of retesting students should be addressed:

- ❖ They won't do retakes because the research says no. Well, if you give retakes, your scores go up. And so they're living through, 'We're taking the high road; but why are other school's scores higher?'

CHARTING STARS: VOICES FROM THE FIELD

THEME 10: NEXT STEPS FOR SUCCESS

With a small student population, we need as many as possible of our 23,000 yearly graduates to be good at what they do and to develop into effective participants in the affairs of their communities, state, and nation.

Douglas Christensen, Nebraska Commissioner of Education (2004)

OVERVIEW

This section highlights the next steps for success in the implementation of the STARS process. This theme emerged when the interviews were coded for the purpose of identifying unanticipated themes. The following quotes were collected from grades 4, 8, and 11 language arts and math teachers, principals, superintendents, and assessment coordinators of districts in collaborations, consortia and, individual districts, and from ESU staff developers serving those districts.

TEACHER INTERVIEW COMMENTS

Teachers in collaborations, consortia and individual districts believed all students should have the opportunity to learn:

- ❖ The biggest thing we have to work on here is getting our kids so they can get the material, have the opportunities. And it's kind of a situation where we may have to kind of force them to take those opportunities.
- ❖ The NDE needs to understand that there are students who will never be able to do advanced math classes. I cannot expect my students who do not take classes like Geometry, Algebra II, or Trigonometry to be able to pass assessments that cover these concepts.
- ❖ At the high school level, we're finding it difficult, because so many of our students are in so many different curriculums, to make sure that all the standards are covered somewhere in their education before they graduate.
- ❖ Before they had only required about two years of math for all four years of high school. Maybe a student took one their freshman year and not again until the senior, so they missed out on a couple of years of math before they needed to have the STARS test. So we're going to try and change it where the freshmen, sophomores and juniors are all required to have a math class.

- ❖ I think . . . there still needs to be modifications made for the special needs students. Everything is all paper and pencil and I don't think we are maybe reaching learners of all types yet.

Teachers in *collaborations* and *individual districts* believed that time must be made available for dialogue with all teachers across the curriculum:

- ❖ But I think more time needs to be spent in the schools today giving teachers time to work and practice those intervention styles. . . We can all benefit from learning something about the other person's area and what they use. And I don't think we have the time to develop that enough.
- ❖ But the teachers need to have a day of going over the standards, picking out key words, writing down ideas of what they do in their classroom, then maybe meet with other teachers and start formulating some ideas for tests, and then start writing their tests.

A **teacher** in a ***collaboration*** district thought that all teachers should share the responsibility and accountability for building assessment literacy skills and participating in the STARS Process:

- ❖ I see some of the math, English, and science teachers becoming bitter about the amount of extra work that is being expected of them and not other subject area teachers.

A **teacher** in a ***consortium*** district indicated that ownership in the STARS process is based on educators understanding the big picture:

- ❖ We need a workshop, or a one-day or a half-day in-service on explaining how this all fits together. I'm just working with some pieces here in the corner of this puzzle and there's this master puzzle and I still don't see how all the pieces fit. I'd like to see the bigger picture and the people who work in it everyday. It's become mechanical to them.

A **teacher** in an ***individual district*** believed that all stakeholders should be knowledgeable about the STARS process:

- ❖ The public, again, is not very well educated on this whole process with Nebraska. I think that's the major fall back right now is that they see it in the paper.

Teachers in *collaborations* and *individual districts* stated that policy should be developed statewide for consistency in retesting:

- ❖ Will we retest until they're proficient? Is that really an accurate reflection of students' abilities to continue to give them the same test as we've heard some districts do?
- ❖ As far as kids not passing the standards, we always do a reteach and then a retest. But after that, if kids aren't passing, we really don't have any kind of program set up, as of now, what to do with these kids that aren't passing the standards. I know that has been a big topic that we've talked about for several years, of what do we

do with kids who aren't passing the standards, aren't meeting that proficiency level.

- ❖ This district struggles with retesting, giving the same test over and over and over (maybe 7-15 times) until students pass. We've had discussions and conferences with experts on the ways to get second retake versions developed.

A *teacher* in a district in a *consortium* believed that STARS assessment results should be aligned with district report cards:

- ❖ We have this report card that's comes out of the post-World War II. . . A, B, C, D. And so I guess I'd like training; we need some real training on how to do this. Pull out the stops, bring in the people, and give us time. Show us how we can align our overall recording and assessment of the students to what we're doing, the standards.

Teachers in *individual districts* thought that data reporting and timeliness of receiving data is critical to the STARS process:

- ❖ I think the next step would be to use the data better. I mean, we have the data, it's out there, but what does it mean? And I think really informing the public, too. You get a lot of these results in the paper where they're comparing districts, but districts give different assessments. How can you compare apples and oranges? And that's one thing we need to do, is inform the public better, I think.
- ❖ I think sometimes an outside person doing some scoring might be a little more valid than somebody who's taught the information and then administered the assessment and then also has to score the assessment. But I don't think that's going to happen.

A *teacher* in an *individual district* believed that the use of technology to administer assessments could enhance the STARS process:

- ❖ But I would like to see some other things that were maybe as valid or more valid than what I've seen with the Stars reading tests. I'd like to be able to get that maybe to a computer so it wouldn't have to be done, necessarily, by hand.

A *teacher* in an *individual district* said that we have to constantly assess the STARS process and ensure that we are doing the right things for students:

- ❖ As far as the state level, I think that we maybe need to take a step back and say are these really what we want to do? Are we really assessing what we need to assess? In our district I definitely think it's a 'yes.' That's a huge 'yes.' At the state level I'm not quite sure.

PRINCIPAL INTERVIEW COMMENTS

Principals in *consortia* and *individual districts* believed all students should have the opportunity to learn:

- ❖ Well, there needs to be accommodations for SPED kids and ESL kids because right now there are none.

- ❖ We have added Algebra 1A and Algebra 1B which is kind of a longer version or slower version of Algebra I. Because it's at a slower pace and it's two years rather than one, it's allowed some of our other kids to start getting into algebra that weren't able to before.

A *principal* in an *individual district* stated that time must be made available for professional development and dialogue with all teachers across the curriculum:

- ❖ So I think our staff development will continue to grow. This year we've gone to all-day staff development rather than just afternoons because it allows us to do a more in-depth presentation and then get teachers some classroom time to kind of digest what it's been about. So we're continually making adjustments to try to improve what we do that will help students.

A *principal* in a district in a *consortium* indicated that policy should be developed statewide for consistency in re-testing:

- ❖ I know one thing we are doing if the kids score low on assessments. The teachers are going back and re-teaching that standard and we are also. . .we've debated a lot on whether or not. . .which score to use. We're retesting them all so whether or not to use the first score or the second time. Just how many times we should test them on a standard.

A *principal* in a *collaboration* district believed that assessment should be embedded into classroom instruction:

- ❖ I think we need to relook at some of the chapter tests that we give. How can we embed our standards into this chapter test so students don't necessarily have the mindset that I'm taking 12.1.1.? I don't think students should ever know, necessarily, that they're doing a state standard.

Principals of districts in *collaborations*, *consortia* and *individual districts* said that we have to constantly assess the STARS Process and ensure that we are doing the right things for students:

- ❖ We haven't done a lot with parent volunteer in-servicing, yet. But that's something that I think we're going to need to start looking at and doing because so many of these kids, especially at my level, need intense intervention one-on-one.
- ❖ We've started something, I believe, having it all online and the consortium that we work with has standardized the assessments in this region. I think that's wonderful because you have students that are moving from one school to another, and standardizing them would be very, very beneficial.
- ❖ Lengthen the school year and have more times in the middle-two or three days-where you can work on some things in between and lengthen the school year that we're here.
- ❖ I think our next step is right where we're going. And that's connecting those assessments to learning, connecting . . . making that big picture just even stronger.

- ❖ We've got to address more issues like bullying and harassment within the school day. You can't do a Friday afternoon in-service and then expect it to take hold with kids. It's something you have to weave within to the curriculum.

SUPERINTENDENT INTERVIEW COMMENTS

Superintendents of districts in *collaborations* indicated that all teachers should share the responsibility and accountability for building assessment literacy skills and participating in the STARS process:

- ❖ I think we need to spread it out across the... across the grade levels, and that's hard to do. But I think we need to share the wealth instead of 4th, 8th, and 11th grade teachers getting hammered. I think we need to spread it out. I think we need to provide time in our calendar where teachers can work on this, yet aren't away from students.
- ❖ So there needs to be a shared responsibility of other disciplines on how are they helping with reading in their content area, writing in their content area. What math do they have? What science? What social studies? It should be . . . we were talking collaboration, I think . . . and there needs to be collaboration amongst all the disciplines at the high school level.

Superintendents in *collaborations* and *individual districts* believed that time and funding must be made available for professional development and dialogue with all teachers across the curriculum:

- ❖ I think many teachers are frustrated with the amount of work that they have to do. We've got to keep funding coming and we've got to keep support coming. Because otherwise, our teachers are going to throw up their arms and say just give me one test, but let me have it over with.
- ❖ There has been talk that we have 186 contract days, that maybe we need to make that 190 and require staff members to come in during those four days. And, of course, they would be compensated accordingly; however you wanted to work that on your salary schedule.
- ❖ The time issue keeps coming back to my mind when you ask that question. How can we create a school year that will not only allow for time? The time's there, but we live in a time, also, when staff members want to be compensated for that extra time.

Superintendents and *assessment coordinators* of districts in *consortia* and *individual districts* thought that assessment should be embedded into classroom instruction:

- ❖ I think incorporating the assessments more as a part of an actual classroom instead of making them seem like they exist in a vacuum, I think that's probably the next step that we need to move to.
- ❖ Probably the main thing is going ahead and having the assessments be a part of the every day-to-day learning where we have the curriculum in place already. The assessments should be a natural part of the classroom learning. But at the

very beginning, I think it was almost an add-on, but we are moving away to that as just a natural part of teaching and learning. So those were the early stages.

Superintendents of individual districts said that data reporting and timeliness of receiving data is critical to the STARS process:

- ❖ They've developed their assessments pretty well. But maybe they'll need now to have some information on looking at the data and developing action plan interventions based on the data that they have received. So I would say that is our next step.
- ❖ I think we've only begun to understand how to read data. And data for the most part has scared us because we didn't really know how to make as much out of it as we could. Better training in how to use data to make good, sound educational decisions is needed.

A **superintendent** of a district in a **collaboration** shared that policy should be developed statewide for consistency in re-testing:

- ❖ So I think our next step, and we've talked a lot about this as an administrative team, is how do we reteach and retest? That's really our next step because we really haven't re-taught or re-tested. We need to go to that next level.

Superintendents of individual districts noted that STARS assessment results should be aligned with district report cards:

- ❖ I think we're looking a little bit at grading practices, as I mentioned earlier, and how they relate to the STARS movements. So that's becoming a bit of a topic of conversation. We've already started that a little bit. So we need to continue to look at that. Average yearly progress . . . it doesn't have to do with the STARS, necessarily, but that's going to create some issues for us, average yearly progress and making sure we meet the standards.
- ❖ Now we're looking at grading practices as it relates to the standards, movement of standards issues.

Superintendents in collaborations, consortia and individual districts believed that we have to constantly assess the STARS process and ensure that we are doing the right things for students:

- ❖ If the Nebraska Department of Education would ever think of revisiting the number of standards schools have . . . we have more science standards and more social studies standards than we actually have to report to the state. I wonder if they'd think of revisiting that for math.
- ❖ I think we need to be on the lookout for those things that we can be more efficient and streamline the process. I guess I'm always looking for better ways to do things. What is the next step? I think we're still doing a dual track assessment system. We're using the standards-based assessment to satisfy the governmental issues but, for individual student indicators, many teachers are still using the chapter tests and pop quizzes.

- ❖ Are there ways that we can reduce the workload, reduce the number of steps, and stay within the spirit of STARS?
- ❖ If STARS and the Department of Education really want to do a good job, they would start bringing teams of people out to districts, look at the system, look at the process, and give people legitimate feedback on how they can improve things.

ESU INTERVIEW COMMENTS

ESU *staff developers* provided the following comments regarding Next Steps for Success:

- ❖ To move forward, we need to continue to build the background in the leadership piece, in the instructional leadership piece. And I'd like to see that happen on a higher level of instruction coming out of our teachers colleges, teacher training places, because I don't know that we totally get it.
- ❖ I think they have to continue to make their public, their stakeholders, be aware of what is happening on those time outs when they're out of school and when there are early outs and things. People need to understand that the teachers have to have the time to work.
- ❖ The STARS updates are overwhelming? Can a summary be developed often for new leaders coming to Nebraska?
- ❖ I like the idea of the NDE doing assessment monitoring through school improvement. External visits and the portfolio are a huge undertaking and continued streamlining and modification is needed. This can be part of the external visits; it will require additional instruction for external team leaders and members.

Primary Research Study References

- Christensen, D. D. (2004). *Schools for our future: Part two: Why do we need them?*
Retrieved on July 20, 2005, from <http://www.nde.state.ne.us/COMMISH/febart2.html>
- DuFour, R., DuFour, R., Eaker, R. & Karhanek, G. (2004). *Whatever it takes: How professional learning communities respond when kids don't learn.* In Foreword, "The Power of Collective Intelligence" (p. xvi) by M. Schmoker, 2004.
Bloomington, IN: National Education Service.
- Fitzpatrick, K. (2003). *System-wide improvement: Focusing on student learning.*
Schaumburg, IL: National Study School Evaluation.
- Kouzes, J. M. & Posner, B.Z. (2002). *Leadership: The challenge.* Jossey-Bass: San Francisco, CA.

APPENDIX a
RESEARCHERS FOR THE PROJECT

**Researchers for the Primary Fourth Year Project
2004-2005**

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 a graduate student at the University of Nebraska-Lincoln. She currently serves as the Coordinator for the Nebraska State Writing Assessment.

Toby Boss is completing his dissertation in partial fulfillment of the requirements to obtain his Doctorate in Educational Administration from the University of Nebraska-Lincoln. He currently serves as a staff developer at Educational Unit 6.

Dr. Wanda Clark received her Doctoral degree in Educational Administration and Supervision from the University of Nebraska at Omaha. She currently serves as the Instructional Research Administrator for Omaha Public Schools.

Dr. Leon Dappen is an Assistant Professor in Educational Administration at the University of Nebraska at Omaha. He received his Doctoral degree from the University of Nebraska-Lincoln and is a licensed psychologist (inactive). He has been a teacher, counselor, psychologist, special evaluation administrator, and assistant superintendent in pre-K through 12 schools. Dr. Dappen is an external evaluator for the Omaha Public Schools Magnet Assistance Grant and Banneker CEMS Math/Science grant. He is actively involved with several schools in their North Central Association School Improvement activities.

Dan Endorf is completing his dissertation in partial fulfillment of the requirements to obtain his Doctorate in Educational Administration from the University of Nebraska-Lincoln. He currently serves as the principal of York High School.

Tammy Heflebower is completing her dissertation in partial fulfillment of the requirements to obtain her Doctorate in Educational Administration from the University of Nebraska-Lincoln. She currently serves as the Director of Staff Development at Educational Service Unit 6.

Shirley J. Mills taught in Nebraska for 38 years prior to working towards her Doctorate in Educational Administration from the University of Nebraska-Lincoln. She is currently working as a graduate assistant with the NCA/UNL School Improvement Specialist Program and the STARS research project.

Phil Warrick is completing his dissertation in partial fulfillment of the requirements to obtain his Doctorate in Educational Administration from the University of Nebraska-Lincoln. He currently serves as Principal of Waverly High School.

Susan Wilson is on the staff of the University of Nebraska-Lincoln and holds an Associate of Science degree in Business Administration from the College of St. Mary, Omaha. She is currently pursuing a Bachelor's degree.

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.

Robert Whitehouse, Ed. Specialist, retired in 1999 after 35 years as a teacher, coach, athletic director, junior high principal, and senior high principal with the Omaha Public Schools. He received the "Outstanding Principal of the Year" in 1997, the "National Milken Family Foundation Award" in 1997, and the "Nebraska Council of School Administrators Distinguished Service Award" in 2000. Bob has served in many capacities on local and state boards in both the private and public sectors. He remains active in education as the coordinator of the Nebraska Educational Leadership Institute sponsored by the Nebraska Council of School Administrators, the University of Nebraska-Lincoln, and the Gallup Organization.

