

DEVELOPING A LOCAL ASSESSMENT SYSTEM FOR STATE ACCOUNTABILITY: A DISTRICT'S REFLECTIONS

Toby Boss - *Educational Service Unit #6*
Daniel P. Endorf - *York Public Schools*
Chad W. Buckendahl - *University of Nebraska, Lincoln*

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, mathematics, science, and social studies. This article briefly describes the state's requirements for local assessment systems and explores language arts teachers' experiences during the development of reading assessments in one suburban school district. The authors used a structured open-ended questionnaire to gather data on teachers' perspectives from this common shared experience. Themes that emerged from the data are reported and recommendations for developing assessments and building capacity at a district level are also provided.

INTRODUCTION

State student assessment and accountability systems received increased national attention with the passage of No Child Left Behind (NCLB) in January 2002. This is the next step in the progression of control over methods of assessment and accountability that has shifted from local school districts to state departments of education and legislative agencies and now to federal jurisdiction. The shift in control occurred over time for a variety of reasons, but many of these focus on the perceived shortcomings of public education. There is a perception that public education has not reached the expectations placed on it. In the 36th Annual Phi Delta Kappa/Gallup Poll, 47% of all respondents rated their community's public schools as deserving of a grade of A or B, with parents' opinions higher at 61% (Rose & Gallup, 2004). More worrisome is that this number plunges to 26% when referencing the nation's public schools (Id.). Although the public may not be able to define these expectations, the perception has remained reasonably consistent for a number of years. These concerns are constant reminders to the educators who are responsible for addressing them in their schools. As evidenced with the passage of NCLB, much of the increased involvement revolves around measuring student performance.

In response to these concerns, some recent

efforts have sought to re-conceptualize the role of assessment in school districts and the classroom to better inform the public about how students are performing relative to state-defined content standards (Diaz, 2001). More importantly, the integration of curriculum, instruction, and assessment is being re-examined. Shepard (2000) suggests that these three areas are interrelated, yet may be treated as mutually exclusive in practice or superseded by external testing frameworks. Other researchers (e.g., Lane, Parke, & Stone, 2003; Stone & Lane, 2003) have examined the interactions between instructional practices and student performance on a state assessment that used performance assessment. The results of these studies suggested that differences among schools could be attributed to instruction that targeted the goals of the reform effort.

Because of the desired connections among curriculum, instruction, and assessment, it is important that educators understand how to collect quality information and then use it to influence instruction and learning. Teachers who possess assessment literacy skills (e.g., basic understanding of measurement concepts such as validity, reliability, and fairness) are better prepared to engage their students in the targeted goals of learning (Black & Wiliam, 1998). Researchers in the area of classroom assessment tend to agree that better classroom assessment contributes to the potential for making better

curricular and instructional decisions (Black, Harrison, Lee, Marshall, & Wiliam, 2004). However, because teachers spend a reasonable amount of time selecting or developing classroom assessments, they need to have the necessary assessment literacy skills to effectively use assessment as a tool in the learning process (Stiggins, 2004). Thus far, preparing teachers and administrators for their assessment roles in classrooms or schools has been limited in their pre-service education (Stiggins, 1999; Trevisan, 1999). However, even if teachers and administrators had this training, most states have already assumed primary responsibility for the student assessment and accountability systems.

Although it may be possible to inform local curriculum and instruction using assessments designed for state accountability, most state assessments are not currently designed with this goal. Locally developed assessments that are designed to inform instruction and report on student performance serve this goal, but may not coincide with the goals of the state accountability model. The competing purposes of most state and local systems create tensions that place school districts in situations where difficult choices must be made. This tension leads to the creation of parallel curriculum, instruction, and assessment systems, each designed with their own purpose.

Popham (1994, 1999) and Stiggins (2001) have suggested that classroom assessment has an opportunity to inform instruction and facilitate learning given the proper conditions. Questions remain, though, about the utility of this information in conjunction with state assessment and accountability requirements. Nebraska's state assessment system allows local school districts to select and/or design assessments that measure students' performance on state content standards in reading, mathematics, science, and social studies (Roschewski, 2001; Roschewski, 2004). Part of a local district's system may include classroom assessment information.

Because each district may create a unique assessment system, part of the state's accountability system incorporates an external review of the technical quality of the district's

assessment system (Buckendahl, Plake, & Impara 2004). The review elements include 1) assessments are in alignment with content standards, 2) assessments are in alignment with curriculum, 3) assessments are free from bias or sensitive language, 4) assessments are written at an appropriate level, 5) reliability evidence supports the desired inferences, and 6) cut scores are set at appropriate levels. These technical criteria are based on the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999) and described as relative to a local assessment system in Plake, Impara, and Buckendahl (2004). A proposed accountability system that incorporates this technical review element was also described and evaluated by Buckendahl, Impara, and Plake (2002).

After the district's assessment systems are employed during the academic year, information about the technical quality of their assessments and their student performance are reported to the state's department of education. This information is used to produce a state report card highlighting the performance of school districts on the state's content standards with respect to the assessment system's technical quality and their students' performance on those assessments. It is with this state assessment model in mind that we took a reflective, in-depth look at one district's efforts to develop and implement a district assessment system in the first year of the state's program. Because this system potentially placed greater responsibility on classroom teachers, we wanted to explore the attitudes and perceptions of one district's teachers as they participated in a common, shared experience. The results of this study rely on a phenomenological methodology whereby a group of individuals experienced a common event (Creswell, 1998).

The purpose of this study was to explore language arts teachers' experiences during the development of reading assessments to measure student performance on state content standards in one suburban school district. Because the district will develop and revise assessments in three other con-

tent areas: mathematics, science, and social studies, these teachers' reflections and suggestions may inform these subsequent development teams. The goal of this study was not to generalize findings to all educators, but to discover the lessons educators in this district learned during their assessment development activities.

SETTING

We conducted this study in a suburban school district in Nebraska using the district's Reading Criterion Reference Assessment (CRA) team as our participants. The district consists of five towns with populations ranging from under 300 to over 1,000. The district has two elementary schools (K-5), one middle school (6-8) and one high school (9-12) with a total enrollment of approximately 1,700 students, instructed by 126 educators. The average age of these educators was 40.1 years with 14.3 average years of experience in education.

The district created the Reading Criterion Referenced Assessment (CRA) Team in the spring of 2000 to develop assessments to measure the state content standards in reading, writing, speaking and listening. Twenty-six eligible teachers in the district comprised the team representing grades K-12. These teachers' experience ranged from 1 to 37 years, with a mean of 18 years. This suggests that the teachers on the CRA Team had more experience than the average teacher in the district. Language arts teachers from each grade level and building (K-5) and all language arts teachers in grades 6-12 were selected for the development team.

It is important to note that Nebraska does not have an assessment literacy element of teacher certification requirements. Thus, most teachers in this district did not receive assessment literacy training in their pre-service programs, so systematic assessment development presented a new challenge. The lack of assessment training among teachers and administrators nationwide contributes to the challenge because there are few models from which to draw. To facilitate professional development for their teachers in assessment literacy, the district relied on regional Educational Service Unit (ESU) professional development staff to provide guidance

during the development process. These ESUs provide a range of services; however, within the scope of this study, we focused on the staff development services related to assessment literacy.

METHOD

Members of the Reading Criterion Referenced Assessment (CRA) Team met in June of 2001 to review curriculum and analyze student performance data. As part of this meeting, the researchers administered a structured questionnaire consisting of 10 items. Most of the questions were open-ended and focused on teachers' perceptions about the assessment literacy activities they engaged in while they were developing a local assessment system. The theoretical framework for the questionnaire was based on Shepard's (2000) description of learning systems that functionally integrate curriculum, instruction, and assessment. The full questionnaire is provided in the Appendix.

Teachers responded anonymously to the questionnaire, writing their responses on the form provided. Although there was no time limit, most teachers completed their questionnaires within an hour. Responses were transcribed by an administrative assistant external to the study to protect the confidentiality of the respondents. Once this transcribed materials were available, the researchers independently coded the teachers' comments into conceptual categories (Miles & Huberman, 1994). After the researchers did this process independently, they met to share their coding schema and came to consensus on the major themes that emerged from the data. Because the questionnaire was constructed to elicit responses to some conceptual themes that were anticipated from literature, there was little disagreement among the researchers about the major themes that were observed in the teachers' responses.