APPENDIX B
EFFECTIVE LEADERSHIP FOR LEARNING

Leadership For Learning
Dr. Jody Isernhagen, University of Nebraska – Lincoln
Dr. Leon Dappen, University of Nebraska at Omaha

Introduction

Wong (2003) in reviewing over 200 studies found strong support for a knowledgeable and skillful teacher being the only way to improve student achievement. However, “while quality teaching is obviously where the rubber meets the road, such teaching cannot be insured in all classrooms for all students without skillful leadership” (Sparks, 2002, p. ii-2). It is increasingly being recognized that leadership, both distributed leadership in learning communities and building/district leadership are also key to improving student achievement (Marzano, 2003).

Student achievement increases in schools where collaborative work cultures foster a professional community among teachers and others (Katzenmeyer & Moller, 2001). Administrators must take a proactive approach by developing a learning community that requires continuous reflection about student learning and takes action on these reflections. Researchers have found that both teachers and school leaders place more value on staff development if the training is clearly linked to improved student learning (Holloway, 2004).

The time has come to take advantage of this need and rethink the relationship between assessment practices and effective schools (Stiggins, 2004). Since the 1960’s we have been dependent upon the path to school improvement being paved with more and better standardized tests. This has blinded school leaders to a different application of assessment – classroom assessment – that has been shown to support remarkable gains in student achievement. School leaders that focus on a vision of classroom assessment for learning contributing to the development of effective schools have largely been ignored.

Nebraska has built a School-based Teacher-led Assessment and Reporting System (STARS), which required each of Nebraska’s 495 school districts to develop a local assessment system to measure student performance on local standards. This process honors educators and trusts their professional judgment but also demands work and a great deal of leadership based on a clear vision for learning. For the last five years, educators in Nebraska have been deeply involved in development and implementation of the STARS model. The effort has dominated the attention, staff development activities, general professional interest and discussions, board of education interest, and significant related financial resources. The efforts have been a primary focus for the Nebraska Department of Education, Educational Service Units, and local districts.

Purpose of the Study

The purpose of this study was to survey perceptions of the leaders involved to examine the status of leadership efforts regarding development and implementation of STARS. The research questions for this study were:

1. What are the perceptions of Nebraska school leaders regarding their ratings of:
 - The amount of standards/assessment/accountability training they have had during their career;
 - Their amount of growth in assessment knowledge over the last five years;
 - The districts use of assessment data for improved student achievement;

- Alignment of their vision of assessment with other leaders in the district/school;
 - The impact of standards/assessments/accountability efforts on teaching and learning in the district/school?
2. What are the perceptions of Nebraska school leaders regarding their ratings of their district/school on the Stiggins Assessment Self-Analysis six doors:
 - Use of achievement standards in the district/school;
 - District commitment to all users of assessment data;
 - Status of development of assessment literacy;
 - Effective communication;
 - Creation of supportive assessment policies;
 - District/school leadership for assessment reform?
 3. What are the implications to the Nebraska STARS program from these findings?

Methodology

A survey was sent to board of education presidents, superintendents, assessment coordinators, elementary principals, and secondary principals in all 495 K-12 school districts in Nebraska. The survey was adapted from an instrument developed by Richard Stiggins. Stiggins has used the instrument in a number of workshop and training sessions across the state as contracted by the Nebraska Department of Education.

The survey used by Stiggins described five general areas, or “doors” of assessment practice. They include: Door 1-use of achievement standards; Door 2-commitment to all users of assessment data; Door 3-developing assessment literacy; Door 4-effective communication; and Door 5-creating supportive assessment policies. The survey culminates with a section on “leadership for assessment reform,” which was included as Door 6 for the instrument used in this study. Leaders were asked to rate items on the survey on a five-point Likert scale with five being the highest and one being the lowest.

Five questions were added to the Stiggins instrument using the same five-point scale. They include (1) Rate the amount of standards/assessment/accountability training that you have had during your educational career, (2) Rate your growth in assessment knowledge over the last 5 years, (3) Rate your use of assessment data for improving student achievement, (4) Rate your vision of assessment with the vision of assessment of other leaders within your school/district, and (5) Rate the impact of standards/assessment/accountability efforts to date on teaching and learning within your school/district. Wording was altered slightly as appropriate to the respondents’ role. Instruments were mailed out in the early fall of 2004 and returned by mid-October.

Results

There were 578 survey respondents, which included 129 superintendents, 51 board of education presidents, 119 assessment coordinators, 210 elementary principals, and 170 secondary principals. The sample included 366 males and 212 females. The years of experience ranged from 282 having one to five years, 127 having six to ten years, 129 having eleven to twenty years, and 61 having over 21 years of experience.

Participants Rating of Standards/Assessment/Accountability Training During Career

All groups' average ratings except school board presidents were between 3 and 4 (3.6 average) on the 5-point Likert scale, indicating, "some" to "quite a bit" of training. School board president's average rating was 2.2, indicating "little" training. Assessment coordinators and elementary principals rated amount of training significantly higher than superintendents. Secondary principals overlapped assessment coordinators, elementary principals and superintendents, indicating a wide range in the amount of training received by individual secondary principals.

Participants Rating of Growth in Assessment Knowledge over the Last 5 Years

The average rating for assessment coordinators was above 4.6, indicating "extensive" growth, significantly higher than all other groups. Elementary and secondary principals' ratings were 4.2 and 4.1 respectively, indicating "quite a bit" of growth, significantly more so than board members. Superintendents were lower than principals (3.9), however not significantly so, still indicating "quite a bit" of growth, significantly stronger than board presidents. While board presidents were significantly lower in growth than all other groups, their average rating was 3.2, indicating "some" growth.

Participants Rating of Use of Assessment Data for Improved Student Achievement

Ratings for principals, assessment coordinators, and superintendents were between 3.7 and 4.0, indicating "quite a bit" of use of data for improved student achievement. Board presidents' rating was 3.0, indicating "some" use of data for improved achievement.

Participants Rating of Alignment of Vision with Other Leaders in District

All groups indicated alignment with others with ratings ranging from 3.7 to 4.0, indicating "quite a bit" of alignment. Elementary principals and superintendents rated highest in alignment, significantly higher than assessment coordinators but not than other groups.

Participants Rating of Impact of Assessment Efforts on Teaching and Learning

All groups rated impact between 3.7 and 4.1, indicating "quite a bit" of impact. Elementary principals rated impact significantly higher than superintendents, with no significant differences with other groups.

Participants Ratings on Stiggins Six Doors of Assessment Leadership

Significant differences in ratings on Stiggins Six Doors of Leadership, (1) use of achievement standards, (2) commitment to all users of assessment data, (3) development of assessment literacy, (4) effective communication, (5) supportive assessment policies, and (6) leadership for assessment reform were examined. Significant differences were found in elementary principals being significantly stronger than other groups in "use of achievement standards," "effective communication," "supportive assessment policies," and "leadership for assessment reform."

Mean ratings for the six doors across all groups on the 5-point Likert scale were 3.9 for "leadership for assessment reform," 3.7 for "use of achievement standards," 3.6 for both "development of assessment literacy" and "commitment to all users," 3.4 for "supportive policies," and 3.3 for "effective communication."

Mean ratings on the total scale for the groups on the 5-point Likert scale were 3.7 for elementary principals, 3.6 for superintendents, 3.5 for secondary principals, and 3.4 for assessment coordinators and board presidents.

Discussion and Implications

A survey of Nebraska board of education presidents, superintendents, assessment coordinators, elementary principals, and secondary principals indicated that data was being used to impact achievement and that standards/assessment/accountability efforts are having a strong impact on teaching and learning. All report having had at least some to extensive related training in their career, particularly in recent years. The survey participants indicated a common vision of assessment with other district leadership. There is a strong base on which to continue the development of the Nebraska STARS program.

On a survey examining areas most important to the support for these efforts, the groups rated leadership highest, followed by use of standards in achievement, then commitment and assessment literacy, and finally supportive policies and effective communication. Those involved with the continued development of the STARS process should recognize that while all components are necessary, it is leadership that is the key to continued success.

The single highest item was in groups rating of gain in assessment knowledge in the last five years being “quite a bit” to “extensive.” Clearly, there has been significant gain in assessment knowledge and all groups recognize it.

While all ratings were similar, elementary principals and assessment coordinators were strongest in their ratings across all areas, followed by superintendents, secondary principals, and then board presidents. This may be a result of elementary principals and assessment coordinators generally having been most involved and affected in the early stages of the STARS process. Secondary principal ratings were the most varied, in their ratings. This may reflect individual interest or variance in district requirements for their involvement in the process. It is important to note that while board of education presidents were lowest in their ratings on all items, they still had an average rating across all areas of 3.2 on a 5-point (with 5 as high) Likert scale. They perceived that they have been involved, were knowledgeable, and aligned with other school leaders regarding the STARS program.

It should be noted that while the sample included a large number of respondents; it was not a random sample and therefore, may not be representative of the populations identified. It is not known if any respondent bias based on this sample would be more or less positive or negative or varies between any of the populations.

Nebraska school leaders have adopted a new vision to develop assessment literacy for improvement of student achievement. Roschewski (2002) summarized a study of promising practices for building quality local assessments by identifying the primary recommendation as the vision being endorsed by leaders and shared throughout the entire school system. This study supported the focus on leadership.

Leaders have demonstrated their commitment to supporting professional development as a part of their vision for assessment literacy for improving student achievement. Quality professional development contributed to the development of teachers who embedded assessment literacy in their instructional practice and collaborated to improve student performance. This professional development has resulted in improved assessment literacy and was perceived by leaders as improving student achievement. Student achievement data reflects this gain in reading, writing, and math (Dappen & Isernhagen, 2005; Gallagher,

2004). This success bodes well for continued efforts, as more staff and curriculum areas are included in the Nebraska STARS program.

References

- Dappen, L. & Isernhagen, J. (2005). STARS: Assessment for Learning. Presentation at 25th Annual Excellence in Education Conference, Nebraska Department of Education, April 21, 2005.
- Gallagher, C. (2004, August). Charting STARS: New Conversations. Year Three Report.
- Holloway, J.H. (2004). *Research Link: Linking professional development to student learning*. Alexandria, VA: ASCD.
- Katzenmeyer, M. & Moller, G. (2001). *Awakening the sleeping giant*. Thousand Oaks, CA: Corwin Press.
- Marzano, R.J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: ASCD.
- Roschewski, P. (2002, December). Promising practices, processes, and leadership strategies: Building quality local assessment. An Executive Summary. Unpublished document.
- Sparks, D. (2002). *Designing powerful professional development for teachers and principals*. Oxford, OH: National Staff Development Council.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86 (1), 22-27.
- Wong, H.K. (2003, October). Collaborating with colleagues to improve student learning. *ENC Focus Review*, 8-11.

APPENDIX C
2002-2004 NEBRASKA CRITERION REFERENCED
tests

**An Examination of 2002 to 2004 Nebraska Criterion Referenced Tests,
Norm Referenced Tests, and District Portfolio Ratings for Math at Grades 4, 8, and 11
Dr. Leon Dappen, University of Nebraska – Omaha
Dr. Jody Isernhagen, University of Nebraska-Lincoln**

Introduction

With the emphasis to demonstrate increased achievement for all students and all schools to be accountable to their constituents, many states have developed or adopted state mandated tests to assess student academic performance. Nebraska stands alone, rated by the 1999 Education Week Report Card as a “C” largely because it does not measure school performance by a statewide mandatory test. According to the report, the state is “lagging behind” in accountability (Editor, 1999). Yet, this seems incongruent as the children in Nebraska rank among the top 10 nationally in most generally accepted measures of academic success (Editor, 1999). Instead Nebraska has made a conscious decision to lead the way in a new system of accountability that focuses upon building assessment literacy among educators and enhancing student performance through the use of a quality assessment system.

Nebraska’s Quality Assessment System

The Nebraska School-based Teacher-led Assessment and Reporting System (STARS) is unique in several ways. As described by Gallagher (2004) this system:

- is a system of local assessments, not a state test;
- promotes a balanced approach to assessment, using multiple measures;
- involves evaluation of achievement and of assessment quality;
- uses classroom-based assessments for reporting; and
- includes no high-stakes testing.

In the STARS process, districts first adopt local or state standards. Districts then submit an assessment plan that includes locally developed criterion referenced tests (CRTs) to assess the district’s standards at the identified grade levels. They may also choose to use some items of their locally chosen norm referenced test (NRT) to assess some of their standards. The districts then submit a District Assessment Portfolio to the Nebraska Department of Education. Portfolios are reviewed and rated by independent experts using six quality criteria outlined by the Buros Center for Testing (Plake & Impara, 2000):

1. Assessments reflect state or local standards.
2. Students have an opportunity to learn the content.
3. Assessments are free from bias or offensive language or situations.
4. The level is appropriate for students.
5. There is consistency in scoring.
6. The mastery levels are appropriate.

The focus of the STARS assessment system is to keep teaching and learning at the center of the educational process, promoting high-impact, not high-stakes assessment (Gallagher, 2004).

STARS began in 2000 and the first scores were released publicly in the State of the Schools Report by the Nebraska Department of Education on their website and in newspapers across the state in the fall of 2004 in the area of reading. The Nebraska

Department of Education collects district and school data including: the percent of students meeting local defined proficiency for criterion referenced assessments; the average percent of students in the top two quartiles on the district chosen norm referenced test; and the District Assessment Portfolio rating, a summary based on scores from the six quality assessment criteria.

Purpose of the Study

This study examined the first achievement data available in math for the Nebraska STARS program, comparing the reading scores in grades four, eight, and eleven from 2002 to scores for 2004. Both locally developed CRT scores and NRT scores were examined. The research questions for this study were:

1. What are the differences in the CRT math scores in grades four, eight, and eleven from 2002 to 2004?
2. What are the differences in the NRT math scores in grades four, eight, and eleven from 2002 to 2004?
3. What are the differences in the math district assessment portfolio ratings in grades four, eight, and eleven from 2002 to 2004?
4. Based on this data, what are the implications for Nebraska STARS?

Methodology

School Districts Studied

Data were included for Class 3, 4, and 5 school districts. Class 3 school districts (209 school districts) are represented by any school district with territory having a population of more than 1000 but less than 150,000 inhabitants (Nebraska Education Directory, 2003-2004). Class 4 school districts (Lincoln only) are represented with a territory having a population of 100,000 or more inhabitants with a city of the primary class within the territory (Nebraska Education Directory, 2003-2004). Class 5 school districts (Omaha only) are represented 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, 2003-2004). 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.

Score Definitions

For CRTs, the score used was the percentage of students meeting the district defined mastery level for their locally developed measure. For NRTs, the score used was the percent of students scoring in the top two quartiles on the nationally standardized test used by that district (e.g., California Achievement, Iowa Test of Basic Skills, etc.). While the NRT used will vary, the data reported (percent of students in the top two quartiles) was constant for all districts. The score used for the reading District Assessment Portfolio was the summary score assigned by the external expert visiting team based on the previously mentioned six quality assessment criteria. Summary portfolio rating categories are “Unacceptable,” “Acceptable,” “Good,” “Very Good,” and “Exemplary” (Plake & Impara, 2000).

District Portfolio Process

Districts complete the portfolio each spring, including requested information supporting each of the six rating criteria. The portfolio is then submitted to the Nebraska

Department of Education in the summer of each year. A contract evaluation agency, Buros Institute for Assessment Consultation and Outreach, arranges for a panel of external reviewers, the District Assessment Evaluation Team. This team is comprised of professionals with an earned doctorate in educational measurement and come from Nebraska and several other states. A rubric has been developed for each of the six quality criteria including the qualifications of those involved, the quality of the process used, and the results reported. The external review team is trained in the rubric review process with significant attention given to inter-rater reliability. Each of the six criteria is rated “Not Met,” “Met - Needs Improvement,” or “Met.” These ratings are then summarized as prescribed by the developers of the system for the reading District Assessment Portfolio score.

Data Analysis

The unit of analysis for this study was the composite of class 3, 4, and 5 school districts for the state of Nebraska in math at grades four, eight, and eleven. Based on the fact that CRT scores were unique to each district, NRT scores varied with the standardized test chosen by that district, and portfolio ratings were based largely on the CRT measures, the data was described as unconventional. Traditional statistical approaches for significance, therefore, were not appropriate. Descriptive data were reported and discussed.

Results

As shown in Table 1, the percentage of students attaining proficiency on the math CRT increased at every grade, fourth (6.58%), eighth (6.90%), and eleventh (6.10%). NRT scores increased at two grades, fourth (3.16%) and eleventh (.53%), and decreased at eighth grade (-.69%).

Table 1: Changes in Criterion Referenced and Norm Referenced Assessments in Math

Grade	Criterion Referenced			Norm Referenced		
	2002	2004	Diff.	2002	2004	Diff.
4	78.30%	85.15%	+6.85%	68.16%	71.32%	+3.16%
8	68.53%	75.43%	+6.90%	67.34%	66.65%	-.69%
11	66.22%	72.32%	+6.10%	67.59%	68.12%	+0.53%

On the math district assessment portfolio, the median score, assigning a score of “1” for Unacceptable, “2” for Acceptable, “3” for Good, “4” for Very Good, and “5” for Exemplary increased at each grade level. Portfolio ratings increased from 2002 to 2004 at each grade, fourth from 3.98 to 4.67; eighth 3.96 to 4.77; and eleventh 3.96 to 4.77. This indicated a strong general increase in applications of assessment knowledge as measured by the portfolio process.

Discussion

Differences in CRT Scores

There were strong gains in the average percent of students mastering the math CRT scores from 2002 to 2004 at grades 4, 8, and 11. While gains in reading last year revealed strong gains at 4th grade, gains at 8th and 11th grades were smaller. This consistent and strong gain in CRT math scores may indicate an increase in success by teachers in applying their STARS training. Students demonstrated this gain most strongly on measures developed to reflect the standards at their level rather than on the norm referenced tests.

It will be important to watch the longitudinal progress of these 4th and 8th grade students as they move forward and continue to be assessed using criterion referenced measures to assess mastery of the math standards at the 8th and 11th grade levels. We will also need to examine other academic areas, as they are included in the system.

Differences in NRT Scores

While the NRT used varies from district to district, the NRT score being examined is the same. There were gains in the percent of students in the top two quartiles of the NRT score from 2002 to 2004 for the 4th and 11th grades. The 8th grade percent of students in the top two quartiles of the NRT score from 2002 to 2004 was slightly lower. There was, however, an overall slight increase in the total percentage of students in the top two quartiles. While the increase is not striking, it is definitely something to watch, as large group NRT scores are generally quite resistant to change. Once again, we must continue to examine the longitudinal effect of these scores and other academic areas, as they are included. It is important to note that, as was true for reading a year ago, the STARS process has not negatively affected NRT scores.

A study conducted by the Buros Center for Testing at the University of Nebraska in 1998 revealed that the five primary standardized achievement tests match only 35% - 40% of the Nebraska State Standards (Roschewski, Gallagher, & Isernhagen, 2001). It would, therefore, not have been surprising to see NRT scores decline somewhat through the local standards and CRT process. The attention paid to CRT assessment activities did not appear to have a negative impact on NRT scores in this study.

Differences in Portfolio Ratings

There were strong gains at all grade levels in the math District Assessment Portfolio ratings from 2002 to 2004. The portfolio is developed each year by local district staff responding to the previous year's external rater comments and any training and improvement activities that have been carried out over the year. This increase would indicate that district staff has increased their knowledge in assessment and its use in instruction and importance in increasing student achievement. This would be expected as the district assessment portfolio is developed based on the six quality criteria that are used by schools when developing their criterion referenced measures and also used when measuring the quality of the district assessment portfolios. The curriculum standard, instruction, and assessment activities tie together to inform each other and build a continuously self-improving educational program. This support for the viability of the portfolio process is the "linchpin" for the STARS system.

Implications for Nebraska STARS

The gains in reading last year, and the strong gains in math this year, continues to build the base of support for the district assessment portfolio process in establishing the creditability of the school-based teacher-led assessment and reporting system (STARS). This is particularly important in Nebraska's accountability system, as student performance is not based on a common statewide measure but on locally developed criterion referenced assessments.

The very nature of STARS and the assessments used makes comparisons very difficult. There is a need to work with the educational measurement community to examine new ways to determine the success of the model. While approaches such as STARS are

strongly supported by the measurement community from a standpoint of student learning, little has been done to support the acceptance of what is referred to as unconventional data.

Much more work is needed in continuing to examine longitudinal results and the impact of refinements and corrections as Nebraskans problem solve concerns and the program evolves. One of the biggest challenges will be the ongoing effort to work with the United States Department of Education to ensure the STARS program will allow Nebraska to meet the No Child Left Behind (NCLB) requirements.

This study supports the continuation of the STARS system. As noted by Roschewski, Gallagher, and Isernhagen (2001), the STARS plan brings together the best of both worlds; student learning is foremost, but public accountability is provided as well. While avoiding the “unintended negative outcomes” of high-stakes assessments, STARS places the responsibility of teaching and learning where it belongs. The teachers and administrators who create high quality learning environments are directly involved in affirming the quality and in measuring the learning of students in their classrooms.

References

- Editor (1999, March 1). Quality Counts-State of the States Report. *Education Week*. Retrieved May 26, 2004, from <http://www.edweek.org/sreports/qc99>
- Gallagher, C. W. (2004, January). Turning accountability tables: Ten progressive lessons From one “backward” state. *Phi Delta Kappan*. 251-261.
- Nebraska Education Directory (2003-2004). *Nebraska Department of Education Directory*. Lincoln, NE. 106th Edition, 170.
- Plake, B. S. & Impara, J.C. (2000). Technical quality rubric for Nebraska’s district performance ratings. Lincoln, NE: Buros Center for Testing.
- Roschewski, P., Gallagher, C. & Isernhagen, J. (2001, April). Nebraskans reach for the STARS. *Phi Delta Kappan*. 611-615.

APPENDIX D
STUDENT PERFORMANCE ON NEBRASKA'S STARS
WRITING

A First Look at Student Performance on Nebraska's STARS Statewide Writing Assessment at Grades 4, 8, and 11

Dr. Leon Dappen, University of Nebraska – Omaha

Dr. Jody Isernhagen, University of Nebraska-Lincoln

Introduction

With the emphasis to demonstrate increased achievement for all students and all schools to be accountable to their constituents, many states have developed or adopted state mandated tests to assess student academic performance. Nebraska stands alone, rated by the 1999 Education Week Report Card as a “C” largely because it does not measure school performance by a single statewide mandatory test. According to the report, the state is “lagging behind” in accountability (Editor, 1999). Yet, this seems incongruent as the children in the Nebraska rank among the top 10 nationally in most generally accepted measures of academic success (Editor, 1999). Instead, Nebraska has made a conscious decision to lead the way in a new system of accountability that focuses upon building assessment literacy among educators and enhancing student performance through the use of a quality assessment system.

Nebraska's Quality Assessment System

The Nebraska School-based Teacher-led Assessment and Reporting System (STARS) is unique in several ways. As described by Gallagher (2004) this system:

- is a system of primarily local assessments, not a single state test;
- promotes a balanced approach to assessment, using multiple measures;
- involves evaluation of achievement and of assessment quality;
- uses classroom-based assessments for reporting; and
- includes no high-stakes testing.

In the STARS process, 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. For the area of writing, a statewide criterion referenced writing assessment has been developed based on the six trait writing approach. Districts may also choose to use items from their locally chosen norm referenced test (NRT) to assess some of their standards. The districts then submit a District Assessment Portfolio to the Nebraska Department of Education. Portfolios are reviewed and rated by independent experts using six quality criteria outlined by the Buros Center for Testing (Plake & Impara, 2000):

1. Assessments reflect state or local standards.
2. Students have an opportunity to learn the content.
3. Assessments are free from bias or offensive language or situations.
4. The level is appropriate for students.
5. There is consistency in scoring.
6. The mastery levels are appropriate.

The focus of the STARS assessment system is to keep teaching and learning at the center of the educational process, promoting high-impact, not high-stakes assessment (Gallagher, 2004).

Nebraska's Statewide Criterion Referenced Writing Assessment

Based on previous involvement by a number of local 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 Nebraska Department of Education works with area Educational Service Units and local districts to carry out the assessment at the same time across the state. A panel of teachers develops, refines, and pilots the prompts to be used at the 4th, 8th, and 11th grades. Trained teachers using rubrics developed for that grade level holistically score the writing assessments. In the first three years, scoring was done at three sites across the state. To improve reliability, scoring is now done at one site. The scoring process and examination of results is carried out by the Buros Center for Testing. A sample is 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 and to the public in the fall. Local district and individual school data shared includes: the percent of students meeting proficiency for criterion referenced assessments (any locally developed as well as the statewide writing assessment); the average percent of students in the top two quartiles on the district chosen norm referenced test; and the District Assessment Portfolio rating, a summary based on scores from the six quality assessment criteria. Because the statewide writing test is developed and administered by the Nebraska Department of Education, districts do not complete an Assessment Portfolio in this area.

Purpose of the Study

This study examined the first achievement data available for the Statewide Writing Assessment for Nebraska STARS, comparing scores available at grades 4, 8, and 11. The research questions for this study were:

1. What are the average percentages of students rated proficient in grades 4, 8, and 11 on the State Writing Assessment?
2. What are differences in district writing scores in grades with pre/post data available?
3. Based on this data, what are the implications for Nebraska STARS?

Research Design and Methodology

School Districts Studied

Data were included for Class 3, 4, and 5 school districts. Class 3 school districts (215) are represented by any school district with territory having a population of more than 1000 but less than 150,000 inhabitants (Nebraska Education Directory, 2003-2004). Class 4 school districts (Lincoln only) are represented with a territory having a population of 100,000 or more inhabitants with a city of the primary class within the territory (Nebraska Education Directory, 2003-2004). Class 5 school districts (Omaha only) are represented 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, 2003-2004). 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.

Score Definitions

Two trained teachers using rubrics based on the six trait writing process holistically score each writing assessment. If there is more than a two-point difference a third scoring is done. The scoring range is from 1 to 4 in + and – intervals resulting in a 10-point scale. The final score is the composite of the two individual scores. The Buros Center for Testing analyzes the data and determines the cut score. The score reported in this study is the average percentage of students in districts across the state that scored above the statewide cut score for that year.

Data Analysis

The unit of analysis for this study is the composite of Class 3, 4, and 5 school districts for the state of Nebraska in writing at grades 4, 8, and 11. While this statewide assessment takes 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 is clearly not a comparison with a separate norm group. Descriptive data was therefore reported and discussed. However, because the assessment is a common measure across districts and is an equal interval scale, inferential statistics were also used to examine statistical significance between pre/post scores.

Results

Significant gains were made at the 4th and 8th grades where pre/post scores were available. At 4th grade, there was a 6.85% gain in the percentage of students demonstrating proficiency in writing, in the 8th grade a gain of 6.90% (Table 1). Both were significant beyond the .001 level. Eleventh grade post score data were not yet available.

Table 1: Statewide Writing Assessment Results

Grade	2002	2003	2004	Diff.	Sig.
4	76.29%		80.70%	+4.41%	P < .001
8		76.27%	85.46%	+6.18%	P < .001
11			85.15%		

Discussion

Fourth (4th) and 8th grades made significant gains in the first pre/post score comparisons on the Nebraska STARS Statewide Writing Assessment. This is very positive for the STARS process and would indicate strong support for continuation of the program. It is interesting to note a difference in the two language arts measures. Reading, which was the first area to be tested in 2001 and 2003 revealed strong gains at 4th grade, but only a slight gain at 8th grade. Writing data for 4th grade, comparing data from 2002 to 2004, and eighth grade, comparing data from 2003 to 2004, revealed strong gains at both grades. Gains in math were also strong for both 4th and 8th grades from 2002 to 2004. This may reflect the history of many school districts involvement with six trait writing preceding the formal statewide assessment process, and/or growing sophistication of staff over the years with the Nebraska STARS process.

Implications for Nebraska STARS

While these are first-year findings for writing and should be interpreted with caution, they do provide an initial base of support for the STARS process and, along with the gains in reading and math, strong credibility for the process and its continuation.

The very nature of STARS and the assessments used presents challenges to the evaluation process. There is a need to work with the educational measurement community to examine new ways to determine the success of the model. While criterion referenced approaches such as STARS are strongly supported by the measurement community from a standpoint of student learning, little has been done to support the acceptance of what is referred to as unconventional data as evidence of school success. While this writing assessment has technical characteristics that enable inferential statistics to be used in analysis, there may be some question from the traditional measurement community concerning this practice. Reading and math data are more clearly criterion referenced (specific to each district, not scored on a common metric) and descriptive analysis only was used in those comparisons. It must be remembered that the philosophy and purpose of Nebraska STARS is to support teaching and learning, not to focus on development of assessments for technical strengths in ranking results.

Much more work is needed in continuing to examine longitudinal results and the impact of refinements and corrections as Nebraskans problem solve concerns and the program evolves. One of the biggest challenges will be the ongoing effort to work with the United States Department of Education to ensure the STARS program will allow Nebraska to meet No Child Left Behind (NCLB) requirements.

This study supports the continuation of the Statewide Writing Assessment as part of the Nebraska STARS system. As noted by Roschewski, Gallagher, and Isernhagen (2001), the STARS plan brings together the best of both worlds; student learning is foremost, but public accountability is provided as well. While avoiding the “unintended negative outcomes” of high-stakes assessments, STARS places the responsibility of teaching and learning where it belongs. The teachers and administrators who create high quality learning environments are directly involved in affirming the quality and in measuring the learning of students in their classrooms.

References

- Editor (1999, March 1). Quality Counts-State of the States Report. *Education Week*. Retrieved on May 26, 2004, from <http://www.edweek.org/sreports/qc99>
- Gallagher, C. W. (2004, January). Turning accountability tables: Ten progressive lessons from one “backward” state. *Phi Delta Kappan*. 251-261.
- Nebraska Education Directory (2003-2004). *Nebraska Department of Education Directory*. Lincoln, NE. 106th Edition, 170.
- Plake, B. S. & Impara, J.C. (2000). Technical quality rubric for Nebraska’s district performance ratings. Lincoln, NE: Buros Center for Testing.
- Roschewski, P., Gallagher, C. & Isernhagen, J. (2001, April). Nebraskans reach for the STARS. *Phi Delta Kappan*. 611-615.

APPENDIX E
STATEWIDE WRITING ASSESSMENT SCORING

**Statewide Writing Assessment Scoring:
Effective Professional Development for the Classroom Teacher
Sue Anderson, NDE Coordinator of Statewide Writing Assessment**

Introduction

Statement of the Project

Statewide Writing Assessment Scoring: Effective Professional Development for the Classroom Teacher was a project of study undertaken to fulfill the requirements for the course, Nebraska Leadership for Learning 2003, at the University of Nebraska-Lincoln. This project was conducted to ascertain the effectiveness of the scoring training procedures for the Nebraska Statewide Writing Assessment through the perceptions of the scoring participants, referred to in this analysis as raters.

The benefits of this project of study were two-fold. First, the insights provided by raters related to the training procedures were beneficial to the Nebraska Department of Education in its implementation processes of the Statewide Writing Assessment. Secondly, the relationship that raters perceived between the scoring experience and their classroom practices may be very informative to the school improvement efforts of schools committed to improving student achievement in writing.

Purpose of the Project

The purposes of the project of study were to learn raters' perceptions of:

1. The effectiveness of the training they received during the scoring of the 2004 Nebraska Statewide Writing Assessment and
2. The effect of the scoring experience on their classroom practices.

Research Questions

The purposes of the project of study provided the basis for the research questions to be answered. They were:

1. What are raters' perceptions of the effectiveness of the training they received during the scoring of the 2004 Nebraska Statewide Writing Assessment?
2. What impact does participation in the scoring of the Statewide Writing Assessment have on raters' classroom practices in the teaching of writing?