We recognize that with any qualitative study, caution is warranted in interpreting these results and we want to discourage readers from attempting to generalize beyond the parameters of this case. This study focused on teachers in one school district. A variety of factors could influence the perceptions of these teachers that could be better controlled with a larger representative sample in a quantitative study. The district in this study

may not sufficiently represent the broader population of language arts teachers, thus it would be inappropriate to attempt to generalize their perceptions.

One of the study's limitations was the timing. This study was limited to teachers' reflections after the first year of the assessment development process. The development and use of systematic criterion referenced assessments to measure student achievement for state reporting purposes was a new concept for all districts. An additional limitation was that one of the researchers is also the district's Curriculum Director. Although the study's design took additional measures to ensure that teachers' responses were kept confidential, some of the responses may be tempered by the fact that the research was conducted by one of the individuals responsible for the assessment development process.

RESULTS

All members (n=26) of the Reading CRA Team responded to the questionnaire during the meeting. All three researchers agreed on the four major themes that emerged from the data. These themes included: 1) evaluation process concerns, 2) resources, 3) instructional time challenges, and 4) professional growth. Each of these themes will be discussed separately.

Evaluation Process Concerns

The most apparent concern of teachers was their suspicion about how the political process would impact the accountability movement. As mentioned above, as part of the Nebraska state assessment and accountability system, districts receive a holistic rating on the technical quality of their district assessment system. The most common premise revolved around how the state department of education and the state legislature would use the data to potentially harm their perceptions of education. One participant indicated that reporting their district results to the state made him or her, "Nervous that the 'higher-ups' are using us as a political pawn." Another respondent reinforced this belief by writing that the governing bodies of the state were using this process as, "Finger-pointing. Political grandstanding." The

respondent continued with, "Assessments aren't here because teachers or parents or students asked for them. This is driven by politics which seeks to place blame on real and mostly imagined problems." There was also alarm that state politics were infringing upon the traditional role of local school districts in the state. "I sense an attitude of mistrust between local districts and 'the state' regarding the quality of teaching," was one respondent's opinion.

Teachers were also emphatic in their disdain for comparing school districts to one another. Respondents reinforced their concern that rampant comparisons of school districts would occur because of the assessment results. Participants did not make special mention of media, parents or students comparing scores; their concern rested at the state level. One respondent perceived a possible adversarial outcome in comparing one district against another district by stating, "I'm concerned that this information may be used to 'guide' school districts, teachers, etc. against each other – and if that happens we have defeated the purpose of these tests." Another participant went a step further when writing, "Regardless of intent, scores will be used to make comparisons." Indeed, the participants made their concerns clear that the state will use the assessment results to compare schools with one another. Although common on the national stage, Nebraska's lack of a state testing program has insulated schools from some of the comparative judgments that are made about schools in other states based on test scores.

Many participants had an apprehensive perspective toward the future of the standards-based accountability movement. Some teachers feared the prospect that someday a district's results on the state report card would have a financial impact on their district. "My guess is that...funding for education will be based on the results." Other respondents took a different view, showing uneasiness about the lifetime of the standard-based accountability movement. "Many new things don't make it, so I hope that doesn't happen." Finally, one respondent summarized the teachers' perspective by stating, "I'm concerned that the test becomes more important than the joy of learning." This statement represents a common sentiment expressed among

educators (even before NCLB); however, it also demonstrates a belief that measuring student performance is not connected to the learning environment. This observation was troubling.

Resources

Although the Nebraska state assessment and accountability system gives individual school districts the latitude to select or create assessments for each of the state standards at benchmark grade levels in elementary, middle, and high school, teachers were neither educated nor experienced in systematically developing assessments. Because the state department of education did not create a prescriptive model or offer other adaptable examples for the development of these assessments, districts chose different paths to progress on the learning curve. The school district in this study used a regional educational service unit and local resources to create their assessments. The service unit relied heavily on Stiggins (1997) to guide the professional development and operational activities. According to respondents, the local district's lack of resources led to additional frustrations for teachers.

An initial hurdle for teachers was obtaining the necessary level of aptitude with technology, specifically word processing and spreadsheet applications. Teachers felt that opportunities for technology training were absent during the year. One respondent wrote, "Teachers needed to be well-trained in the use of technology if they are going to use it effectively." Another source of contention for teachers was the difficulty in finding compatibility among software programs used in various buildings throughout the district, "Technology was the most difficult part of the process. I lacked the knowledge needed to create spreadsheets, etc. Computers, when we were working, did not coincide with one another and with the many problems, we lost material and spent many hours re-doing it." Another aspect of the technological requirements related to the assistance they received from district personnel. One participant wrote, "On the days when there wasn't any help, we accomplished very little." However, not all teachers struggled with technology. Some were generally pleased with the technology available to teachers. One respondent

wrote, "The district already had the technology necessary to assist in the process." Obviously, technology was one resource that caused conflict for some teachers.

Another focal point for the teachers was the training necessary to write good assessments. One teacher recommended, "[I] needed more direction in how to develop assessments. As a teacher who has to develop tests for English 11 students,² I received no prior training about how to develop criterion referenced assessment tests. This meant a lot of hit and miss work before the actual tests were completed." Other teachers wondered about the quality of their assessments prior to the actual administration, "I would have liked a critiquing process to review the assessments before I gave them to students." Another respondent reflected on the perceived lack of guidelines by simply stating, "We had to start from scratch." Although teachers struggled with the development process, they were complimentary of the regional educational service units. One participant wrote, "We had a group of people from the service unit that has helped us tremendously." Respondents' reflections upon their experience suggest the need for additional knowledge prior to development of assessments.

The largest source of frustration for teachers was their lack of time to develop, administer, and internally evaluate the assessments. All teachers mentioned the challenge of finishing the many tasks associated with developing assessments in a timely fashion. As one teacher stated, "You will give more hours to developing tests and then after you pilot you will have more revisions to make to the test." Other respondents reiterated this claim, but pointed to a specific time period, "More time needs to be provided in the development stage for teachers." Another educator coupled teaching English with that of developing her assessments when stating, "Teaching English is simply very time-consuming. So is writing criterion referenced assessments."

Teachers were adamant in their need for additional time to develop, implement, and revise assessments. The challenges of working with teachers who struggle with technology and the education necessary to write good assessments are a concern for teachers, but the issue of time appears to reign supreme. As one respondent wrote,

“Never forget – this process takes a toll on all involved.”

Instructional Time Challenges

A third theme that emerged from the teachers’ responses was the perception that instructional time needed to be reduced to allow for the addition of further testing. Teachers approached this theme from three different perspectives: struggles in the past year, the rearrangement of their teaching objectives, and worries regarding the future.

Respondents displayed contempt for the amount of class time devoted to testing in the academic year. One participant stated, “I’ve had to cut several things to fit the testing in.” Another participant reiterated the argument by writing, “We ‘stopped’ more frequently to do these formal assessments.” Teachers chose to sacrifice learning activities used in the past to meet the requirements of the district’s assessments. One teacher wrote, “Teachers gave students independent work to free teachers to test. This led to lots of lost instruction/learning time.” This also suggests a perception that learning could not occur unless the teacher was leading it. Students also noticed the change in teaching and learning, according to some respondents. “Students started to get tired of taking these tests and verbally complained,” wrote one person.

Teachers shuffled their prior schedule of instruction to massage the assessments into the curriculum. The result was a modification of the scope and sequence of the curriculum. “During the first semester the amount of material was reduced considerably because of the criterion referenced assessment process,” wrote one teacher. Another respondent wrote more explicitly about the lessons that were lost to the accountability requirements. “We had to use a lot of Accelerated Reader [individual computerized reading testing] time to do the criterion referenced assessments tasks. This is compromising one of the goals that standards are trying to measure—reading level and comprehension.” Others focused displeasure with the perceived interference criterion referenced assessments created in teaching and learning. “In order to meet the test dates I was continually adjusting my goals and lesson plans. At times the test restraints affected the flow of teaching and disrupted the continuity of lessons.” Clearly,

teachers demonstrated contempt for the interruptions caused by the criterion referenced assessments.