Review of Related Literature

The national dialogue about accountability for student learning frequently moves to a local debate over the effects of state level testing on student learning. The merits or lack thereof of state level writing tests have not been omitted from the debate. Of performance assessments in state testing programs, Linda Mabry (1999) reported that direct writing assessments were the most prevalent and longest running. While that may be true, much has been written about how they are not accurate measures of how well students can write. The primary reason cited was standardization measures that some believed reduce student writing to formulaic texts that demonstrated little creativity or original thought. Standardization measures of large scale writing assessments that come under 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 (1999) reported that all states administering direct writing assessments used rubrics to score student performance. In her discussion of rubrics she maintained that rubrics

“promote reliability by standardizing scoring, but they also standardize writing” by limiting the writing to “pre-determined standards.” She further argued that rubrics “overwhelm the curriculum,” and the result was that teacher-driven writing curricula, which may be superior, were overridden, an example of what she called “the de-professionalism of teachers effectively being prevented from making decisions about curriculum and pedagogy for their students” (p. 35).

Opponents of state writing assessments also argued 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 George Hillocks’ (2003) conclusions in *The Testing Trap: How State Writing Tests Control Learning* was that when states determined categories of writing to assess, those categories received 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 echoed 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. Grant Wiggins (1998) warned 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) maintained 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 (2003) 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 reported that teachers in each of the five states believed 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 believed their state writing test was not an adequate measure of student ability. Only teachers from Kentucky believed, however, that the state test, which was a portfolio of writing that students create over time, was a valid measure of their writing skills.

Hillocks’ (2003) 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 utilized their own teachers to score the state writing test, reported 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 enlisted classroom teachers as raters. They maintained that teachers who assessed students’ work in the large scale setting gave themselves an education both in 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, were higher expectations about 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.

Nebraska's Statewide Writing Assessment was implemented in 2001 as a result of state legislation (LB812) mandating school accountability in assessing and reporting student performance in meeting state content standards in reading, mathematics, and writing. Policy guidelines for the writing assessment issued by the State Board of Education required that the assessment be a direct assessment of student writing requiring students in grades 4, 8, and 11 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 sample of student writing to be read and scored by two raters who assign a single holistic score within allowable ranges as prescribed by the rubric. The rubric criteria are identified as ideas and content, organization, voice or tone, word choice, sentence fluency, and conventions. Raters assign a score based on how the writing meets these criteria overall. The rubric criteria are aligned to the qualities of writing identified in the Nebraska content standards for writing.

Nebraska teachers are recruited by the Nebraska Department of Education to score the writing assessment each year. The qualifications to become a scorer are that the teacher is currently teaching or has taught at or near the grade level being assessed, be familiar with student writing at the grade level being assessed, and have some basic knowledge of the six trait writing assessment model.

Scoring of the state assessment is held at a central location in the state and scorers come to the site for three days during which the scoring and all training for the scoring occurs. All training is coordinated and facilitated by State Department of Education personnel.

In "Charting STARS – Sustainability as Challenge and Opportunity," Chris Gallagher (2003) reported 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 (2003) reported that 88% of teachers agreed or agreed strongly 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 (2003).

Methodology

Population and Sample

The population of this study consisted of 392 raters in the scoring of the 2004 Nebraska Statewide Writing Assessment, with 130 raters at Grade 4, 130 raters at Grade 8, and 132 raters at Grade 11. A scoring session was conducted for each grade level.

Among raters in the 2004 scoring sessions 110 Nebraska school districts were represented. The 110 school districts represented all geographic regions in the state. Only raters were asked to complete a survey at the close of each scoring session in order to ascertain their perceptions about the scoring experience. Raters were not asked to provide their names or address specific information.

The survey, entitled 2004 Statewide Writing Assessment Scoring Rater Evaluation, contained 15 questions across the following categories: general rater information; perceptions of various components of the scoring training and schedule; impact of the scoring experience on classroom practices; general comments about the scoring experience (see Appendix A). The questions that asked for general rater information pertained to previous scoring experience at the state level, previous scoring experience at the regional or local level, participation in six trait writing training, years of classroom teaching experience, and current teaching status including grade level(s) and subject area(s). The questions related to scoring session information asked raters to rate the effectiveness of various components of the scoring training, scoring schedule, and scoring site on a scale of poor to excellent. A single question asked raters to indicate the way(s) they thought the scoring experience would affect their classroom teaching and the learning of their students. For this question raters could select one, two, three or four responses. A final item on the survey invited raters to provide additional comments on any aspect of the scoring experience that they wished to share.

Data Collection and Analysis

The data for this study were collected from rater responses to a survey created by the Coordinator of Statewide Writing Assessment at the Nebraska Department of Education. The questions were designed to gather information about the writing assessment scoring experiences of raters and their perceptions of the effectiveness of the various training components provided during the scoring sessions.

The survey was designed to be completed by raters in no more than 10 minutes. Raters 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 responses were coded and entered onto a spreadsheet using Microsoft Excel 2001 at the Nebraska Department of Education. The data was then tabulated at the Nebraska Evaluation and Research Center at the University of Nebraska-Lincoln. Analysis included the tabulation of frequency of responses for all items and cross-tabulation of responses to selected items for the purpose of determining the relationship between previous state level scoring experience, years of teaching experience, current teaching assignment, previous

training in the six trait writing model, and the impact of the scoring experience for scorers on their classroom practices.

Findings

Analyses of the survey results provided information related to 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 392 raters participating in the scoring across all grade levels, 314 raters responded to the survey, for a return rate of 80%.

General Rater Information

Among raters who responded to the survey, 64% were currently classroom teachers in Nebraska schools; 35% were retired educators (5%), substitute teachers (10%), or others (17%) with education-related responsibilities or experiences (i.e. school administrators, curriculum specialists, guidance counselors).

Raters' teaching experience by grade level ranged from elementary (grades 1-4), 38%, to middle (grades 5-8), 16%, to high school (grades 9-12), 43%. Two percent of raters reported teaching at both middle and high school levels or in special programs such as special education or federal Title programs. Raters currently in the classroom reported the subject or content areas they teach as language arts only, 4%; a general curriculum including language arts and one or more subject or content areas, 38%; and subjects or contents other than language arts, 56%.

Among raters there was a range of years of teaching experience with 19 % reporting 1-5 years of experience, 13% with 6-10 years of experience, 19% with 11-19 years of experience, and 47% with 20 or more year of experience. Raters' years of teaching experience reported by grade level are included in Table 1.

Table 1: Raters' Years of Teaching Experience by Grade Level Group

	1-5 Years	6-10 Years	11-19 Years	20 Years or More
Grade 4	23%	10%	21%	45%
Grade 8	14%	10%	23%	54%
Grade 11	18%	17%	16%	48%
Average of All Grade Levels	19%	13%	19%	48%

Across all grade levels a majority of raters, 66%, had not participated in previous Statewide Writing Assessment scoring sessions. Of the of raters who had previously participated in state level scoring, 23% participated in 2001, the pilot year of the assessment, 13% in 2002, and 17% in 2003.

A majority of raters across all grade levels, 69%, reported previous experience within their school districts or at regional Educational Service Units in large scale writing assessment scoring. Similarly, a majority, 89%, reported having participated in some form of training in the six trait writing assessment model prior to the 2004 state scoring.

Raters' Perceptions of the Effectiveness of Scoring Training Procedures

Training for raters in the 2004 Statewide Writing Assessment scoring was included in the schedule for each grade level scoring session. Raters were asked to evaluate the

effectiveness of the various training processes and procedures as well as the appropriateness of the scoring schedule and scoring site.

The primary training of raters occurred on the first day of each scoring session. The content of this 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 primary training format was consistent across grade levels. The content of the training varied from grade level to grade only with relationship to the grade specific writing prompts and mode specific anchor papers. Across all grade level scoring sessions 65% of raters rated the effectiveness of the primary training as excellent, 31% as very good, 4% as good, 1% as fair. No raters rated the primary training as poor. Raters' perceptions of the effectiveness of the primary scoring training by grade level are included in Table 2. See also Appendix B.

Table 2: Raters' Perceptions of the Effectiveness of Primary Scoring Training

	Poor	Fair	Good	Very Good	Excellent
Grade 4	0%	1%	2%	36%	61%
Grade 8	0%	2%	2%	32%	64%
Grade 11	0%	2%	5%	24%	69%
Average of All Grade Levels	0%	1%	4%	31%	65%

In addition to the primary training raters received, scoring room leaders conducted re-training activities throughout the remainder of the scoring session. The content of these re-training procedures included reviewing the rubrics and anchor papers and scoring of additional blind scored papers for purposes of rater re-calibration. Across all grade levels 65% of raters rated the effectiveness of these re-training procedures as excellent, 32% as very good, 7% as good, and 1% as fair. None of the raters in any grade level rated the re-training procedures as poor. Among grade level groups the Grade 8 raters rated the re-training procedures highest with 72% rating them as excellent. For ratings of retraining procedures by grade level groups see Table 3. See also Appendix B.

Table 3: Raters' Perceptions of the Effectiveness of Rater Re-Training Procedures

	Poor	Fair	Good	Very Good	Excellent
Grade 4	0%	0%	8%	37%	55%
Grade 8	0%	2%	4%	23%	72%
Grade 11	0%	2%	7%	30%	60%
Average of All Grade Levels	0%	1%	%	32%	65%

A variety of scoring materials were provided by the Nebraska Department of Education (NDE) for use by raters throughout the training and scoring. These materials included the "2004 Nebraska Statewide Writing Assessment Scoring Manual," which contains explanations of all scoring procedures, rubrics, anchor papers, information about the characteristics of writing modes, score point scale worksheets, scoring keys, pencils, and

highlighter markers. Use of these materials by raters was coordinated during the training and scoring by scoring room leaders and through a video presentation by the NDE Coordinator of the Statewide Writing Assessment. Sixty-four percent of raters across all grade levels rated the effectiveness of the scoring materials as excellent, 28% as very good, 6% as good, and 2% as fair. Less than 1% of raters rated the scoring materials as poor. For rater perceptions of the effectiveness of the scoring/training materials by grade level group see Table 4. See also Appendix B.

Table 4: Raters' Perceptions of the Effectiveness of Training/Scoring Materials

	Poor	Fair	Good	Very Good	Excellent
Grade 4	0%	0%	4%	35%	61%
Grade 8	0%	4%	7%	21%	68%
Grade 11	.8%	2%	8%	25%	64%
Average of All Grade Levels	.3%	2%	6%	28%	64%

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 2004 there were five scoring leaders, one for each of the scoring rooms and 13 table leaders, one for every 2-3 table groups of raters.

Scoring room leaders' responsibilities were to facilitate the training and 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.

In each scoring room, table leaders were assigned to assist with re-training procedures, 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 scored, and their understanding of the characteristics of mode specific writing.

Seventy-four percent of raters across all grade levels rated the effectiveness of the scoring room leaders as excellent, 21% as very good, 5% as good, and 1% as fair. Similarly 74% rated the effectiveness of the table leaders as excellent, 21% as very good, 4% as good, and 1% as fair. No raters rated scoring leaders or table leaders as poor. For raters' perceptions of the effectiveness of the scoring room leaders and table leaders by grade level group see Table 6. See also Appendix B.

The scoring sessions were held at Educational Service Unit #3 in Omaha, Nebraska in March and April of 2004 (Grade 4, March 18-20; Grade 8, March 25-27; Grade 11, April 1-3). Each scoring session consisted of three days with a daily schedule of 8:30 a.m. - 4:00 p.m. A 45-minute lunch break and 15-minute mid-morning and mid-afternoon breaks were included in the schedule. All scoring sessions were held at Educational Service Unit #3 in Omaha, Nebraska. In rating the appropriateness of the

scoring facility, raters were asked to consider such items as scoring room arrangements, room temperature and lighting, meals, restrooms, overall neatness and cleanliness, and site support staff. Raters were also asked to rate the appropriateness of the scoring session schedule and the scoring facility. Sixty-two percent of raters rated the scoring schedule as excellent, 29% as very good, 7% as good, and 1% as fair. The scoring facility ratings by raters were as follows: 77% excellent; 20% very good; 2% good; and 1% fair.

Table 6: Raters' Perception of the Effectiveness of Scoring Room Leaders and Table Leaders

Scoring Leaders	Poor	Fair	Good	Very Good	Excellent
Grade 4	0%	2%	2%	23%	73%
Grade 8	0%	0%	4%	21%	76%
Grade 11	0%	2%	6%	18%	75%
Average of All Grade Levels	0%	1%	5%	21%	74%
Table Leaders					
Grade 4	0%	1%	3%	24%	73%
Grade 8	0%	2%	0%	23%	77%
Grade 11	0%	3%	8%	22%	68%
Average of All Grade Levels	0%	1%	4%	21%	74%

Effect of the Scoring Experience on Raters' Classroom Practices

Raters were asked in what way(s) they thought their state level scoring experience would affect their teaching and the learning of their students. Response options for this question were: the scoring experience broadened my knowledge or expectations of students writing; the scoring experience informed my understanding about teaching writing; the scoring experience affirmed my current teaching practices; the scoring experience provided other benefits (i.e. effective professional development or support for local school improvement efforts). Raters could select any one or all of these response options.

Across all grade levels 81% of raters reported the scoring experience broadened their knowledge or expectations of student writing; 58% said the scoring experience informed their understanding of the teaching of writing; 62% responded that the scoring experience affirmed their current teaching practices; and 39% indicated the scoring experience provided other professional benefits. Among raters across all grade levels, Grade 4 raters associated the greatest impact of the scoring experience on their classroom practices with 87% reporting the experience broadened their knowledge and expectations of student writing, 60% said the experience informed their knowledge of teaching writing; 70% indicated the experience affirmed their current classroom practices; and 43% said the experience provided other professional benefits. For responses to this question by each grade level group see Table 7. See also Appendix C.

When responses to this question were compared with responses of raters based on their years of teaching experience, a greater number of raters with more years of teaching experience reported benefits from the scoring experience than raters with fewer years of teaching experience. Across all grade level groups 44% of raters with 20 or more years of teaching experience reported benefits associated with the scoring experience; 23% of raters

with 11-19 years of teaching experience reported benefits associated with the scoring experience; 20% of raters with 1-5 years of teaching experience reported benefits associated with the scoring experience; and 14% of raters with 6-10 years reported benefits of the scoring experience.

Table 7: Raters' Perceptions of the Effect of the Scoring Experience on Their Classroom Practices

	Broadened Knowledge/Expectations of Student Writing	Informed Knowledge of Student Writing	Affirmed Current Classroom Practices	Other Professional Benefits	Average of all Responses
Grade 4	87%	60%	70%	43%	65%
Grade 8	77%	55%	64%	40%	59%
Grade 11	78%	55%	54%	33%	55%
Average of All Grade Levels	81%	58%	62%	39%	60%

When the responses to this question were cross tabulated with the responses of raters who were currently teaching in classrooms the results showed that across all grade level groups raters found the scoring experience very beneficial. Among raters, who are currently teaching, a greater percentage, 73%, reported benefits associated with the scoring experience; 68% of Grade 8 raters reported benefits; and 65% of Grade 11 raters. For percentage of rater responses to all response options by grade level groups see Table 8. See Appendix D.

Table 8: Raters' (Currently Teaching) Perceptions of the Effect of Scoring Experience on Classroom Practices

	Broadened Knowledge/Expectations of Student Writing	Informed Knowledge of Student Writing	Affirmed Current Classroom Practices	Other Professional Benefits	Average of All Responses
Grade 4	69%	71%	83%	67%	73%
Grade 8	65%	67%	81%	60%	68%
Grade 11	62%	58%	81%	57%	65%
Average of All Grade Levels	65%	65%	82%	61%	68%

The final item on the survey asked raters to write comments related to any aspect of the scoring experience they wished to share. The results of this item were not tabulated for purposes of this study, but all comments were transcribed and organized according to grade level.

Summary of Findings

This project of study was designed to ascertain the effectiveness of the training procedures for raters participating in the 2004 Nebraska Statewide Writing Assessment and to learn their perceptions of the relationship between the scoring experience and their classroom practices as teachers.

A majority of raters, 64%, in the 2004 state assessment scoring were currently classroom teachers with 11 or more years of teaching experience (67%). For most raters, 66%, the 2004 scoring was their first state level scoring experience. However, a majority, 70%, reported having previous experience participating in large scale writing assessment

scoring at the local school district or regional educational service unit level. Eighty-nine percent of raters also reported having participated in six trait writing training prior to the 2004 scoring.

Hillocks' (2003) study of five states' writing test provided important information about the relationship between state level writing tests and their effects on local curriculum decisions. Even though he concluded that the expectations of state writing tests tend to override local decisions of pedagogy which may be based on the 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 did 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.