Teachers viewed the future of teaching and learning in their classroom with a cynical undertone. “I’m concerned [with] what’s going to happen as we add the other criterion referenced assessments (mathematics, science, social studies),” one individual wrote. “This will mean lots of assessing, and too much time taken away from learning.” The writer continued, “This is going against everything we’ve spent the last couple of years trying to build: multiple intelligences, integration, modes of learning, and brain research. The standards are starting to drive everything we do.” Other writers’ opinions resonated with the same belief that the accountability movement is leading to the concession of additional instruction time in the future. “We are trying to teach stuff too quickly, instead of teaching a few things to a high degree of mastery,” wrote a respondent. Clearly, teachers viewed the incorporation of district assessments as another mandate that compromises the time they spend educating students.

Professional Growth

Teachers wrote candidly about their concerns with the accountability movement and their suspicion of how the state would use their local information. Intermixed with their concerns about the state, the lack of resources provided to them, and the loss of instructional time, was a sense of fulfillment in completing an enormous task. Respondents wrote proudly about their increased understanding of curriculum, instruction, and assessment. The rich professional dialogue and collegiality the teachers experienced through the challenges of developing assessments were unifying factors.

Teachers became more knowledgeable about curriculum, instruction and assessment in the past year. One individual focused on newfound curricular knowledge by stating, “I feel I have a better understanding of what our grade level curriculum is.” Another respondent concentrated on a positive twist that entered their classroom, “Now there is a new dimension to the class that holds students and teachers more accountable for information and test taking skills that had not been

present.” A third person mentioned that the process made them, “...more aware of my curriculum—always thinking about aligning to the standards.” The respondent turned to assessment when continuing, “It made me more aware of how effective my lessons were—did the test [scores] plot out as I figured or was there more teaching needed?” Using the assessment results to impact instruction grabbed another teacher’s attention when she wrote, “I realized kids didn’t know things I thought they should have. I also have implemented a few more structured activities to do this.”

In addition to curriculum, instruction and assessment changes, teachers wrote about the rich professional dialogue that transpired throughout the year. One teacher noticed an atmosphere of teamwork throughout the assessment development process when writing, “I feel we have a good foundation of skills developed to help students achieve.” Other respondents gained insight from conversations with colleagues. One educator wrote of an experience during the assessment development process that, “...enabled me to talk with another district teacher about how they taught the curriculum.” Finally, a connection between curriculum, instruction and assessment and professional dialogue occurred when one participant stated, “Many different styles of teaching have come up from discussing and brainstorming with different teachers.” Although the road was long and winding, teachers appreciated the opportunity to professionally discuss educational issues with their colleagues.

Some teachers felt this long process created a paradigm shift in their outlook of assessment. One educator wrote, “I learned that less is more, and that focusing on a few important skills, practicing those skills until students learn them, and then testing students worked.” A second participant reinforced this claim by stating that assessments, “...don’t have to be long and hard to write in order to assess.” Another professional growth opportunity occurred that supported good teaching practices. One respondent wrote, “These test results told me what I knew in my ‘gut’ and reinforced that belief.” The respondent added, “I do believe I have grown as a professional.” However, most respondents wrote with tempered enthusiasm

for assessments, though. Instead, teachers focused on the utter importance of teaching and learning. As one teacher wrote, “...student assessment is crucial, however, I do not think it should be the focal point of education. Learning should be.” Teachers in this district underwent a strenuous process, but one which facilitated a deeper understanding of being a professional educator.

DISCUSSION

Language arts teachers across Nebraska underwent a tremendous change in the past three years. From aligning curriculum to the state content standards to developing, implementing, and administering assessments to measure those standards, the role of language arts teachers in the state changed.

Noteworthy across the themes was the general feeling of frustration that the teachers expressed about the process in which they engaged. Many of the comments elicited and describe above suggest that the staff and assessment development processes could have been handled more efficiently. Although much of the responsibility for this trauma was ascribed to the state’s department of education, local administrators were also included in their displeasure because they were responsible for the local decisions that guided them through the process. There was a perception among the teachers the rules regarding the assessment and accountability process changed midstream. Teachers were glad that this phase of the ordeal was over and did not want to go through it again. It will be interesting to learn if these teachers’ attitudes on the process are similar to those who experience initial change/learning processes in general or if they believe that they will not periodically revisit their curriculum, instruction, and assessment strategies because they are “done”. Politically, the teachers were fearful that this first round of state assessment and accountability would lead to further mandates and higher expectations for teachers.

These teachers also felt overburdened by their professional lives. They struggled to meet their personal expectations in the classroom because of the immense commitment revolving

around state standards and accountability. Regarding the instructional challenge theme, teachers felt like they forfeited their philosophy of education in order to jump through hoops created by the state department of education. Furthermore, these language arts teachers believed that the integration of the district's criterion referenced assessments created unnecessary upheaval in their classrooms. The same scope and sequence that many teachers followed for a number of years was upset by the many assessments that now have to be infused into instruction, and teachers did not like this change. However, the process forced them to critically examine their curriculum and challenged past practices that may have become stale.

Within the resources theme, there were greater concerns expressed by elementary language arts teachers as opposed to the middle level teachers and high school teachers. The elementary teachers' comments suggested a greater struggle with technology. In further exploration of this difference, the researchers discovered that within this district, the middle and high school levels received recent upgrades in technology that had not yet extended to the elementary level. An additional explanation is that for the elementary level, the assessment development process appears to be continuous because they will also be developing the assessments for mathematics, science, and social studies in subsequent years. At the middle and high school levels, teachers tend to specialize within their content area.

There were positive elements within the professional development theme as teachers expressed a better understanding of the connection between assessment and instruction. As a result of the development and subsequent implementation of district assessments, teachers in the district also learned much about merging accountability and school improvement. In a state model that relies on locally selected and/or developed assessments, the state requirements become the district (or perhaps even classroom) requirements, allowing teachers more immediate access to information for the improvement of instruction. Under this scenario, state assessments, district assessments, or classroom assessments should not be mutually exclusive; assessment is then based on individual student learn-

ing and proximate to instruction. The possible dual use of district assessments for targeting both improved instruction and accountability reporting also has the potential to decrease additional, redundant layers of testing.

CONCLUSIONS

Although the purpose of this study was not to generalize the results beyond this district, other school districts, or perhaps states, may benefit from the perceptions that teachers in this district expressed. A critical component of any new system is professional development opportunities. In education we are typically very concerned about providing students adequate time and opportunities to learn prior to being held responsible for that learning. State and federal policymakers should extend that concern to reform efforts that require meaningful professional development. Many of the comments made by the teachers in this district demonstrated a limited understanding of assessment literacy. Until these skills become part of pre-service educational requirements, we cannot expect to train teachers once they are in the classroom. There is likely a perception that if the skills were important, they would have been in their undergraduate (or sometime graduate programs). Once teachers have opportunities to develop basic assessment literacy skills, it is appropriate to push them to the next level of understanding to help them see the connections among the curricular, instructional, and assessment activities they are responsible for in their classrooms. This is a non-trivial goal.

Given the uncharted waters into which districts are sailing with Nebraska's state assessment and accountability system, a follow-up study of the teachers that develop and implement the mathematics assessments in the current school year would provide comparative data about perceptions of the assessment process. It would also be appropriate to examine these same teachers' perceptions after additional opportunities to develop their skills.

Because this study focused on a single district, other researchers are encouraged to examine these questions in their own districts or states. This research is also needed for districts in other states

that have considered varying levels of locally developed assessments as part of the state assessment and accountability system (e.g., Maine, Vermont) to discover whether teachers' experiences are comparable. It is unknown whether teachers from other states would have similar perceptions about developing assessments because this assessment and accountability model is still unique nationally. The use of locally developed assessments to inform state accountability is just beginning and the body of knowledge must grow if more states undertake the process.