Across all grade levels of the 2004 Nebraska Statewide Writing assessment scoring, raters' perceptions of the effectiveness of the scoring training was very positive with 96% rating the primary training procedures as very good or excellent, 97% rating the re-training procedures as very good or excellent, 92% rating the training and scoring materials as very good or excellent, and 97% rating the effectiveness of the scoring room leaders and the table leaders as very good or excellent. In addition, 91% of raters across all grade levels rated the scoring schedule as very good or excellent, and 97% rated the scoring site facility as very good or excellent.

Raters' perceptions about the effect of the scoring experience on their current classroom practices indicated that 65% felt the scoring experience broadened their knowledge and expectations of student writing. Similarly across all grade level groups, 65% of raters currently in the classroom felt the scoring experience informed their knowledge of the teaching of writing, and 82% reported that the experience affirmed their current classroom practices and associated other professional benefits with the scoring experience. Among all raters, including those not currently in the classroom, 62% associated positive professional benefits with the scoring experience. Raters with 11 or more years of teaching experience associated benefits with the scoring experience more than raters with 10 or fewer years of teaching experience.

Conclusions

Based on the results of this project of study it is possible to draw important conclusions about the raters and their perceptions of the scoring process. First, a majority of raters, while not previously experienced in state level scoring, possessed other types of experience valuable to the scoring process, including years of classroom teaching experience, participation in local or regional level large scale writing assessment scoring, and training in the six trait writing assessment model prior to the scoring. Secondly, raters' perceptions about the effectiveness of the scoring training processes 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. Thirdly, raters' perceptions of the benefits of the scoring to their classroom practices were positive in that the majority of raters across all grade levels, particularly those currently teaching, felt the scoring

experience broadened their knowledge or expectations of student writing, informed their knowledge of teaching writing, affirmed their current classroom practices, and or provided other professional benefits such as aid to local school improvement efforts.

Recommendations

The results of this project of study provided 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 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 so doing, their students will reap the benefits of the teachers' learning when they return to their classrooms.

References

- Gallagher, C. (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*. New York: Teachers College Press.
- Freedman, S.W., and Diaute, C. (2001). *Instructional methods and learning in teaching writing*. In *subject-specific instructional activities and methods*. 8. New York: Elsevier Science.
- Mabry, L. (1999, May). Writing to the rubric. *Phi Delta Kappan*. pp. 673-679.
- 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.

Appendix A

**Nebraska Department of Education
2004 Statewide Writing Assessment Scoring Rater Evaluation**

Thank you for taking a few moments to complete this questionnaire. Your responses will assist the overall implementation of the Statewide Writing Assessment scoring process.

Rater Information

Have you participated as a rater in the Statewide Writing Assessment before 2004?

yes no

If yes, when? 2001 (Pilot Year Grades 4, 8, 11) 2002 (Grade 4) 2003 (Grade 8)

Have you participated as a rater for your district/building/ESU writing assessment? yes no

Have you ever attended a workshop or training in Six Trait Writing? yes no

Years of classroom teaching experience: 1-5 6-10 11-19 20 or more

Are you currently a classroom teacher? yes no

If yes, what grade level(s) do you currently teach _____

If yes, what subject area(s) do you currently teach _____

If no, what is your status? retired teacher substitute teacher other

Scoring Session Information

1. How would you rate the effectiveness of the training for scoring (overview of the scoring process, understanding the scoring criteria, using the rubric to assign scores, using anchor papers, practice scoring rounds, etc.)?

Poor Fair Good Very Good Excellent

2. How would you rate the effectiveness of the refresher training procedures (use of re-calibration/practice papers, re-training on the rubric) for raters throughout the scoring session?

Poor Fair Good Very Good Excellent

3. How would rate the effectiveness of the training materials (rubrics, anchor papers, rater scoring manual, training video) used during the training and scoring sessions?

Poor Fair Good Very Good Excellent

4. How would you rate the effectiveness of your scoring room leader in facilitating the rater training?

Poor Fair Good Very Good Excellent

5. How would you rate the effectiveness of your table leaders in assisting with re-training and 3rd reads?

Poor Fair Good Very Good Excellent

6. How would you rate the appropriateness of the daily scoring schedule (length of scoring day, length and frequency of breaks, time allotted for lunch)?

Poor Fair Good Very Good Excellent

7. How would rate the scoring site facility (scoring rooms, room temperature, lighting, hospitality, meals, restrooms, overall neatness and cleanliness, site support staff)?

Poor Fair Good Very Good Excellent

8. In what way(s) do you think your experience as a rater for the 2004 Statewide Writing Assessment will affect your teaching and the learning of your students?

Broadened my knowledge/expectations of student writing

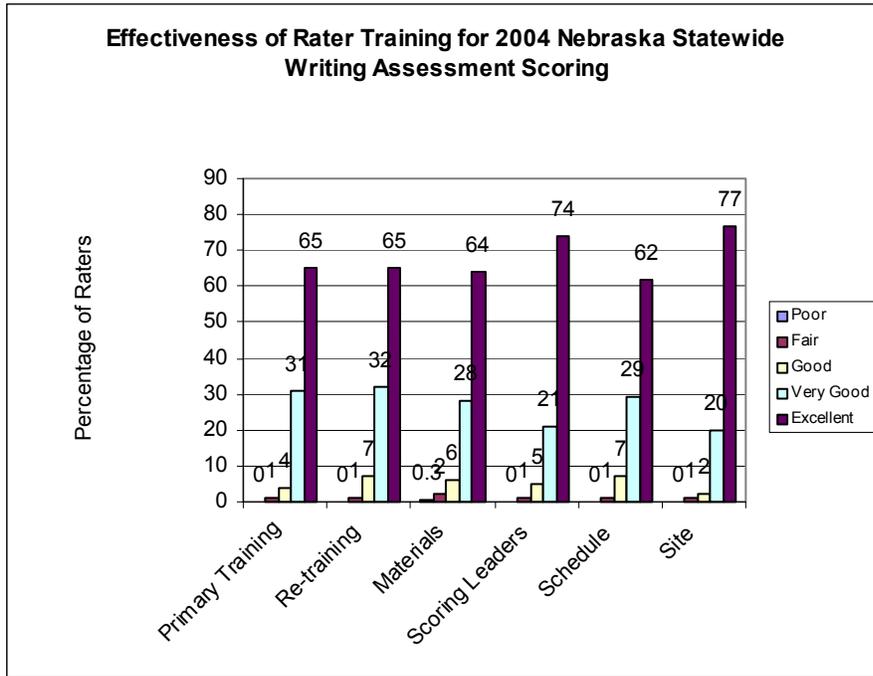
Informed my understanding about teaching writing

Affirmed what I am doing in my classroom in the teaching of writing

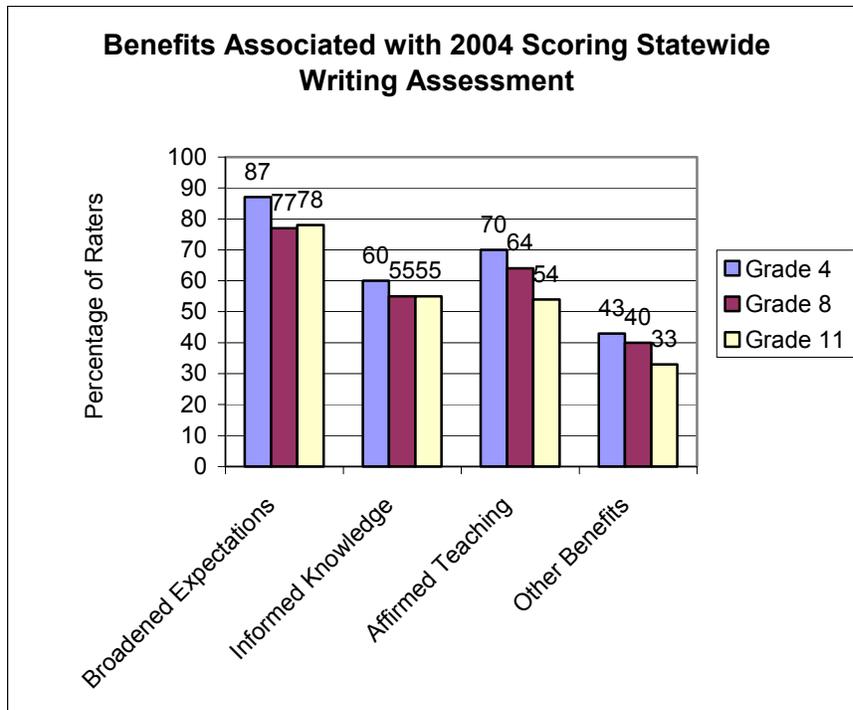
Other (i.e. effective professional development, aid to local school improvement)

9. Additional comments/thoughts you would like to share about your experience as a rater or about the scoring process.

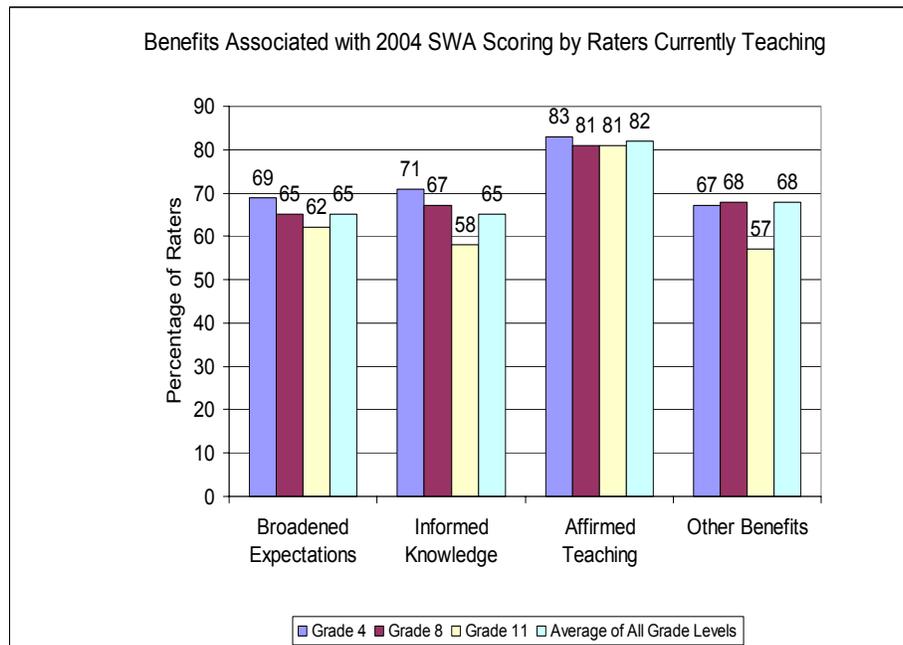
Appendix B



Appendix C



Appendix D



APPENDIX F
EDUCATORS' PERCEPTIONS OF STARS

Educators' Perceptions of STARS

**Toby Boss, Staff Developer, ESU#6; Dan Endorf, Principal, York High School;
Tammy Heflebower, Staff Development Director, ESU#6; and Phil Warrick, Principal,
Waverly High School**

Introduction

Although most states use a statewide strategy for student assessment, Nebraska school districts are responsible for determining the strategies for measuring and reporting students' performance in reading and mathematics, with science (2006) and social studies (2007) added in the next two years. The School-based Teacher-led Assessment Reporting System (STARS) empowers educators within each school district, with assistance from respective educational service units, to develop, implement, and manage data from the assessments. Within each Nebraska school district are many STARS stakeholders with various job titles and certain responsibilities connected to the assessment model (Boss, Endorf & Buckendahl, in press; Roschewski, 2004).

Seven hundred eighteen (718) educators from across the state of Nebraska participated in this research. This web-based study of public school educators included four groups heavily involved with the assessment model, with the group referred to as "educators" hereafter: assessment coordinators, Educational Service Unit staff developers, principals and teachers.

A 59-question survey asked educators to examine three aspects of the STARS initiative. The first area of concentration was the overall impact of STARS on education in Nebraska. The impact on curricular, instructional, and assessment practices used by educators to implement STARS was the second area of concentration. The final portion of the survey analyzed the perceptions of educators about the impact of STARS on the professional abilities of other educators across the state of Nebraska.

Purpose of Study

The purpose of this study was to determine the perceptions and practices of educators across the state of Nebraska as they implemented STARS. The research team identified the following three research questions to guide the study:

1. What are the perceptions of educators about STARS as it relates to education in Nebraska?
2. What are the curricular, instructional, and assessment practices used by educators to implement STARS?
3. What are the perceptions of educators about the impact of STARS on the professional abilities of educators across the state of Nebraska?

Research Design and Methodology

Designed as a descriptive, quantitative study, this research specifically analyzed the perceptions of educators involved with the STARS process. Collection of data occurred from January 17, 2005 to February 10, 2005 through a self-designed, web-based survey. Web-based surveys have the potential of bringing efficiencies to self-administered questionnaires not possible with paper-pencil surveys, all the while reducing implementation time (Dillman, 2000).

The survey design allowed for a numeric description of the sample by asking educators questions, which then empowered the researchers to generalize to the larger population (Fowler, 1988). This study used a cross-sectional survey procedure to gather data from the sample population. The cross-sectional methods sought to gather data from a particular group at a single point in time (Ay, Jacobs & Razavieh, 2002; McMillan, 2000).

Two sampling techniques determined the eligibility of educators to participate in the study. First, the researchers surveyed the entire population of assessment coordinators and ESU staff developers in the state with accessible e-mail addresses. In all, 386 assessment coordinators and 54 ESU staff developers in the state received the opportunity to participate in the study.

A stratified purposeful sampling technique was the strategy employed to select the principals and teachers eligible for the study. The incorporation of two strata, first from a geographical perspective and then according to NDE school classification, ensured a representative sample of principals and teachers from across the state. The respondents were first selected from five geographic regions in the state based upon a cluster of service units. For the second sampling strata, each school district in the five geographic regions was categorized according to NDE school classification. A proportionate number of educators were selected from each geographic region to participate based upon the accessibility of e-mail addresses. At least 60 principals from each region and 150 teachers from each region received an opportunity to participate in the study, with oversampling procedures used with those educators working in school classifications with smaller student populations. Due to the nature of an online survey, the sample included only those educators with accessible e-mail addresses and/or on-line capability.

One hundred ninety-six (196) assessment coordinators completed surveys (55% return rate), 43 ESU staff developers participated (80% return rate), 156 public school principals responded (45% return rate) and 323 teachers (36% return rate). These return rates are similar to other on-line surveys (Dillman, 2002).

Survey data were collected from the web host and imported into Excel spreadsheets and a Statistical Package for Social Sciences (SPSS). The research team used the services of the Nebraska Evaluation and Research Center for assistance with data analysis.

Results

Research Question 1: What are the perceptions of educators about STARS as it relates to the improvement of education?

This section of the survey included questions regarding the perceptions of educators about the quality of education due to STARS. The survey section included 10 questions on a 7-point Likert Scale: “Much Worse” (1); “Worse” (2); “Slightly Worse” (3); “About the Same” (4); “Slightly Better” (5); “Better” (6); and “Much Better” (7). The total score for the section could range between 10 and 70, with a higher score indicating greater agreement that STARS improved education. Table 1 shows the mean, mode, standard deviation, and variance statistics for the first section of the survey:

Table 1: Quality of Education: Mean, Mode, Standard Deviation, & Variance

Educator Group	Completed Surveys	Mean	Mode	Standard Deviation	Variance
Assessment Coordinators	196	50.99	58	10.43	108.79
ESU Staff Developers	38	58.68	60	5.97	35.68
Principals	154	50.71	59	10.61	112.67
Teachers	330	44.71	39	9.42	88.80

Based upon the mean score, the average answer in Research Question 1 for ESU staff developers were 5.9, which indicated that scores hovered around the “Much Better” range. The average score for teachers (4.5) fell between “About the Same” and “Slightly Better,” while both assessment coordinators (5.1) and principals (5.1) fell between the other two groups with average responses between “Better” and “Much Better.”

Research Question 2: What are the curricular, instructional, and assessment practices used to implement STARS?

Two sub-sections of the survey gathered data regarding the curriculum, instruction, and assessment practices used to implement STARS. The first sub-section of the survey asked about the importance of sound curriculum and assessment practices due to STARS. Survey questions stemmed from a meta-analysis of curricular practices in *What Works in Schools* (Marzano, 2003), as well as sound assessment practices determined by the Six Quality Assessment Criteria (NDE, 2001). The first survey sub-section included 21 questions on a six-point Likert Scale: “Strongly Disagree” (1); “Disagree” (2); “Slightly Disagree” (3); “Slightly Agree” (4); “Agree” (5); and “Strongly Agree” (6). The total score for the section could range between 21 and 126, with a higher score indicating greater agreement that STARS implementation led to more effective curricular and assessment practices in Nebraska schools. Table 2 shows the mean, mode, standard deviation, and variance statistics for the first sub-section.

Table 2: Curriculum and Assessment Practices: Mean, Mode, Standard Deviation, & Variance

Educator Group	Completed Surveys	Mean	Mode	Standard Deviation	Variance
Assessment Coordinators	178	91.79	105	6.38	268.52
ESU Staff Developers	36	98.83	104	9.33	87.05
Principals	144	92.14	105	17.74	314.88
Teachers	307	82.31	82	16.80	282.40

ESU staff developers gave the highest ratings on sub-section 1 of Research Question 2 a mean score of 98.83 or an average rating of 4.7 per question, which indicated that the typical answer fell closer to “Agree” than “Slightly Agree.” Assessment coordinators and principals both gave scores of 4.4 on average, which means their average answers were between “Agree” and “Slightly Agree.” Teachers, however, averaged a mean of 3.9, which represented an average answer of almost “Slightly Agree.”

The second sub-section of Research Question Two included 12 questions about the development and implementation of instructional practices due to STARS. Survey questions emerged from *Classroom Instruction That Works* (Marzano, 2001), a meta-analysis of classroom instructional practices and a natural connection to the aforementioned curriculum and assessment framework. Respondents ranked the frequency of each instructional strategy

on a three-point Likert Scale: “Less Often” (1); “About the Same” (2); and “More Often” (3). Total score for the section could range between 12 and 36, with a higher score indicating that implementation of effective instructional strategies occurred more often as a result of STARS. Table 3 shows the mean, mode, standard deviation, and variance statistics for the second sub-section:

Table 3: Instructional Practices: Mean, Mode, Standard Deviation, & Variance

Educator Group	Completed Surveys	Mean	Mode	Standard Deviation	Variance
Assessment Coordinators	175	25.74	24	5.25	27.60
ESU Staff Developers	36	28.50	26	6.43	41.40
Principals	145	27.36	24	4.032	16.026
Teachers	313	25.47	24	3.63	13.22

Three of the four groups appeared to think STARS had little effect on instructional practices. All groups averaged scores of “About the Same” except ESU staff developers, which fell between “About the Same” and “More Often.”

Research Question 3: What are the perceptions of educators about the abilities of other educators as a result of STARS?

Section Three of the survey was aligned to Research Question 3, regarding the perceptions of educators about the professional abilities of other educators as a result of STARS. This section of the survey asked questions concerning the degree to which STARS changed the knowledge of educators about curriculum, instruction, assessment and educational leadership among the four groups (assessment coordinators, ESU staff developers, principals and teachers). This design strategy allowed each of the four survey groups to rate the abilities of their group as well as the other three groups. The survey section included 16 questions on a 7-point Likert Scale: “Much Worse” (1); “Worse” (2); “Slightly Worse” (3); “About the Same” (4); “Slightly Better” (5); “Better” (6); and “Much Better” (7). The total score for the section could range between 16 and 112, with a higher score indicating that the knowledge and skills of educators improved as a result of STARS. Table 4 shows the mean, mode, standard deviation, and variance.

Table 4: Professional Abilities: Mean, Mode, Standard Deviation, & Variance

Educator Group	Completed Surveys	Mean	Mode	Standard Deviation	Variance
Assessment Coordinators	164	88.31	96	12.76	162.84
ESU Staff Developers	35	96.60	96	8.45	71.54
Principals	132	89.62	100	13.50	182.35
Teachers	282	79.28	64	14.57	212.56

Using the mean scores of the third section as a guide, ESU staff developers (6.0) averaged a response rate of “Better,” whereas both assessment coordinators (5.5) and principals (5.6) gave average scores between “Better” and “Slightly Better.” Teachers gave an average score of 4.9, which represents an average answer of almost “Slightly Better.”

Discussion and Implications

Several general conclusions can be made from this data. The major finding of this study is that educators were generally positive in their perceptions of STARS. ESU staff

developers gave the STARS model the most positive responses of the four educator groups surveyed; conversely, teachers gave STARS consistently lower marks relative to the other groups. Scores from assessment coordinators and principals were generally similar and placed between the scores of ESU staff developers and teachers.

All groups believed that public education in Nebraska improved due to STARS. Aligned with the generally positive impression of STARS mentioned above, educators reported that the procedures involved with STARS are generally understood and that the practices to implement these procedures occur in a successful manner.

The second research question asked educators about the curricular, instructional and assessment practices due to STARS. Sub-section one focused on curriculum and assessment. Six specific questions from this sub-section concentrated on assessments for STARS being aligned to the Six Quality Criteria (one question for each requirement). Data from all of the groups strongly suggests that STARS assessments are generally aligned to the Six Quality Criteria.

The second sub-section of Research Question Two revolved around classroom instruction. Assessment coordinators and ESU staff developers responded to the frequency in which training on effective strategies took place, while principals and teachers responded to the frequency with which implementation of the instructional strategies occurred. Tables 5 and 6 below show how the groups responded:

Table 5: Frequency of Training

Educator Group	Less Often	About the Same	More Often
Assessment Coordinators	12.59	59.97	19.53
ESU Staff Developers	13.1	34.93	51.97

Table 6: Frequency of Implementation

Educator Group	Less Often	About the Same	More Often
Principals	3.98	64.15	31.88
Teachers	7.15	73.32	19.53

The mean scores for assessment coordinators, principals, and teachers indicated there was little change with instructional practices; however, the response from ESU staff developers points toward ample opportunities for training in classroom instruction.

The third research question asked educators how they felt about the abilities of the four educator groups due to STARS. Each of the groups responded favorably regarding the curricular, instructional, and assessment knowledge level of educators and also provided positive impressions of the leadership each group exhibited since the inception of STARS.

This data is useful to NDE as they contemplate next steps with STARS. First, NDE should consider surveying other educational groups, such as superintendents or students, to compare their impression of STARS with the other four groups. It may also be of interest to drill deeper into one or all of the educator groups by analyzing specific demographic information such as years of experience, geographic location in the state, and NDE school classification.

A second recommendation is for NDE to investigate the discrepancy in data between instructional training and instructional practices. While ESU staff developers believed they

are providing the training, those working in the districts believed otherwise. NDE may want to consider additional research in this dimension of STARS and/or extensive training opportunities for teachers to learn and grow in the area of instruction.

It would also behoove education throughout the state to collect and maintain a large e-mail address database of educators from these respective groups. The researchers went to great lengths to collect e-mail addresses of assessment coordinators, principals, and teachers for our web-based survey.

Lastly, STARS deserves an opportunity to prove itself over the course of time. The data revealed a sense of ownership and understanding in the system from each of the four educator groups. NDE is worthy of commendation for its many years of labor with this different approach to statewide assessment.

References

- Ay, D., Jacobs, L. & Razavieh, A. (2002). *Introduction to research in education*. Belmont, CA: Wadsworth.
- Boss, T., Endorf, D. P., & Buckendahl, C.W. (in press). Informing state assessment from the local level: A district's reflections. *The Researcher*.
- Dillman, D. A. (2002). *Mail and Internet surveys, The tailored design method*. New York, Wiley & Sons, Inc.
- Marzano, R. J. (2001). *Classroom instruction that works*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McMillan, J.H. (2000). *Education research: Fundamentals for the consumer*. New York: Longman, Inc.
- Nebraska Department of Education. (2001). *School-based teacher-led assessment and reporting system: A summary*. Lincoln, NE: Author.
- Roschewski, P. (2004). History and background of Nebraska's school-based teacher-led assessment and reporting system (STARS). *Educational Measurement*, 23(2), 7-9.

APPENDIX G
NEBRASKA SCHOOL SUPERINTENDENTS'
PERCEPTIONS

**Nebraska Public School Superintendents' Perceptions Of Nebraska's
Assessment/Accountability System's Effect On High Schools
Wanda Clark**

Introduction

As the 49th state to adopt an assessment/accountability system, Nebraska is unique. Its 517 school districts range in size from one to more than 46,000 students. These districts cover a combined area of over 77,000 square miles. More than 300 are elementary-only districts, while approximately 400 have a population of less than 100 students. Sixty percent (60%) of the students in the state are enrolled in the 20 largest school districts (Nebraska Department of Education, 2002b).

Nebraska is progressive. As a result of Nebraska Legislative Bill 812, passed in the spring of 2000, the state developed and implemented an assessment/accountability system called STARS (School-based Teacher-led Assessment and Reporting System) that defined content standards in the areas of communication (reading, writing, listening, and speaking), mathematics, science, and social studies/history. These content standards are to be assessed at the end of four grade divisions: (1) Kindergarten- 1st grade; (2) 4th grade; (3) 8th grade; and (4) 11th grade (Christensen, 2001; Roschewski, Gallagher, & Isernhagen, 2001).

Purpose and Methodology

This study examined public school superintendents' perceptions of how the Nebraska system (STARS) has affected high schools. Data were gathered using an on-line survey developed from instruments originally created by Weichel (2002), Duke et al. (2000), and Johnson (1981) and modified for use with superintendents. E-mail requests were sent to public school superintendents in Nebraska whose districts have high schools inviting them to participate in the survey. The response rate was 50% (129/259). The survey questions were constructed on a five-point Likert scale where one represented "strongly disagree" and five represented "strongly agree."

The variables measured were (1) district size, (2) years of experience as a superintendent, (3) percentage of district students on free and reduced lunch, (4) rating received on the 2002-2003 district assessment portfolio for communication at the eleventh grade, (5) amount of funding provided from the Nebraska Department of Education (NDE) to implement STARS, and (6) amount of additional district financial support required to implement STARS. Statistical analyses included descriptive statistics and one-way analyses of variance (ANOVAs).

Several statistically significant results were identified through subscale analyses. Taken together, they indicate a relationship between district size and a superintendent's perception of how STARS has affected high school teachers and instruction.

Three main themes emerged from the data:

1. Overall, superintendents perceived STARS as having had no significant effect on high school teachers, instruction, students, or resource allocations.
2. Overall, superintendents perceived STARS to have had a positive effect on high school building principals' leadership practices ($M=3.78$, $SD=.62$).
3. Specifically, superintendents in districts with student populations of 100-1000 (81% of respondents) perceived the effect of STARS on high school teachers and on

instruction as less positive than do superintendents in school districts of 2001-4000 students (10% of respondents).

Results

Theme 1: No Perception of Significant Effect

Overall, superintendents perceived that the Nebraska system (STARS) did not have a significant effect on the high school teachers, instruction, students, or resource allocations in the district. Of the 42 survey questions, only nine individual mean scores showed an average difference of one point or more from the neutral score. Of the six subscales, no averaged mean scores were one point or more from the midpoint.

Superintendents felt that STARS had created more committee work responsibilities ($\underline{M}=4.60$, $\underline{SD}=.63$) for high school teachers and certainly had not reduced the number of workshops they must attend ($\underline{M}=1.49$, $\underline{SD}=.72$). This is understandable since the Nebraska system hinges on teacher input in developing local district assessments. Its very name explains this outcome: School Based Teacher Led Assessment and Reporting System. Workshops and committees are integral parts of this process.

On the other hand, superintendents also perceived that many teachers are gaining knowledge about assessment development and becoming better judges of assessment quality because of STARS ($\underline{M}=4.15$, $\underline{SD}=.74$). This probably is a by-product of the system established in Nebraska. Assessment literacy courses have emerged in several institutions and the state university system has created a program of 18 semester credit hours to educate teachers and administrators in assessment development and the statistical analysis techniques needed to support STARS (Lukin et al. 2004). In addition, the state mandates that at least one person in each district have a working knowledge of assessment literacy and the requirements of meeting the quality criteria outlined by STARS before it will accept a district's assessment portfolio.

Still, committee participation, workshop attendance, and a budding growth in assessment literacy do not lead superintendents to perceive any truly important changes in teachers' lives. Overall, superintendents felt that STARS had relatively little effect on high school teachers ($\underline{M}=2.66$, $\underline{SD}=.54$). Whether this perception is accurate may be open to question since it contradicts the responses obtained from high school principals (Weichel, 2002) and fourth grade teachers in Nebraska (Beran, 2003). To whatever extent this might be a misperception may reflect the fact that organizational leaders often find it difficult to assess the attitudes of front line employees (Fulk & Mani, 1986; Tesser & Rosen, 1975). Superintendents may not regularly interact with the teachers in their high schools and their perceptions may be shaped more by what they hear – or the absence of comment – than by what they observe or are directly told. It also may be that teachers are not yet speaking up to superintendents about the effect STARS has had on their day-to-day lives.

Another possible reason superintendents see little effect on teachers is because individual teachers are not matched with low student performance in STARS reporting, so there is no individual accountability pressure (Plake, Impara, & Buckendahl, 2004). Teachers may not feel enough concern to provoke comments sure to reach the superintendent's ears, such as through union activities, contract negotiations, or grievance proceedings.

Since superintendents perceived that STARS had not had a great effect on high school teachers, it makes sense that they don't see much effect on instruction either. While superintendents perceived that teachers see record keeping as a major time constraint ($\underline{M}=1.64$, $\underline{SD}=.72$), they don't generally believe that the STARS process takes so much time that it prevents teachers from helping individual students. STARS also is not perceived as causing teachers to move more quickly through the curriculum, or to spend less time teaching the broader content of their subjects and more time specifically preparing students for the tests. This perception, though, might also be a result of superintendents' minimal knowledge of teacher day-to-day activity. A major problem for every organizational leader is understanding and appreciating the pressures faced by those on the firing line (Norton, Webb, Dlugosh, & Sybouts, 1996).

In the superintendents' eyes STARS did not limit the curriculum to those items assessed in the exams. This perception appears to be in conflict with national surveys (Pedulla, 2003) and with the perceptions of fourth grade teachers in Nebraska who felt that the standards were leading to a narrowing of the curriculum (Beran, 2003). This discrepancy may result from the differences between teaching at an elementary and at a high school. Teachers at the elementary level generally teach across more subject areas than do subject-specific high school teachers. The multiple subject area demands may cause some elementary teachers to feel they must narrow the curriculum in order to teach all of the required standards.

Superintendents also did not perceive that STARS had a measurable effect on high school students ($\underline{M}=2.83$, $\underline{SD}=.68$). Scores from this subscale vary no more than one point from neutral. Weichel (2002) obtained similar responses from high school principals. In fact, this perceived lack of student effect may range across all grades. Fourth grade teachers in Nebraska didn't perceive any effect on their students either when Beran surveyed them in 2003. Rising scores on district-created assessments may not necessarily reflect increases in learning overall (Abrams & Madus, 2003; Popham, 2004c).

One possible explanation for this perception is that there may not be an effect as yet. It might be that the STARS program has not been in place long enough to produce a discernible difference in high school student achievement. Since the STARS assessments are aimed at influencing and improving curriculum and instruction over the years from fourth grade on, effects may not be realized until the current fourth graders reach high school.

Another reason may be that superintendents do not see significant high school achievement changes in norm referenced tests (NRT). The only NRT analysis conducted in Nebraska (Isernhagen & Dappen, 2004) compared class 3, 4, and 5 school districts (representing 94% of the Nebraska student population) in reading at grades four, eight and eleven. The analysis compared the number of students in the top two quartiles of a reading NRT from 2001 to 2003. The results showed only a 1.19% ($n=180$ districts) improvement at the eleventh grade level. The modest increase in these norm referenced scores seems unlikely to lead superintendents to assume that STARS assessments have a measurable effect on student achievement.

If STARS did have an effect on high schools, it probably was found in finance. Superintendents strongly believed that the costs to implement the Nebraska system (STARS) have increased at the district level ($\underline{M}=4.19$, $\underline{SD}=.85$), and did not think that the state

provided adequate financial support ($\underline{M}=1.83$, $\underline{SD}=.87$). Although questions about STARS financing resulted in strong perceptions from individual superintendents, the general opinion revealed by the subscale as a whole was that STARS does not have a greatly negative effect on district resources ($\underline{M}=2.87$, $\underline{SD}=.61$).