¹ The authors would like to acknowledge Amy Roberts and the three anonymous reviewers who provided excellent feedback on an earlier version of this article.

² Students enrolled in the 11th grade English course in the district.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Black, P. & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-73.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 9-21.
- Buckendahl, C. W., Plake, B. S., & Impara, J. C. (2004). A strategy for evaluating district developed assessments for state accountability. *Educational Measurement: Issues and Practice*, 23(2), 17-25.
- Buckendahl, C. W., Impara, J. C. & Plake, B. S. (2002). District accountability without a state assessment: A proposed model. *Educational Measurement: Issues and Practice*, 21(4), 6-16.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: SAGE Publications, Inc.
- Diaz, M. E. (2001). Will reform based on standards and assessment make a difference in the 21st century? *Mid-Western Educational Researcher*, 14(1), 22-27.
- Lane, S., Parke, C. S. & Stone, C. A. (2003). The impact of a state performance-based assessment and accountability program on mathematics instruction and student learning: Evidence from survey data and school performance. *Educational Assessment*, 8(4), 279-315.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook (2nd ed.)*. Thousand Oaks, CA: SAGE Publications, Inc.
- Plake, B. S., Impara, J. C., & Buckendahl, C. W. (2004). Technical quality criteria for evaluating district assessment portfolios used in the Nebraska STARS. *Educational Measurement: Issues and Practice*, 23(2), 12-16.
- Popham, W. J. (1994). The instructional consequences of criterion-referenced clarity. *Educational Measurement: Issues and Practices*, 13(2), 15-18, 30.
- Popham, W. J. (1999). Where large-scale assessment is heading and why it shouldn't. *Educational Measurement: Issues and Practices*, 18(3), 13-17.
- Roschewski, P. (2004). History and background of Nebraska's school-based teacher led assessment and reporting system (STARS). *Educational Measurement: Issues and Practice*, 23(2), 9-11.
- Roschewski, P. with Chris Gallagher & Jody Isernhagen (2001). Nebraskans reach for the STARS. *Phi Delta Kappan*, 82(4), 611-615.
- Rose, L. C. & Gallup, A. M. (2004). The 36th Annual Phi Delta Kappa/Gallup Poll of the Public's Attitudes Toward the Public Schools. *Phi Delta Kappan*, 86(1), 41-58.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Stiggins, R. J. (1997). *Student-centered classroom assessment (2nd ed.)*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education pro-

grams. *Educational Measurement: Issues and Practices*, 18(1), 23-27.

Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practices*, 20(3), 5-15.

Stiggins, R. J. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 23-27.

Stone, C. A. & Lane, S. (2003). Consequences of a state accountability program: Examining the relationships between school performance gains and teacher, student, and school variables. *Applied Measurement in Education*, 16(1), 1-26.

Trevisan, M. S. (1999). Administrator certification requirements in student assessment competence. *Applied Measurement in Education*, 12(1), 1-11.

Toby Bossis a professional staff developer for Educational Service Unit #6 in Milford, Nebraska. His research interests include classroom assessment, professional development, and educational reform.

Daniel P. Endorf is the principal of York High School in York, Nebraska. His research interests include leadership, educational reform, and school improvement.

Chad W. Buckendahl is the director of the Buros Institute for Assessment Consultation and Outreach at the University of Nebraska, Lincoln. His research interests include applied measurement and computerized testing.

Appendix

Perceptions of Assessment Development Questionnaire

1. Which grade level team did you work with (elementary, middle, high school)?
2. Please estimate how many hours you spent developing your district's criterion referenced assessments this year? Of these hours, how many were compensated?
3. What resources would have been beneficial to you prior to the assessment development process?
4. What did you learn about your grade level curriculum through the process?
5. How was your classroom affected by this process?
6. How were you able to integrate criterion referenced assessments into classroom instruction?
7. How did your understanding of assessment change because of this process?
8. What assistance did you receive in the assessment development process?
9. How do you feel about the state's evaluation of your district's assessments that will occur later this summer?
10. Based on your experience, what would you recommend to other teachers who will be going through this process