One possible reason for the individually strong negative opinions that did surface is that tests are expensive to create and administer. First, districts are expected to create and publish at least six tests. This is one test in each of two subject areas at each of the grade levels (grades 4, 8, and 11) following specific guidelines for assessment development. Second, the Nebraska Department of Education has set requirements for how districts may spend STARS funding. The guidelines mandate that 85% of the funds be spent on salaries and fringe benefits, while the remainder may be used to pay for contracted services (10%) and administration (5%). This allocation may not be congruent with what a superintendent sees as needed for staff development registration costs, data analysis, or test construction and printing. Superintendents may object to redirecting district funds slated for other needs to assessments.

At the same time, however, superintendents did not perceive that the costs of STARS had required them to reassign high school teachers, reduce elective courses, reduce field trips, or lower expenditures in other areas. Still, even though districts may not have been forced to reduce expenditures in other areas, some may have had to cap allocations in certain areas to meet the cost of STARS implementation.

An interesting consideration obtained from this research was the mild satisfaction with which the superintendents perceived the whole STARS process. They just don't see the portfolio ratings, student performance ratings, assistance from the Nebraska Department of Education, or STARS requirements as having had any great effect on the high schools in their district ($\underline{M}=2.96$, $\underline{SD}=.82$). Although the state mandated timeline for reporting student achievement and submitting the portfolio seems reasonable to them, the superintendents did not perceive that the Nebraska system (STARS) had a significant effect on high school student achievement. Again, this could be due to the relative newness of this process in Nebraska.

Theme 2: Positive Effect on Principals' Leadership Practices

Although superintendents didn't see much effect of STARS on high school students or teachers, they did generally perceive STARS to have a positive effect on high school building principals' leadership practices ($\underline{M}=3.78$, $\underline{SD}=.60$). Overall, they believed that STARS was causing high school building principals to gain knowledge about assessments ($\underline{M}=4.12$, $\underline{SD}=.69$), send more communications than in the past regarding school progress to staff ($\underline{M}=4.02$, $\underline{SD}=.74$), and provide more in-service assessment training for teachers than in the past ($\underline{M}=4.01$, $\underline{SD}=.84$).

One likely reason for the superintendents' positive perception of improved leadership among their high school principals is that they are likely to be the leader and contact person in compiling the data required for the STARS state reporting. Superintendents may rely on building leadership to complete the high school portion of the district portfolio and to assemble student achievement data. In turn, the building leadership probably has to work more with classroom teachers to develop assessments and to ensure that accurate information is collected and submitted for state reporting. This squares with the superintendents'

perceptions reported earlier, that high school teachers have increased their committee work participation and seen no reduction in the number of workshops they must attend.

Weichel (2002) found that high school principals in Nebraska expected the state system to have a negative effect on their role as the school leader because it would reduce their discretionary time and add pressure. Weichel reasoned that because principals are looked to for direction and guidance when new processes are implemented (Lashway, 2000), they will likely have to commit more time to faculty development. Additionally, many Nebraska high school principals were required to learn the STARS reporting process in order to implement it since they will be required to administer tests, interpret test scores, help teachers increase their assessment literacy, arrange professional development opportunities, and encourage change in the classroom (Weichel, 2002). While principals may see these demands as further burdens in an already overloaded job, superintendents may see these activities as fostering greater interaction and shared decision-making.

Theme 3: Perception Differences Based on District Size

The perceived effect of the STARS system differed among superintendents by district size. Specifically, superintendents in districts with population ranges of 100-1000 students (81% of respondents) are less positive about STARS and its effect on high school teachers and daily instruction at a statistically significant level than are those leading school districts with ranges between 2000-5000 students (10% of respondents). This is worrisome since a majority of the superintendents responding to this survey in Nebraska oversee districts with populations that range between 100-1000 students. On the contrary, perceptions were generally neutral from superintendents of districts above 5000 students (4% of respondents).

It may be that much of the discrepancy in perceptions can be attributed to funding differences. Those who are receiving less money for implementation are also those who are less positive about the state system. According to Harmon and Branham (1999), small schools may have fewer resources to support assessment development and planning. This may also partly explain why Weichel (2002) found that high school principals' perceptions of STARS also varied with school size.

Recommendations for Practice

Increase State Funding to Support STARS

The results of this study, like Beran's (2003) study of how STARS had effected fourth grade teachers, argued for a distinct need to increase the amount of funding allotted to Nebraska districts for assessment development, administration, and data analysis. Over 80% of the responding superintendents were in districts with populations under 1000 students. The perception of inadequate funding for assessment/accountability requirements may be why they believed STARS had less of a positive effect on both teachers and instruction in their districts. The collective perception of superintendents in rural Nebraska school districts suggested, as Harmon and Branham (1999) argued that implementing standards brings focus and clarity to student expectations, but the efforts required sufficient resource support to produce improved results.

On the other hand, superintendents in large urban districts are also concerned about the lack of funding. Due to the complexities that come with size, large districts often do not have the resources available to adequately facilitate reform processes (Hannaway & Kimball, 1998) such as the assessment development required in Nebraska (Glissman, 2005).

Currently, school districts are provided STARS funding based on their student populations. Seven of the eight state categories are used to group districts with populations up to 17,999 students. The eighth category is comprised of three districts with 18,000 or more students. These systems have populations of approximately 46,000, 32,000, and 20,000 students. The grouping is artificial and flawed. Because the number is small, the grouping is expedient, and it masks critical differences. Not only does the number of students differ widely across these three, but so do the cultural diversity and poverty levels. Obviously, the needs of each of these school districts vary, and providing them with the same amount of funding is not appropriate.

The larger urban school districts also objected to the rules for how to spend STARS funding. In November of 2004, the Metropolitan Omaha Education Consortium (MOEC) Assessment Task Force submitted a letter to the Nebraska Commissioner of Education requesting additional funds and flexibility in meeting the requirements of STARS reporting. Many of the districts in the metro area were concerned that restricting spending for contracted services to ten percent of the total funding was too limiting. Contracted services may include, but not be limited to, paying for staff development consultants, paying for teachers to attend professional development training on assessments, hiring additional staff to process and analyze data, or covering the costs of printing tests. The commissioner responded that increasing spending flexibility was not possible at this time. He said that the rules would remain the same since the funding is supported by federal monies and must conform to the intentions of the No Child Left Behind legislation. He reiterated that STARS funding must be used for standards development, assessment processes, and school improvement efforts *conducted by teachers* (emphasis added). He encouraged districts to spend their local funds for contracted service costs running over the state-allotted amount (D. Christensen, personal communication, November 8, 2004). This urban district concern needs to be readdressed by the State Department of Education.

Assessment Literacy

Is it possible that teachers across Nebraska are sufficiently literate in assessment to make this statewide system work? Assessment literacy means knowing how to improve learning by responding to needs defined by student assessment data. Teachers must identify which students are mastering the required knowledge, and then use the data from those assessments to make changes in teaching practices (Jerald, 2003; Popham, 2004a). According to Swaffield and Dudley (2003), educators needed to become assessment literate in order to make educated decisions about both assessment methods and the use of assessment data. The State of Nebraska and the Buros Center for Testing have provided assistance to teachers in understanding the technical aspects of the six quality criteria required for submitting the portfolio for state reporting of STARS (Plake, Impara, & Buckendahl, 2004), but much work is still needed to provide them with a practical understanding of how assessment relates to student learning and their own teaching. Many educators are not sufficiently literate in basic assessment to understand whether their achievement results show significance (Swaffield & Dudley, 2003), and a national survey revealed that only 30% of teachers believed that a state-mandated testing program is worth the time and money (Pedulla, 2003).

Maybe an assumption is being made by the Nebraska Department of Education that teachers in Nebraska are embracing standards-based teaching and instruction. It is

appropriate to ask if teachers are focusing on standards or only following the outline of the textbook on what should be discussed, learned and assessed. A State Board of Education goal is to have all teachers and administrators trained in the use of STARS by 2008. Stiggins (2004) argued that while it is important for administrators to be grounded in assessment literacy before embarking on a school improvement plan, no such plan can succeed without informed teacher participation and commitment.

Teachers in Nebraska were scattered in terms of their involvement with and understanding of STARS. Less than 50% of the teachers surveyed or interviewed in focus groups in Nebraska reported involvement in the alignment of the curriculum to the state standards, the scoring of assessments, or the development of assessments (Bandalos, 2004). Even fewer teachers (less than 25%) had been involved in re-teaching activities following the assessments, or in assisting with the district's portfolio (Bandalos, 2004).

Nebraska has tried to address the low level of teacher understanding by investing in assessment literacy courses that could lead to an 18-hour assessment endorsement. Three cohorts, approximately 150 teachers from across the state, have completed the endorsement program since 2001. The program, called the Nebraska Assessment Cohorts (NAC), involved increasing the assessment literacy of teachers and administrators to improve classroom assessment practices. Since the creation of the assessment endorsement, a few localized learning teams and pre-service assessment literacy programs have been established (Lukin et al. 2004). Although the efforts to increase assessment literacy are noted, more needs to be done to capture a critical mass of teachers. Even if all 150 NAC graduates were high school educators, the overall effect would be next to nothing in Nebraska's 299 high schools.

In addition, more emphasis should be placed on assessment for learning – that is, assessment that promotes students' learning (Black et al., 2004) – among Nebraska educators. This promotes the use of formative assessments to increase student understanding. Many current high school classroom educators were not required to learn assessment techniques as part of their pre-service training. High school teachers need to incorporate more meaningful formative assessment into their programs in order to monitor student progress toward mastery of a given standard. It is when teachers review the results of these assessments with other teachers that they begin to respond to student needs with corrective instructional practices (Rettig, McCullough, Santos & Watson, 2003).

Recommendations for Future Research

Small versus Large Districts

The results of this study showed that superintendents of small and large districts differed in their perceptions of state-mandated assessment reporting, but this may only be scratching the surface. Many rural school district superintendents felt they needed more guidance (Davis, 2004) and funding. Small rural districts often felt forced to pool their resources with neighboring school districts in order to reap any benefits from state funding (Tyler, 2003; Tyler & Kibby, 2004; Williams, 2003). Small rural district superintendents may perceive that STARS places a greater burden on them, their principals, and their teachers than on those in larger districts (Tyler, 2003) since they do not have the staff to specifically dedicate to assessment development and coordination. Assessment demands put enormous pressure on small districts in finding the needed resources, time, and expertise (Gallagher, 2003).

Unable to centralize the problems, many small districts leave the coordination of assessment development responsibilities to small groups of classroom teachers (Bandalos, 2004; Gallagher, 2003).

Large districts, on the other hand, have more staff to help with assessment development and implementation, but are not as able as small districts to implement changes quickly (Hannaway & Kimball, 1998). Large districts are also more likely to have less teacher participation, proportionate to their size, in their assessment design process (Gallagher, 2003). This may also mean less real contributions by the teachers who are involved, and hence less decision-making ability. The results from this study showed that more research is needed on identifying the differences between the effect assessments have on small and large districts.

Elementary versus Secondary School Impact

Superintendents' perceptions of the effect of STARS on elementary schools are unknown as yet. The results from this study showed a need to research whether the effect is different at the elementary level in comparison to secondary. Pedulla (2003) showed that elementary and middle school teachers felt a greater amount of stress and pressure as a result of teaching in a state with statewide testing programs than do high school teachers. The pressure on elementary teachers to teach multiple subjects well is greater than at the high school level where teachers may teach only one state reported subject.

In addition, future research may focus on whether superintendents perceive the match or alignment of standards in elementary versus secondary classrooms differently, therefore the effect of STARS differently. Local control in Nebraska creates various methods and forms of standards alignment and articulation. This appears to be more notable in the elementary grades. Since elementary curricula may not align with the state standards as well as the high school curriculum does, more instructional change might be needed at the elementary level to master the standards (Pedulla, 2003).

Assessment Development Comparisons

The Nebraska system is by far a better choice than state-generated assessments that do not align with the curriculum development cycle established in each district. Superintendent responses, however, were driven by their perceptions of their district-created assessments. If they perceived their assessments as more rigorous and demanding than other districts, the system failed to hold each district to the same expectations of excellence.

Nebraska may need to address concerns from superintendents and the public that district-generated assessments do not necessarily mean districts create relatively equivalent assessments (Gallagher, 2003). To date there has not been a substantive review of the content of teacher-developed assessments in terms of question and task quality (Plake, Impara, & Buckendahl, 2004). We cannot effectively move forward until this is accomplished.

Summary

In summary, the results of this study showed that Nebraska superintendents generally did not see that the STARS system had any great effect on the high school teachers, students, instruction, or resource allocations in their districts. They did perceive that STARS generates funding concerns and had some positive effect on high school building level leadership

practices. The results also showed that superintendents in large districts perceived STARS's effect differently from those with small populations.

Final Thoughts

Do superintendents feel that the process of STARS is reasonable, but the results aren't worth it? The answer really isn't clear, but maybe more time is needed to see the effect of STARS in Nebraska. The process is still a new phenomenon. We may need to wait to see how our fourth grade students fare through the complete STARS process before the ultimate verdict is in on the worth of this system.

A limitation to this study was that it cannot be generalized beyond Nebraska, since STARS is a state program established on the premise that each district should create its own assessments for reporting purposes. The success of the Nebraska system resides heavily on its teachers and principals. The ultimate test for Nebraska's assessment/accountability system is whether other states will see this process as superior to the single statewide test system. If other states begin to adopt a School-based Teacher-led Assessment and Reporting System (STARS), then the Nebraska system may increase its credibility with superintendents, principals and teachers both here in Nebraska and nationally.

References

- Abrams, L., & Madaus, G. (2003). The lessons of high-stakes testing. *Educational Leadership*, 61 (3), 31-35.
- Bandalos, D. (2004). Can a teacher-led state assessment system work? *Educational Measurement: Issues and Practice*, 23 (2), 33-40.
- Beran, R. (2003). *Fourth grade teachers' perceptions of the Nebraska's state standards process in elementary schools*. Unpublished dissertation, University of Nebraska at Omaha.
- Black, P., Harrison, C., Lee, C., Bethan, M., & Wiliam D. (2004). Working inside the black box: assessment for learning in the classroom. *Phi Delta Kappan*, 86 (1), 8-21.
- Christensen, D. (2001- December). Building state assessment from the classroom up. *The School Administrator*. 58 (11), 27-28, 30-31.
- Davis, M. (2004). Rural schools in need of guidance, GAO says. *Professional Development Collection*, 24 (6), Retrieved November 17, 2004, from http://web10.epnet.com/citation.asp?tb=1&_ug=sid+B1BB21C4%2D31D0%2D4DB7%2...
- Duke, D., Tucker, P., & Heinecke, W. (2000). *Initial responses of Virginia's high schools to the accountability initiative*. Charlottesville, VA: Thomas Jefferson Center for Educational Design.
- Fulk, J. & Mani, S. (1986). Distortion of communication in hierarchical relationships. In: *Communication Yearbook 9*, McLaughlin, M.L. (ed.), pp. 483-510. Sage Publications, Beverly Hills.
- Gallagher, C. (2003). Charting STARS: Sustainability as challenge and opportunity, year two report. Retrieved January 21, 2005 from World Wide Web <http://www.nde.state.ne.us/stars/index.html>.
- Glissman, B. (2005, February 26). Nebraska test system questioned. Omaha World Herald, pp. B1, B2.
- Hannaway, J & Kimball, K. (1998). Big isn't always bad: School district size, poverty, and standards-based reform. Washington, D. C.: Urban Institute. (ERIC Document Reproduction Service No. ED439184).

- Harmon, H., & Branham, D. (1999). Creating standards for rural schools: A matter of values. *The High School Magazine*, 7 (4), 14-19.
- Isernhagen, J. & Dappen, L. (2004). An examination of 2001 to 2003 Nebraska criterion referenced tests, norm referenced tests, and district portfolio ratings for reading at grades 4, 8, and 11. In C. Gallagher (Ed.), *Charting STARS: New conversations, year three report* (pp. 80-86). Retrieved January 21, 2005 from World Wide Web <http://www.nde.state.ne.us/stars/index.html>.
- Jerald, C. (2003). Beyond the rock and the hard place. *Educational Leadership*, 61 (3), 12-16.
- Johnson, M.(1981). *An analysis of administrator perceptions on mandatory evaluation programs in elementary and secondary schools*. Pullman, WA: Washington State University.
- Lashway, L. (2000). Who's in charge? *Principal Leadership*, 1(3), 9-13.
- Lukin, L., Bandalos, D., Eckhout, T., & Mickelson, K. (2004). Facilitating the development of assessment literacy. *Educational Measurement: Issues*
- Nebraska State Department of Education (2002b, July). Nebraska's comprehensive plan for school improvement. Nebraska State Department of Education [Homepage](http://www.nde.state.ne.us). Retrieved August 1, 2002, from the World Wide Web: <http://www.nde.state.ne.us>
- Norton, M.S., Webb, L.D., Dlugosh, L.L., and Sybouts, W. (1996). *The School Superintendency: New Responsibilities, New Leadership*. Needham Height, MA: Allyn & Bacon.
- Pedulla, J. (2003). State-mandated testing what do teachers think? *Educational Leadership*, 61 (3), 42-46.
- Plake, B., Impara, J., & Buckendahl, C. (2004). Technical quality criteria for evaluating district assessment portfolios used in the Nebraska STARS. *Educational Measurement: Issues and Practice*, 23 (2), 12-16.
- Popham, J. (2004a). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62 (1), 82-83.
- Popham, J. (2004c). A game without winners. *Educational Leadership*, 62 (3), 46-50.
- Rettig, M., McCullough, L., Santos, K., & Watson, C. (2003). A blueprint for increasing student achievement. *Educational Leadership*, 61 (3), 71-76.
- Roschewski, P., Gallagher, C., & Isernhagen, J. (2001). Nebraskans reach for the STARS. *Phi Delta Kappan*, 82 (8), 611-615.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86 (1), 22-27.
- Swaffield, S. & Dudley, P. (2003). Wising up to assessment literacy. *Education Journal*, (67), p. 9.
- Tesser, A. & Rosen, S. (1975). The reluctance to transmit bad news. *Advances*
- Tyler, C. (2003). NCLB: Tall order for small districts. *Leadership*, 33 (1), 26-27.
- Tyler, C. & Kibby, R. (2004). Small districts, sharp axes. *Leadership*, 33 (5), 16-20.
- Weichel, M. (2002). *Nebraska public high school principals' perceptions of how state standards impact schools*. Unpublished dissertation. Omaha, NE: University of Nebraska at Omaha.
- Williams, D. (2003). Rural routes to success. *Educational Leadership*, 61 (3), 66-70.

APPENDIX I
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 and 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** The Nebraska accountability system (both state and AYP) classifies schools or school systems correctly (avoids misclassifications.)
- Goal Five** The consequences of the accountability system are positive ones and contribute to the achievement of the districts in the state accountability goals.

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>Is the DAP process reliable and accurate?</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>Inside Evaluation of Process</p> <p>On-site visitations</p> <p>Surveys of local educators</p> <p>Surveys, case studies</p>	<p>Buros</p> <p>Ellen Forte-Fast/Dave Frisbie</p> <p>Eight Nebraska Participants</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>
Consortiums, Collaborations, Individual Districts	How do educators working in groups handle local assessment differently from educators in independent districts?	Survey/case studies	UNL evaluation

GOAL 2

2004-2005

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 Jody Isernhagen & Leon Dappen Jody Isernhagen & Leon Dappen NDE/UNL

GOAL 3

2004-2005

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>Have STARS 2001 reading data improved? 2005 reading data?</p> <p>Have NRT and CRT reading data improved from 2001, 2002, 2003, 2004? 2005?</p> <p>Have STARS 2002 mathematics data improved from 2002 – 2004 mathematics data? 2005?</p> <p>Have NRT and CRT mathematics data improved from 2001, 2002, 2003, 2004? 2005?</p>	<p>Correlational Studies</p> <p>Correlational Studies</p> <p>Correlational Studies</p>	<p>UNL</p> <p>UNL</p> <p>UNL</p>
<p>Student performance on Statewide Writing</p>	<p>Have statewide writing results improved? Grade 4 - 2002/2004 Grade 8 – 2003/2004</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>UNL</p> <p>UNL</p> <p>UNL Evaluation</p>
<p>Graduation Rate</p>	<p>Has the graduate rate improved from 2003-04 to 2004-05?</p>	<p>Correlational Studies</p>	<p>NDE</p>

GOAL 4

2004-2005

Does the Nebraska accountability system (both state and AYP) classify or misclassify schools?			
Indicators for Study	Research Questions	Method	Who?
Internal Data Validation	What does NDE do internally to insure the accuracy of its data?	AAR Business Rules Flow Chart	NDE
Decision Verification	How does NDE know its accountability decisions are accurate?	Internal Audit	

GOAL 5

2004-2005

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	What kinds of things are happening in schools where accountability goals aren't being met?	Case Studies Visitations Survey	UNL Evaluation NDE Jody Isernhagen & Leon Dappen
Emerging consequences	What kinds of things are happening in schools where accountability goals are being met?	Case Studies Survey	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 Jody Isernhagen & Leon Dappen

APPENDIX I
IRB APPROVAL LETTER

November 3, 2004

Dr. Jody Isernhagen
H&E Admin
132 TEAC 0360

IRB # 2001-12-108 BX

TITLE OF PROPOSAL: **Study of School-based Teacher-led Assessment and Reporting System**

Dear Dr. Isernhagen:

The Institutional Review Board for the Protection of Human Subjects has completed its review of the Request for Change in Protocol submitted to the IRB.

1. Enclosed is the IRB approved Informed Consent form for this project. Please use this form when making copies to distribute to your participants. If it is necessary to create a new informed consent form, please send us your original so that we may approve and stamp it before it is distributed to participants.
2. It has been approved to make Jody Isernhagen the principal investigator and Chris Gallagher the secondary investigator. Other personnel changes have been noted as well.
3. Changes to the methods and participant categories have been approved.
4. It has been approved to enlarge research to include reanalysis of data collected as Component C, year three.
5. The research instruments have changed. You will now be using STARS Survey for Teachers, School Administrators, District Administrators, and Educational Service Units. There are three versions of each: Collaboration, Consortium, and Individual Districts.
6. Year three research will continue. Component B changing from language arts to math and will become Component A.
7. It has been approved to include Component B - survey superintendents of ~259 school districts across Nebraska.

This letter constitutes official notification of the approval of the protocol change. You are therefore authorized to implement this change accordingly.

APPENDIX J
2004-2005 PRIMARY RESEARCH SURVEY SAMPLE

[Surveys were changed to reflect appropriate configuration and role of the participant]

STARS Survey Sample

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:**
 Language Arts Teacher **Grade Level:** Elementary Middle School High School
 Math Teacher **Grade Level:** Elementary Middle School High School
 Other—Please identify: Job Title _____ Grade _____
2. **My Gender:** Male Female
3. **Years of experience in my primary role:** _____
4. **My school is participating in a STARS-related collaboration:** YES NO
My collaboration’s name is: _____
My school joined the collaboration (Approximate date): _____
5. **My school’s primary reason for joining the STARS collaboration:**

6. **The main advantage of my school’s participation in our present collaboration is:**

7. **The main limitation of my school’s participation in our present collaboration is:**

	None of the time	Very little of the time	Some of the time	Most of the time	All of the time
DIRECTIONS: Please circle the number that best describes your response for your collaboration.					
Collaboration Support					
8. My collaboration supports “school-based teacher-led” assessment.	1	2	3	4	5
9. My collaboration provides services that my school needs for the implementation of STARS.	1	2	3	4	5
10. My collaboration provides on-going assessment training for school administrators.	1	2	3	4	5
11. My collaboration provides on-going assessment training for teachers.	1	2	3	4	5
Directions: Please circle the number that best describes your response for your school or collaboration.					
	None of the time	Very little of the time	Some of the time	Most of the time	All of the time

Assessment Literacy					
1. In my school, educators are committed to improving their own assessment competence.	1	2	3	4	5
2. Students in my school are involved in understanding their own progress and achievement status.	1	2	3	4	5
3. Teachers in my school write their own STARS assessments at the classroom level.	1	2	3	4	5
4. Teachers in my school select assessment items from a common bank/pool provided by the collaboration.	1	2	3	4	5
5. Administrators in my school support "school-based teacher-led" assessment.	1	2	3	4	5
6. Teachers in my school support "school-based teacher-led" assessment.	1	2	3	4	5
7. My collaboration develops standards-based assessments for my school.	1	2	3	4	5
8. My collaboration develops the STARS assessment portfolio sent to NDE** for my school.	1	2	3	4	5
9. Teachers in my school are involved in designing assessment items for the collaboration.	1	2	3	4	5
10. All the schools in my collaboration use the same grade-level assessments to meet STARS requirements.	1	2	3	4	5
11. Teachers in my school collaborate on designing assessments for STARS.	1	2	3	4	5
12. Teachers in my school are confident in their ability to design valid and reliable assessments.	1	2	3	4	5
13. Since the inception of STARS, teachers in my school use rubrics in assessment.	1	2	3	4	5
14. My collaboration supports the growth of teachers' assessment literacy.	1	2	3	4	5
15. My collaboration supports the growth of administrators' assessment literacy.	1	2	3	4	5
16. Teachers in my school participate in learning teams to improve their assessment skills.	1	2	3	4	5
Data					
17. My school receives assessment data in a format that allows for disaggregation by:					
a) district results	1	2	3	4	5
b) school results	1	2	3	4	5
c) classroom results	1	2	3	4	5
d) student results	1	2	3	4	5
18. My school receives data in a format that allows for disaggregation by relevant populations (free and reduced lunch, gender, etc.) for individual students.	1	2	3	4	5
19. My collaboration scores all STARS assessments for my school.	1	2	3	4	5
20. My collaboration provides timely return of data.	1	2	3	4	5
21. My collaboration interprets assessment results.	1	2	3	4	5
22. My collaboration provides recommendations for responding to assessment results.	1	2	3	4	5
Instructional Impact					
23. My collaboration has helped my school align curriculum to state standards.	1	2	3	4	5
24. My collaboration has helped my school establish benchmarks for meeting state standards at kindergarten through 12 th grade.	1	2	3	4	5
25. Teachers in my school reviewed units of study for alignment to state/local standards.	1	2	3	4	5
26. Instructional units are assigned to appropriate grade levels in my school.	1	2	3	4	5
27. Teachers in my school develop clear and appropriate instructional targets based on assessment results.	1	2	3	4	5
28. In my school, lesson planning is aligned to assessment data and state standards.	1	2	3	4	5
Directions: Please circle the number that best describes your response for your school or collaboration.	None of the time	Very little of the time	Some of the time	Most of the time	All of the time
29. In my school, assessment is integrated into instruction and is used to inform	1	2	3	4	5

teaching and learning.					
30. My school's assessments accurately measure what my students know and can do.	1	2	3	4	5
31. Teachers in my school modify their instruction in cases where students did not perform well on an assessment.	1	2	3	4	5
32. Teachers in my school use rubrics in instruction since the inception of STARS.	1	2	3	4	5
33. The results of my school's assessments are helpful in identifying individual student strengths and weaknesses.	1	2	3	4	5
34. Teachers in my school share successful instructional strategies.	1	2	3	4	5
35. My school's assessments give teachers important feedback about how effectively they are teaching.	1	2	3	4	5
Support					
36. My district provides time and resources for "school-based teacher-led" assessment.	1	2	3	4	5
37. My ESU* provides on-going leadership training for school administrators for implementing the STARS process.	1	2	3	4	5
38. My ESU* provides on-going assessment training for teachers in my school.	1	2	3	4	5
39. My ESU* provides data retreats for my school.	1	2	3	4	5
40. My ESU* provides software for scoring, analyzing, and reporting data to support my school.	1	2	3	4	5
41. NDE **provides leadership training for all administrators in my school for implementing the STARS process.	1	2	3	4	5
42. NDE** provides assessment training for teachers in my school.	1	2	3	4	5
43. List any other organizations/sources from which your school receives assessment training and leadership support for implementing the STARS process. <hr/>					
44. Please use the back of this page to make any additional comments you might have regarding the STARS process.					

*Educational Service Unit

**Nebraska Department of Education

PLEASE COMPLETE AND RETURN NO LATER THAN **DECEMBER 15, 2004**. 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 1
2004-2005 PRIMARY RESEARCH INTERVIEW
PROTOCOL

STARS RESEARCH INTERVIEW PROTOCOL FOR ESU STAFF DEVELOPERS
January – March 2005

Qualitative Research Purpose: Explore and understand the differences identified in perceptions of STARS on assessment literacy, use of data in classroom settings, leadership, and support from external agencies on consortia, collaborations, and independent school districts.

Date of interview: _____ **Time of interview:** _____

Location of interview: _____

Interviewer: _____

Participant Profile

Participant: _____

Educational Service Unit: _____

Position: ___ Director of Staff Development ___ Staff Developer

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 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 ESU staff is reported or released in aggregated form only. Districts, schools, individuals, and ESUs are not identified.
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 educational service unit. 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 the districts served by your Educational Service Unit. Questions about assessment literacy, use of data in classroom settings, instructional impact, leadership, and support from external agencies within the districts served within your educational service unit are specific topics of interest.

Interviewer Only:

If there is a question about the school district's identification as a collaboration/consortium /district, you want to clarify by using the NDE definition and stating that you have spoken to their contact person and this is the correct terminology for your district.

8. Please feel free to discuss your views openly. From time to time, I may have additional questions to further understand a concept or idea that you have shared.
9. Let's begin. Please state your name, educational service unit and indicate permission to record this interview by repeating this statement, "I (your name) a staff developer at (educational service unit number) located in (name of city) 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. As a staff developer what has been your primary role in the STARS process with each collaboration/consortium/individual district?

Probes

- ___ a. What was the role of your educational service unit with collaborations/consortia/individual districts in the assessment process?
- ___ b. Can you tell me a little bit about your role in the process that was used for the development of classroom assessments for STARS with collaborations, consortia, and individual districts?
- ___ c. What does your educational service unit see as the value for school districts in belonging to a collaboration/consortium/district?

Descriptive Notes:

Reflective Notes

2. What new learnings have you had due to your educational service unit's involvement in the STARS assessment process within collaborations/consortia/individual districts?

Probes:

- ___ a. You have stated (one, two or what ever has been stated) new learnings due to your involvement with districts in collaboration/consortium/district. Are there others?
- ___ b. What was the value of your new learnings to you as a professional and to your educational service unit?

Descriptive Notes:

Reflective Notes:

3. What type of data do you collect from district assessments for school districts and how do you use that data in improving student performance in collaboration/consortia/individual districts?

Probes:

- ___ a. How have you been involved in the scoring and reporting of assessments in collaborations/consortia/individual districts?
- ___ b. How are data disaggregated in collaborations, consortia, individual districts (free and reduced lunch, gender, Title I) and what is your role?
- ___ c. How do you interpret the data for/with districts once it is given to you? Do your collaborations/consortia/individual districts support your interpretation of data?
- ___ d. How do you use data to develop and assist in the implementation of interventions in collaborations/consortia/individual districts?
- ___ e. What is the role of your Educational Service Unit in scoring and reporting, disaggregation, interpreting, and using data for interventions for collaborations, consortium, and individual districts?

Descriptive Notes:

Reflective Notes:

4. What major obstacles have you faced with your collaborations/consortia/individual districts during the implementation of STARS?

Probes:

___ a. Are there other obstacles that I should know about?

Interviewers Only: Ask this probe only if the participants bring one of these topics up.

___ b. Tell me about what you mean when you say you need more time, leadership, NDE support, or change etc. What would you do with extra time, etc.?

Descriptive Notes:

Reflective Notes:

5. What would you like for your collaboration/consortia/individual districts next steps to be to better implement the STARS process?

Probes:

___ a. Can you tell me more about that? (If you need additional clarification on a topic)

___ b. Do you have some ideas of how these steps could be implemented?

Descriptive Notes:

Reflective Notes:

6. How could the collaboration/consortium/individual districts in your ESU better implement the STARS process?

Probes:

___ a. What would you like to see to the collaboration/consortium/districts do in the future that they are not doing today?

Descriptive Notes:

Reflective Notes: