12-2016

Teacher-Created Student Growth Assessments Used For Teacher Evaluation And The Effects On Teacher Efficacy

Perry A. Finch

TEACHER-CREATED STUDENT GROWTH ASSESSMENTS USED FOR TEACHER EVALUATION AND THE EFFECTS ON TEACHER EFFICACY

Perry A. Finch

Educational Leadership Doctoral Program

Submitted in partial fulfillment

of the requirement of

Doctor of Education

in the Foster G. McGaw Graduate School

National College of Education

National Louis University

December 2016

Document Origination Statement Digital Commons @ NLU

This document was created as *one* part of the three-part dissertation requirement of the National Louis University (NLU) Educational Leadership (EDL) Doctoral Program. The National Louis Educational Leadership EdD is a professional practice degree program (Shulman et al., 2006).

For the dissertation requirement, doctoral candidates are required to plan, research, and implement three major projects, one each year, within their school or district with a focus on professional practice. The three projects are:

- Program Evaluation
- Change Leadership Plan
- Policy Advocacy Document

For the **Program Evaluation** candidates are required to identify and evaluate a program or practice within their school or district. The "program" can be a current initiative; a grant project; a common practice; or a movement. Focused on utilization, the evaluation can be formative, summative, or developmental (Patton, 2008). The candidate must demonstrate how the evaluation directly relates to student learning.

In the **Change Leadership Plan** candidates develop a plan that considers organizational possibilities for renewal. The plan for organizational change may be at the building or district level. It must be related to an area in need of improvement, and have a clear target in mind. The candidate must be able to identify noticeable and feasible differences that should exist as a result of the change plan (Wagner et al., 2006).

In the **Policy Advocacy Document** candidates develop and advocate for a policy at the local, state or national level using reflective practice and research as a means for supporting and promoting reforms in education. Policy advocacy dissertations use critical theory to address moral and ethical issues of policy formation and administrative decision making (i.e., what ought to be). The purpose is to develop reflective, humane and social critics, moral leaders, and competent professionals, guided by a critical practical rational model (Browder, 1995).

Works Cited

Browder, L.H. (1995). An alternative to the doctoral dissertation: The policy advocacy concept and the policy document. *Journal of School Leadership, 5,* 40-69.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Shulman, L.S., Golde, C.M., Bueschel, A.C., & Garabedian, K.J. (2006). Reclaiming education's doctorates: A critique and a proposal. *Educational Researcher, 35*(3), 25-32.

Wagner, T., et al. (2006). *Change leadership: A practical guide to transforming our schools.* San Francisco: Jossey-Bass.

ABSTRACT

The landscape surrounding teacher evaluation is changing as teacher accountability becomes more prevalent in this process. This study looks at how teacher-created student growth assessments (SGAs) for teacher evaluation affects teacher efficacy with respect to instructional practice. Survey data were collected from 110 kindergarten through eighth-grade teachers; additional data were gathered from three interviews. A majority of these teachers reported the work in creating SGAs benefitted their instructional practice. Teachers reported increased knowledge in learning standards, curriculum development, and assessment literacy as a result of creating their own SGAs. Qualitative findings suggest that when teachers are provided an opportunity to collaborate, they learn from each other. Other findings suggest teachers may need professional development to increase their ability to collaborate efficiently and effectively.

PREFACE

As I finish this program evaluation of my district's student growth assessments, I realize I have affirmed some beliefs and formed some new thoughts about the world of education. In my current position as principal of Prairie Middle School (pseudonyms are used in this document), I find myself in a position to act on some of these affirmations and new thoughts with the hope of making our school a more productive place where student learning is increased. My journey in education started as a special education teacher and later dean at an inner-city Chicago high school prior to landing in District 32 where I worked as a middle school dean, then assistant principal, and later a principal. All of these experiences have helped form my perspectives about education and how I can do my part to try to improve it.

One of my primary understandings that developed from completing this research is that a vast majority of teachers find value in conversations with their colleagues. Though teachers respect and value such conversations, a major misassumption has been made regarding educators. Teachers make their living talking to and with students; however, this does not necessarily mean they know how to talk with *each other*. While so many of us in education crunch numbers in daily schedules to find time for collaboration or debate the value of collaboration during teacher contract negotiations, it appears the cart is being put before the horse for many in the field. Even with growth mindset-educators who see the value in collaboration as a way of broadening their abilities and effectiveness with students, teachers frequently admit they are ill prepared to problem-solve collectively, give constructive criticism, or receive critical feedback. This

disconnect between what teachers find an effective source of professional development and their ability to take part in the activity is something educators must address.

Another finding that struck me as I completed this research involved the manner in which we support adult learning for teachers. So much time and effort is rightfully spent discussing effective methods for teaching and assessing students. Clear learning targets and success criteria are common best practices when teachers instruct students. Exemplars, rubrics, and explicit formative feedback are other ways that good teachers get desired results from their students. Why then should these practices of quality instruction that have proven to support child-learners not be used with adults? A major takeaway from my research is that when adults are not afforded the necessary scaffolds to support new learning such as collaboratively creating and scoring their own student growth assessments, frustration levels increase while efficiency and productivity decrease.

I also learned that when traditional approaches to instruction shift to more progressive instructional practices, these challenges and new methods could affect a school or district's culture. Having teachers create and score their own student growth assessments produced new expectations for teachers. Teachers were now expected to collaboratively create common assessments based on the Common Core State Standards. Once teachers became familiar with creating standards-based assessments and then formulated standards-based instruction, an organic progression occurred, as teachers desired standards-based grading of their students. Teachers in our district realize that just as instruction is best when based on standards, so too is assessment. This research revealed that standards-based grading could be the next area of change in my district. This is an area I am interested in investigating and possibly researching in the future.

TABLE OF CONTENTS

LIST OF TABLES

SECTION ONE: INTRODUCTION

Purpose

There has been much discourse and deliberation regarding the scope and methodology of teacher evaluation. The state of Illinois has determined that demonstrated student growth is a required component of teachers' evaluations in upcoming years. According to Section 50.110 Student Growth Component in the Illinois Administrative Code (2014), "Each school district, when applicable (see Section 50.20 of this Part), shall provide for the use in the performance evaluation plan of data and indicators on student growth as a significant factor in rating teacher performance" (Illinois State Board of Education, 2014). While many districts may rely on standardized tests such as Measure of Academic Performance (MAP) to gauge their student growth for teacher evaluation, Broadridge Elementary School District 32 has chosen to have its teachers create their own teacher evaluation assessments. According to the Broadridge School District 32 Assessment Overview and Administrative Manual (2014):

> District 32 believes that to employ a standardized test to measure student growth as part of a teacher's evaluation is misaligned with the curriculum work that has been conducted over the last three years. For example, units of study created around standards that follow the integration guidelines of the PARCC Model Content Frameworks engage students in authentic performances that a simple, standardized test cannot effectively measure. The purpose of standardized tests is not to inform instruction at the classroom level, which is why we have embraced teacher-made assessments whose purpose *is* to inform classroom level decisions.

> They are the most informative measure to ensure students' progress on the college and career readiness trajectory. (p. 4)

The purpose of this study is to examine the effects that the creation of these assessments had on teachers' sense of efficacy with respect to their classroom instruction.

Results from standardized tests such as MAP are norm-referenced, not criterion-referenced, and may not be aligned to daily curriculum and instruction that students experience. Chappuis and Chappuis (2006) defined criterion-referenced tests to be assessments used to measure student progress toward specific curriculum goals or standards that produce scores that are compared to a pre-defined acceptable level. Norm-referenced tests measure one student's performance against another student's performance or other students of the same age or grade (pp. 150-151). Leaders in the educational community have found problems with the use of norm-referenced assessments for teacher evaluation. Stiggins (2013) stated, "At best, such tests are designed to support inferences about the *probability* that a student has mastered the domain sampled; such tests will not support inferences about student mastery of any specific learning targets within that domain" (p. 22). Though MAP and other standardized assessments may seem precise and effectual from a distance, a closer look at such assessments yield another perspective. Darling-Hammond (2013) stated:

> The promise of efficiency and objectivity are seductive, but they are not a sufficient rationale if it turns out that the measures are not accurate or their use undermines the main objectives of teacher evaluation: developing and retaining excellent teachers and continually improving teaching and learning. (pp. 73-74)

As stated, the data from MAP assessments do not yield individual student growth results. Norm-referenced assessment data provides a *comparison* growth component based on the growth of the norm-reference group. The assessments created by teachers for the purpose of teacher evaluation in District 32 do provide a direct measure of student growth and are aligned to the daily curriculum and instruction that takes place in classrooms. Beyond measuring the growth of students in District 32, I believe these assessments may be a way of increasing growth in teachers as well.

While simultaneously meeting the student growth requirements of the Performance Evaluation Review Act (PERA), collaborative design of common student growth assessments provide a professional development opportunity for participating teachers. This process allows for collaboration and sound assessment practice that include the collection of student baseline data, followed by targeted instruction based on this data. There is a valuable unintended consequence of the PERA that goes beyond measuring student growth for evaluating teachers. This process allows for teacher-learning to occur. Having teachers complete this cycle yearly may have a positive impact on their efficacy as educators. If the results of these findings suggest that teachers' effectiveness in the classroom increase due to creating their own student growth assessments (SGAs), then this program is something to be replicated and shared with others in the field of education.

It was necessary to gain insights from teachers who created SGAs in order to ensure the effectiveness of this program's impact on instruction. Teacher involvement in this evaluation was crucial as they hold valuable insights that can help shape the future direction of how SGAs are designed and implemented in District 32. Involving teachers

in this process broadened their horizons, validated their opinions, and reaffirmed their commitment to making this program a success. According to Patton (2008), "To evaluate something means determining its merit, worth, value, or significance" (p. 5). The purpose of this research is to determine the value of this program and its impact on teacher efficacy.

Many districts see the MAP assessments as the simplest way to measure student growth. MAP may be the easiest in terms of efficiency in providing some degree of student data, but it may not connect to what all teachers do with their instruction on a daily basis. An alternative option for meeting the demands of the new teacher evaluation law in Illinois includes allowing teachers to create their own common SGAs. By doing so, an opportunity for teacher growth and increased efficacy may result.

Rationale

PERA Law

In 2010, Illinois governor Pat Quinn mandated a new era of evaluation for teachers across the state. The Performance Evaluation Reform Act (PERA) became law, requiring teachers to demonstrate academic growth in the students they teach. The primary goal of all PERA student growth assessments (SGAs) is to provide administrators evidential data regarding teachers' ability to facilitate student growth through their instruction. The state of Illinois has provided a model of how districts might fulfill this mandate through a complex framework that assigns assessments to different "types." A Type 1 assessment includes standardized tests such as the Illinois Standard Achievement Test (ISAT), which was replaced by Partnership for Assessment of Readiness for College and Careers (PARCC) in 2014, or the Measure of Academic

Progress (MAP). A Type 2 assessment includes common assessments created across districts by departments or grade levels administered as a local common assessment. A Type 3 assessment is an assessment created by an individual teacher to be administered by that individual teacher to their students. The state of Illinois determined districts would have the option to choose either a Type 1 or Type 2 assessment and that all teachers would need to administer Type 3 assessments (Illinois Board of Education & Illinois Performance Evaluation Advisory Council, 2011).

Danielson Framework for Teaching

The process described above expects teachers to delve deeply into the professional practices espoused by Charlotte Danielson (2013) in her widely accepted publication, *The Framework for Teaching Evaluation Instrument.* According to Danielson, "The Framework for Teaching identifies those aspects of a teacher's responsibilities that have been documented through empirical studies and theoretical research as promoting improved student learning" (p. iv). Teachers in District 32 have been using the Danielson framework as a basis for their evaluation for years. The implementation of teacher-created SGAs has provided a platform for deeper understandings and greater opportunities to bring to life some of the concepts embedded in the Danielson framework.

This teacher-created SGA program has the potential to go beyond its intended purpose of evaluating teachers. As an administrator in this district, I have seen the teacher-created SGAs change the conversations among teachers. They now speak about instruction using terminology directly pulled from the Danielson framework. As a result, their efforts to create SGAs have increased discussions regarding individual learning

needs of students, which are so vital for increased student achievement. Beyond providing feedback regarding teacher performance, these teacher-created assessments introduced the use of formative assessment into teachers' instruction, potentially increasing their effectiveness with students. The use of formative assessment was not a new practice for some District 32 teachers. However, teachers creating their own SGAs has increased the use of formative assessment, thus bringing the hope of uniformity with these best practices to our entire teaching staff.

A final reason I find this program valuable and worthy of evaluation involves the implementation of multiple layers of inter-rater reliability for scoring these assessments. The teachers score their own assessments, then grade-level colleagues score samples as a group to ensure consistency. Finally, a district-wide organization known as the Inter-rater Reliability Team scores sample assessments to increase scoring consistency. This district-wide team approach provides not only corroboration about scoring but also feedback about the validity of the assessments and their alignment to the Common Core State Standards (CCSS). If this process continues to prove valuable to District 32 teachers and increases student learning, it could be a model that other districts may wish to adopt.

Goals

An evaluation of District 32's teacher-created SGAs is necessary in order to determine if my beliefs and perceptions about increased teacher efficacy are in fact valid. Through surveys and group interviews with District 32 teachers, the primary goal of the study is to provide evidence that either supports or refutes the connection between increased teacher efficacy and teachers collaboratively creating SGAs as a component of teacher evaluation. It may turn out that this process is beneficial for some teachers and

not others. If this were to be the case, another goal of this evaluation would be to attempt to identify exactly how this process benefits teachers with the hope of replicating the positive experience with more teachers.

The creation of SGAs by teachers also presents an opportunity for teachers to examine the CCSS that they teach and assess. This forms the background for the third goal: teachers must construct their assessments using the CCSS providing a greater understanding of these standards. If a teacher creating his or her own SGAs is an organic form of professional development, a reasonable outcome could be increased student-learning. Teachers creating their own SGAs, going through the baseline data collecting process to inform instruction requires regular use of good assessment practices. The mandate of PERA can be met by having teachers involved in creating the assessments, which in turn raise the opportunity for increased student learning.

Research Questions

Primary Question

The primary question of this evaluation is: What is the effect of teacher-created SGAs on teacher efficacy in the classroom?

Secondary Questions

Secondary questions to evaluate the teacher-created SGA program are as follows:

- Does the act of creating SGAs increase student learning?
- Does the act of creating SGAs increase teachers' understanding of the value of common assessments?
- Does the act of creating SGAs as teams increase positive collaboration amongst teachers?

- Does the act of creating SGAs as teams increase reflectiveness regarding their professional practice?

SECTION TWO: LITERATURE REVIEW

Introduction

Including student growth in teacher evaluation is not a new idea. States such as Illinois have made student growth a mandated part of teacher evaluation. There is research that examines specific components related to the creation of SGAs used in teacher evaluation and their effects on teacher efficacy. There is literature to be reviewed regarding the following topics: teacher efficacy, teacher collaboration, creation and use of common assessments by teachers, and a professional learning community called Critical Friends Groups.

Teacher Efficacy

Many scholars have focused their work on teacher efficacy, including Bandura (1993); Moolenarr, Sleegers, and Daly (2011); Calik, Sezgin, Kavgaci, and Kilinc (2012); Powell-Moman (2011); and Darling-Hammond and Richardson (2009). Bandura's work involved defining self-efficacy with a focus on how it impacts student achievement. The work of Moolenarr et al. and Calik et al. focused on collective and self-efficacy while Powell-Moman et al. focused on how pairing scientists and teachers affects teacher efficacy. Darling-Hammond and Richardson (2009) explored the relationship of professional development to teacher efficacy.

Bandura (1993) was highly referenced in all of the recent research related to teacher efficacy. This work is pertinent in education because it addressed how individuals with a strong sense of efficacy manage problematic circumstances as challenges, set and maintain challenging goals for themself, sustain focus in their work, increase their efforts when they fail, and connect failure to lack of effort or understanding of acquirable skills

(Bandura, p. 144). These attitudes not only allow good teachers to persevere and succeed, they also provide a positive model for students to develop into confident successful learners. Bandura espoused the importance of a teacher's sense of efficacy and the direct connection to student achievement. According to Bandura, "Teachers who believe strongly in their instructional efficacy create mastery experiences for their students. Those beset by self-doubts construct classroom environments that are likely to undermine students' sense of efficacy and cognitive development" (p. 140). Teachers with low efficacy negatively impact the efficacy of their students, instilling increased feelings of doubt when it comes to academic successes (Bandura, p. 144). A teacher's sense of confidence in teaching ability can impact student achievement.

In Moolenarr et al. (2011), the research involved 53 Dutch elementary schools. The researchers hypothesized that allowing teachers to collaborate, set shared goals, and increase social interactions with other teachers increased their sense of efficacy, which in turn increased student achievement (Moolenarr et al., p. 258). The unique approach to this research involved the collaboration amongst teachers occurring virtually. The researchers believed a correlation could be made between the collective efficacy of teachers and their effectiveness with students. The study discerned that teachers using social networks allowed for sharing of advice, instructional and content support, and social support among other things (Moolenarr et al., p. 252). Only quantitative data were collected and analyzed in this study. The researchers stated this was done in order to address the large sample size needed to establish general trends in the 54 schools involved. Moolenarr et al. reported that, "Findings suggested that the density of work related [to] personal advice networks affected teachers' perceptions of collective efficacy,

which in turn was associated with increased student achievement" (p.258). This research also contended that, "Collective efficacy served as an intervening variable that may explain how dense social networks among educators may ultimately benefit student achievement" (Moolenarr et al., p. 258). This research demonstrated that teacher communication and collaboration online could lead to increased teacher efficacy and thus greater student learning.

A study conducted by Calik et al. (2012) examined the effects of school administrators' instructional leadership on teachers' self-efficacy and collective efficacy. According to Goddard (2001), "Collective teacher efficacy is defined as 'teachers' perceptions that their effort, as a group, can have a positive impact on students'" (as cited in Calik et al., p. 2499). Researchers explored the effects of instructional leadership and teachers' self-efficacy and how self-efficacy subsequently affected the collective efficacy of a teaching staff. Quantitative data were collected from 328 primary classroom and branch[1] teachers in schools in Ankara, Turkey. The results of the study demonstrated that instructional leadership had a positive impact on teachers' self-efficacy and thus an indirect, strong, and positive impact on teachers' collective efficacy. "In other words, teachers' self-efficacy plays a mediator role between instructional leadership and collective efficacy" (Calik et al., p. 2500).

Powell-Moman et al. (2011) conducted a study of a specific relationship offered by the Kenan Fellows Program in North Carolina[2]. The purpose of the program was to

[1] Demir (2013) described branch teachers as teachers of subjects such as "social studies, science and technology, foreign language, physical education" (p. 74).

[2] Kenan Fellows Program Mission: To improve kindergarten through grade 12 STEM education by providing relevant, professional learning and leadership development for exceptional teachers through innovative collaborations with research partners in industry, higher education, and government (Powell-Moman et al., 2011, p. 49).

increase the retention of STEM (science, technology, engineering, and math) teachers through an improved curriculum that heightened student engagement. The professional relationship called "scientist-teacher partnerships" and the way in which such relationships increased teachers' sense of efficacy was the focus of this study. Twenty-three Kenan fellows took pre- and post-surveys yielding a positive correlation between their work with their partner-scientist and their sense of efficacy. "After participation in the program, fellows indicate increases in their efficacy for inquiry-based teaching and greater focus on the depth of content than on the coverage of all objectives" (Powell-Moman et al., p. 52). The Kenan fellows reported a greater sense of efficacy with their instructional delivery and their overall knowledge of the content in which they taught. The scientist-teacher relationship provided limited increases in teachers' knowledge of all science curricular expectations.

Darling-Hammond and Richardson (2009) addressed the issue of quality professional development. It is imperative that teacher professional development changes as the new demands on teachers emerge. Providing instruction to meet twenty-first century learning challenges requires a shift in professional development involving embedding teacher learning opportunities into the everyday work of teachers (Darling-Hammond & Richardson, pp. 46-48). According to Darling-Hammond and Richardson, "Hands-on work that enhanced teachers' knowledge of the content and how to teach it produced a sense of efficacy—especially when that content was aligned with local curriculum and policies" (p. 47). A special emphasis was placed on the integration of professional development and the collective school improvement process. To ensure consistency in teacher professional development and daily work, it is important to

smoothly connect standards, assessment, and curriculum to professional learning (Darling-Hammond & Richardson, p. 48).

Teacher Collaboration

There is considerable research available to review in the area of teacher collaboration and how it positively affects teacher efficacy. The value of teacher collaboration and its impact on student learning was addressed by Williams (2010), Adie (2010), Hargreaves and Fullan (2012), Darling-Hammond (2013), and Ross and Bruce (2012). Williams (2010) focused on the benefits of teacher collaboration and teacher efficacy in a high school setting. Adie (2010) explored how virtual collaboration online benefits teachers' efficacy, and Hargreaves and Fullan (2012) claimed collaboration provides teachers with greater confidence and opportunity for growth. Through the lens of teacher evaluation, Darling-Hammond (2013) explored the value of teacher and evaluator collaboration through the review of student artifacts. Ross and Bruce (2012) investigated how teachers' efficacy was affected by collaborative action research.

According to Williams (2010), there is a preponderance of evidence to support the value of collaboration as a means of increasing teacher efficacy. Schmoker established in DuFour, Eaker, and DuFour (2005) that, "Teachers do not learn best from outside experts or by attending conferences or implementing 'programs' installed by outsiders. Teachers learn best from other teachers, in settings where they literally teach each other the art of teaching" (Williams, 2010, p. 38). According to Williams, "Research consistently points to collaboration as a model of professional development that substantially impacts instructional practice and improves student achievement outcomes" (pp. 11-12). In her research, Williams conducted a case study of 20 secondary teachers within the same

school that previously had very limited teacher collaboration. This study purported that collaboration is highly valuable to newer teachers as it provides opportunities for less experienced teachers to share ideas with others. More seasoned teachers also found collaboration beneficial as it eliminated frequent feelings of dissatisfaction and seclusion (p. 18).

Adie (2012) conducted research with 24 teachers in Australia who participated in online meetings to discuss consensus on assessment standards. Data collection consisted of six recorded online conversations, 13 pre- and post-interviews, 12 surveys, and hundreds of emails between participants. Adie stated the purpose of the study to be:

> An ongoing research project that is investigating the formation of a common understanding of defined standards when teachers meet to moderate student work within a synchronous online environment. The paper is focusing on the process of online social moderation, and the factors that may support or hinder teachers in their judgment role. (p. 14)

Results of this research demonstrated support for online collaboration and its ability to increase teachers' sense of efficacy. According to Adie, "The confidence gained by teachers as a result of their judgment being agreed upon by someone outside of the school or cluster is significant" (p. 24). This research demonstrated that teachers value the validation of their ideas even through online forms of communication.

Hargreaves and Fullan (2012) wrote extensively about the concept of "professional capital." In their book they explained that professional capital was made up of human, social, and decisional capital. Human capital involved the talent and skills of individuals in an organization. Social capital was the collective knowledge and skills of

an organization and its ability to collaborate and use knowledge as a collective (Hargreaves & Fullan, p. 3). Decisional capital entailed the ability to make effective decisions in the moment of need (Hargreaves & Fullan, p. 5). When human, social, and decisional capital is actively apparent in an organization, that group is said to possess "professional capital." Hargreaves and Fullan concluded, "Support from and communication with colleagues led teachers to have greater confidence and certainty about what they were trying to achieve and the best ways to achieve it" (p. 111). They continued:

> Collaborative cultures build social capital and therefore also *professional capital* in a school community. They accumulate and circulate knowledge and ideas, as well as assistance and support, that help teachers become more effective, increase their confidence, and encourage them to be more open to and actively engaged in improvement and change. (p. 114)

The increase in professional capital in a school or district allows for increased teacher efficacy that can lead to greater student learning.

Darling-Hammond (2013) examined what was needed in order to create quality teacher evaluation. This work explored the value of teachers and administrators collectively discussing and examining student work. For schools to be effective, knowledge-sharing needs to be created in schools to promote a system that generates on-going learning (p. 112). Darling-Hammond wrote:

> Teachers and principals need the time and guidance to develop a shared understanding of effective teaching, to examine artifacts of practice for evidence of learning, to explore one another's assumptions about how learning occurs and

what counts as evidence of learning, to promote reflection, and to learn how to provide effective feedback. (p. 111)

Darling-Hammond broadened the scope of collaboration to go beyond just teachers to also include administrators and evaluators. This fresh approach increases the boundaries of collaboration to include more stakeholders and increase the adult learning within a school.

Ross and Bruce (2012) conducted two studies in order to measure the effects that teachers' participation in collaborative action research had on teacher efficacy. The research was based on an action research initiative in an elementary teachers' union in Ontario, Canada. Teams of four to six kindergarten through eighth-grade teachers were formed and assigned an academic researcher. These teams participated in summer training in which they created an action research question, established validity of the question through research articles, developed a plan, enacted the plan including data gathering, reached conclusions, and presented findings (Ross & Bruce, pp. 538-539). A collaborative action research (CAR) model was followed providing an equal opportunity for researchers and teachers to collectively conduct the research (Ross & Bruce, p. 539). Data were gathered using pre- and post-surveys based on such research questions as, "What are the effects of CAR on teacher efficacy (beliefs about their professional ability)?" (Ross & Bruce, p. 543). Study one provided data to suggest that action research does increase a teachers' sense of efficacy in some capacities. Ross and Bruce found, "In study 1 we found that one dimension of teacher efficacy improved (confidence in engaging students in learning activities) while two others remained unchanged (confidence in classroom management and confidence in instructional strategies)" (p.

546). The arrangement of study two—research team structure and instruments—was the same as the previous study, with the exception being the action research was math-focused. (Ross & Bruce, pp. 546-547). "Study 2 confirmed Study 1 findings that there was a statistically significant pre-post- improvement on teacher attitudes toward educational research and on teacher efficacy" (Ross & Bruce, p. 551).

Common Assessments

The value of teachers creating common assessments has been the focus of articles and studies in the academic community. Stiggins and DuFour (2009), McTighe and Emberger (2006), and Psencik and Baldwin (2012) professed teachers' professional growth can stem from work on common assessments. Good (2012) and Darling-Hammond and Falk (2013) explored the value of teachers working in partnership on common assessments with a focus on scoring common assessments. Stiggins (2013) defined assessment literacy and how assessment literacy plays a part in teachers' and administrators' work with common assessments with respect to teacher evaluation.

According to Stiggins and DuFour (2009), assessments could be classified into three categories: classroom assessments, school-level assessments, and institutional-level assessments (p. 641). When describing school-level assessments, the authors highlighted how collaborative teacher teams create common assessments together for a variety of reasons. One of these reasons was "[common assessments] help each team member clarify strengths and weaknesses in his or her teaching and create a forum for teachers to learn from one another" (Stiggins & DuFour, p. 641). The scope of this learning from colleagues included many elements incorporated into the teaching profession. Stiggins and DuFour stated:

> To create a common assessment, team members must build shared knowledge of relevant state standards, district curriculum guides, state assessment frameworks, and the expectations of the teachers in the next course or grade level in order to classify the intended learning for students. (p. 644)

The work of Stiggins and DuFour reinforce that teachers creating common assessments is constructive work and professional development for teachers.

Good (2012) explored the value of common assessments at the postsecondary level. In an attempt to shore up assessment scoring consistency for student papers, this study made comparisons among different professors and their scores on common writing assessments. Approximately one-third of tenured teachers and those on track to be tenured teachers took part in an extensive 30-hour professional development training program that provided common elements to include in teaching writing. Common rubrics and an emphasis on five dimensions—focus, content, organization, style, and language conventions—helped add unity to the writing instruction for teachers in multiple disciplines. Initial data collected to test the instructors' inter-rater reliability levels proved to show inconsistency in scoring writing assessments. Adjustments were made to the professional development of teachers scoring common papers (Good, p. 26). Scoring consistency increased after these changes were made and data provided new direction for future adjustments in the professional development for teachers scoring writing samples. Good concluded, "Having the opportunity to share their ratings on student-produced writing assignments from colleagues' discipline-specific classes and discuss and defend that rating they provided in training created an opportunity for interdisciplinary growth and teamwork" (p. 29).

McTighe and Emberger (2006) discussed three ways in which teachers collaborated to create common assessments. The first component of this work was to collaboratively create assessments that were aligned to learning targets. The second step in this process involved having peers review the assessment and offer critical feedback. The final stage in this work was when the teachers collectively evaluated student work based on the tasks they created in collaboration. McTighe and Emberger deduced, "By designing performance assessments, educators enhance their understanding of content standards and of the evidence needed to show that students really understand the important ideas and processes contained in those standards" (p. 44). The work of McTighe and Emberger highlighted the value of common assessment creation as such assessments increase teachers' understanding of elements of teaching.

Stiggins (2013) stated that assessment literacy included demonstrating "clear purpose" with the assessment or a well-defined idea of how the assessment would be used and by whom. Assessment literacy also included "clear targets" for the learner and "good assessment design" that yielded accurate results. "Good communication" in assessment literacy involved sharing assessment results with stakeholders in a meaningful fashion that met the needs of all stakeholders (Stiggins, pp. 74-75). These components were necessary in order to create quality assessments and when done in a collaborative manner, this process created learning opportunities for the teachers and administrators when the assessment involved teacher evaluation. Stiggins wrote:

> Over and above the learning benefits of assessment literacy, the development of this capacity within a faculty can permit enhanced teacher/supervisor relationships, ongoing opportunities for faculty team collaboration, enhanced

> opportunities for ongoing teacher and administrator growth, and better overall instructional decision making at all levels. (p. 95)

Teachers collaborating on common assessments particularly that are used in teacher evaluation can provide learning opportunities for teachers as well as administrators.

According to Darling-Hammond and Falk (2013), teachers who design and score common assessments create a learning opportunity for themselves that better prepares them to teach the complex skills found in the CCSS and better prepares students for work in the twenty-first century (p.1). This work explored how educators from New York Public Schools, the Silicon Valley Math Collaborative, the Ohio Performance Assessment Pilot Project, and the Boston Arts Academy all found benefit when teachers were afforded time to collaboratively create and score common assessments. The consistent finding across all of these educational organizations reinforced the concept that outsourcing assessment development and scoring to an independent entity is a missed opportunity for teacher learning. Darling-Hammond and Falk declare, "Instead of having to rely on testing companies to judge the outcomes of students' work, teacher involvement in scoring places assessment back into the domain of teaching, where it can be readily accessed to inform and support learning" (p. 8). This insightful statement suggests that the best method to determine and inform instructional practices involves being a part of student misconception determination, which is assessment scoring. Beyond the benefits of professional development for teachers and increased awareness of students' understandings, teachers collaborating on common assessments produce an opportunity for tighter instructional focus and collegial relationships. Darling-Hammond and Falk stated that, "In addition, the scoring experience helps [teachers] develop a

shared understanding and common language about the essentials of their discipline, which develops a sense of professional community and can facilitate more coherent instruction across classrooms" (p. 9). All of these findings reiterate how teachers working together on common assessments may lead to increased teacher efficacy, which could lead to an increase in student learning.

Psencik and Baldwin (2012) explored the work of Douglas County Public schools in Douglasville, Georgia. The focus of this work involved teachers using common district assessments in order to ensure consistent learning experiences for students (Psencik & Baldwin, p. 30). Two of the assumptions of the district leadership team included the following: "Teacher-developed common assessments driven by the Georgia Performance Standards for all units are central to teaching teams' efforts to offer results-oriented instruction that meets students' needs," and "When teachers analyze student achievement trends in a variety of data and reflect on their own practice, they are better informed to set relevant professional learning goals for strengthening their performance" (Psencik & Baldwin, p. 31). To lead this new assessment initiative, a district assessment team consisting of teachers, coaches, building administrators, and district-level cabinet members was established. This team created the guidelines for the numerous grade-level teams that developed the common assessments. One fourth-grade team's work on common assessments increased team members' knowledge of lesson planning, student interventions, and their own professional development (Psencik & Baldwin, p. 31). Efficacy increases were noted in the fourth-grade team members in the areas of classroom management, lesson alignment to standards, and self-assessment as an educator (Psencik & Baldwin, p. 32).

Critical Friends Groups

Though the concept of Critical Friends Groups (CFG) is not a new idea, limited research can be found on this topic. Since 2000, National School Reform Faculty (NSRF) has coordinated all CFG trainings (NSRF, 2012, p. 8).[3] NSRF touts, "CFGs result in greater student learning and success by helping teachers and administrators intentionally develop and implement 'best practices'" (p. 9). CFGs' protocols explore improving student work, improving teacher work, strategic planning, exploring professional dilemmas, and implementing peer observations (NSRF, p. 11). Bambino (2002) explored the value of CFGs in three separate school settings while Cox (2010) examined how teachers in CFGs benefit from the same quality learning conditions afforded students. Dunne, Nave, and Lewis (2000) observed and conducted interviews with CFGs in order to draw conclusions illustrating the program's benefits. Curry (2008) gathered data supporting and refuting the value of CFGs at the high school level. Harrington (2009) examined the effects of a CFG on a group of seven elementary teachers.

Bambino (2002) provided insight regarding the value of CFGs implemented in three different schools. Bambino first described the makeup of typical CFGs to be approximately 12 teachers who desired to meet monthly to examine student work and the teacher work that inspired it (p. 25). The first school in Bambino's study, Felix Varela Senior High School in Miami, Florida, had 11 of the 145 teachers trained as CFG coaches. From the CFG, a science teacher received constructive advice regarding an assessment that was received enthusiastically by students but in her view lacked the intended rigor she had intended. Through input from the CFG, this teacher was able to

[3] The CFG concept began from the work of the Annenberg Institute for School Reform in 1994 (NSRF, 2012, p. 8).

make the appropriate adjustments to secure rigor by clarifying the purpose of the assessment (Bambino, p. 26). The second school in this study, Manual Education Complex in Denver, Colorado, involved the entire staff of 70 teachers in the CFG process. The fruits of this work produced a common student portfolio assessment for students in tenth and twelfth grades (Bambino, p. 26). The final school in this study, C. W. Henry Elementary School in Philadelphia, Pennsylvania, was able to use CFG protocols to expand teachers' perceptions and preconceived notions of students to better meet student needs (Bambino, pp. 26-27). Bambino concluded this analysis of CFGs by stating, "The work [of implementing CFGs] is *critical* because it challenges educators to improve their teaching practice and to bring about the changes that schools need, but the process is neither negative nor threatening" (p. 27).

Cox (2010) claimed that CFGs could lead to effective professional development for teachers, particularly when a librarian led the CFG. The author stated, "It has taken us decades, but educators are finally realizing that teachers learn best under the same conditions that are advocated for students" (p. 33). A typical CFG is described as a collection of six to 10 teachers who commit to meeting over a two-year period and must recognize that involvement in CFGs involves open, honest sharing that requires a level of trust (Cox, p. 33). CFG protocols such as "chalk talks" and "tuning" are discussed and deemed valuable for teacher development. According to Cox, "The tuning protocols could be an ideal way to find areas of collaboration between teachers, identify missing content or skill development, or prevent unwanted/unintended student outcomes" (p. 34).

The work of Dunne et al. (2000) involved data gathered during the observation of CFG meetings at 12 schools (five high schools, five elementary schools, and two middle

schools) from various areas of the country. In addition to observing CFG meetings, interviews were conducted with CFG teachers and non-CFG teachers. Student and teacher work samples were also collected prior to and after CFG implementation for comparison (Dunne et al., p.2). In interviews, teachers who took part in CFGs stated they found the CFG form of professional development more fulfilling because it was continual, related to their teaching and their students' work, and it involved small groups of trusted colleagues (Dunne et al., p. 4). Through classroom observations and interviews, teachers in CFGs revealed an increased execution of student-centered instruction and increased student expectations (Dunne et al., p. 5). According to Dunne et al., "One theme that emerged from teacher interviews was that many teachers became more thoughtful about the connections among curriculum, assessment, and pedagogy as they participated in the CFG activities" (p. 4).

Curry (2008) gathered mixed results in a study that examined the effects of CFGs with 25 participants that included both teachers and administrators in a three-year study conducted at Revere High School located in the Pacific Northwest. Data collection included observation and videotaping of nine CFG meetings, two CFG coaches meetings, and 42 semi-structured interviews with a combined total of 25 Revere teachers and administrators (Curry, p. 379). Curry described the purpose of CFGs to be as follows: "Together, CFG members seek to increase student learning and achievement through ongoing practice-centered collegial conversations about teaching and learning" (Curry, p. 735). Additional conclusions by Curry included, "Schools cannot be intellectually engaging places for students unless their teachers are likewise actively engaged in learning, thinking, reading, and discussing" (p. 735). The Revere staff was described by

Curry as a mature professional community and a collection of teachers who recently participated in a national teacher action research initiative, which may have positively influenced the effects CFGs had on the teachers (pp. 734, 741).

Curry's (2008) data organization included a diverse menu of activities, a decentralized structure, interdisciplinary membership, and a reliance on protocols. Some benefits of the diverse CFG menu included a wide variety of issues that could be addressed from the micro to the macro. A drawback to that menu included the feeling that too wide of a span of work was undertaken resulting in a lack of coherence and depth (Curry, pp. 743-745). The decentralized structure of CFGs afforded teachers an informal forum to review and debate issues in the school that required teachers to define and defend their opinions as well as consider other perspectives. A disadvantage of the decentralized structure of CFGs concerned an overall confusion as to the purpose of their work and the group's inability to connect their reform conversations with action and change within the school (Curry, pp. 749-754). The interdisciplinary structure of the CFG allowed for interdepartmental connections and a sense of collegiality reaching into various departments of the school. Limits to the interdisciplinary element of the CFG included a shallow exploration of content-specific discussions (Curry, pp. 756-760). CFG protocol structure enabled teachers to question colleagues' beliefs and practices in a safe yet critical fashion and provide constructive feedback. Constraints of the protocol structure of CFGs involved the limits to spontaneous topics or needs that were dismissed due to the group's loyalty to the protocols (Curry, pp. 764-767). Curry found specific benefits to the highly structured format of the CFG while recognizing restraints in this structure as well.

Harrington (2009) conducted a study of CFGs at Basal Elementary School in Wilmington, North Carolina. The primary question in this study dealt with the influence participation in CFGs had on teacher practice and collaboration (Harrington, p. 16). Seven teachers took part in the study, which consisted of pre- and post-surveys with open-ended questions regarding collaboration. Additionally Harrington gathered, narrative and descriptive notes to add to the qualitative data (pp. 17-19). The CFG met every other week for six weeks. Team members identified areas of instruction in which they wanted members of the CFG to observe and provide feedback. CFG members were all given notebooks in which to record information from observations, meeting notes, and reflections (Harrington, pp. 20-21). Results of the pre- and post-surveys demonstrated that CFGs provided educators with opportunities to introduce and share instructional ideas (Harrington, p. 33). Other findings by Harrington included: "The positive communication and trust within the group led to the creation of [a] learning community. The learning community created a high morale and had a positive impact on the student's [*sic*] learning and the teacher's [*sic*] practice" (p. 39). According to Harrington, "While some of the experiences that participants had during the study may be similar to other professional development activities, the CFG has demonstrated the ability for teachers to collaborate, share, and reflect on a deeper level" (p. 42).

SECTION THREE: METHODOLOGY

Research Design Overview

The research for this program evaluation was a mixed method approach utilizing both quantitative and qualitative data. Data were collected through one survey and three group interviews.

The survey was used to gather an overview of the District 32 teaching staff's perceptions of the teacher-created SGAs. According to James, Milenkiewicz, and Bucknam (2008), "The quantitative methodology allows educators to understand the characteristics, opinions, attitudes, or previous experiences of groups in the school community" (p.101). This quantitative information will help shape the focus of additional surveys regarding teacher-created SGAs. Survey data also informed the group interviews and will provide potential assistance regarding the future of the teacher-created SGA program. Additional data were gathered from remarks made in comment boxes from three items on the survey, providing additional insights and explanations to the quantitative data.

Survey results were sent to all survey participants through District 32 email addresses at the start of the 2014 school year. Participants were also sent an invitation to attend a data analysis forum at each of the three District 32 school campuses. Survey participants were encouraged to attend these meetings to discuss the survey findings and to ask questions about the survey results. Survey participants attending data analysis forums were encouraged to participate in group interviews.

Qualitative data were collected through group interviews. The purpose of collecting this qualitative data was to add deeper insight into teacher perceptions

regarding the SGA program that could not be gleaned solely from the survey. The group interviews explored how the SGA program potentially increased a teacher's understanding and execution of elements of the Danielson framework and why this might be occurring with some teachers and not others. The end result of this information could lead to this program having a greater impact for more teachers and increasing student learning. According to James et al. (2008): "Rigorous collection and analysis of the words and pictures, gathered as evidence about a topic, enhance the position of educators to build a convincing body of knowledge on which to improve educational practices" (p. 66).

Results of this quantitative and qualitative data have the potential to assist in greater teacher efficacy for District 32 teachers. Recognizing the professional development value of teachers creating their own SGAs may lead to increased student learning.

Demographics

District 32 is a small kindergarten through eighth-grade district located just outside Chicago, in Broadridge, Illinois. The per pupil operational expenditure for District 32 is $13,617 while the per pupil instructional expenditure is $8,013. The student demographics for the district include the following: 32% of the student population is English learners, 63% receive free or reduced lunch, and 11% receive special education services. The district student mobility rate is eight percent. There are two elementary school buildings—Williams Elementary School and Greenview Elementary School—that instruct students from grades kindergarten through fifth grade with one of those schools also having a pre-kindergarten program. Those schools feed into Prairie Middle School,

which provides instruction for grades six through eight. At the time of the survey, the district was home to 2,157 students, 185 certified staff, seven building-level administrators, and five district-level administrators.

Survey Participants

Survey participants in this study included the 166 teachers in the district that were involved in the creation of SGAs as outlined in the new Illinois PERA law. Participants in the survey included 56 middle school teachers and 110 elementary school teachers. Fourteen teachers participated in the three group interviews in this study. At the time that SGA writing had begun, the teachers of this district had three years of experience with common instructional unit planning and two years of experience with common assessment design for these common units.

Elementary school teachers created English Language Arts SGAs, as they are self-contained teachers of all core classes. Middle school teachers created SGAs based on the subjects they taught due to the departmentalized model of instruction implemented at the elevated grades. Specialty classes such as music, art, and technology created assessments within their individual disciplines at both the elementary and middle school levels.

Data Gathering Techniques

Survey

In the spring of 2014, the survey was sent, using District 32 email addresses, to all teachers who create SGAs. SurveyMonkey was used to gather this data from May 26, 2014 through June 9, 2014. The survey consisted of 10 multiple-choice items with a four-choice Likert scale response option for each item (Appendix A). Three of the 10 items on

the survey included response boxes for additional data to inform future surveys and group interview questions. Eight of the survey items began with the common stem, "Being a part of creating PERA assessments…" Six of the 10 survey items referenced elements of the Danielson framework such as knowledge of standards, unit lesson planning, formative assessment, summative assessment, and collaboration. Survey participants were encouraged to notify the researcher if they were interested in partaking in a group interview during the following school year.

Group Interviews

Three separate group interviews took place during October 2014. It is worth noting that the term used for measuring student growth in District 32 was transitioning from "PERA assessments" to "student growth assessments" (SGAs) throughout the time of this research. Thus PERA assessment and SGA are used synonymously. The questions posed in group interviews went beyond the initial prompts put forth by the survey. Group interview questions were informed by feedback in comment boxes on the survey (Appendix B).

The first interview included teachers who were members of the district-wide Inter-rater Reliability Team that was charged with spot-checking the scoring of SGAs throughout the district. All three District 32 schools were represented in this six-teacher interview. Three of the teachers in this interview were from the middle school and three of the teachers were from the elementary schools. Participants in this interview included one kindergarten teacher, two fourth-grade teachers, one middle school math teacher, one middle school science teacher, and one middle school Encore teacher.[4]

[4] Encore classes constitute classes that meet for nine rotations and include such classes as art, music, communication/technology, Spanish, and additional math courses.

The second interview consisted of six middle school teachers of various teaching assignments. Participants in this interview included two ELA teachers, one social science teacher, one science teacher, and one Encore teacher.

A third interview took place with three elementary school teachers representing both of the district elementary schools. Participants in this interview included one interventionist and two third-grade teachers.

Members of the Inter-rater Reliability Team were interviewed as their own group because of their different level of participation in the SGA process. Elementary and middle school teachers were interviewed in separate groups due to the differing circumstances in which they created SGAs. For example, the departmentalized model of instruction at the middle school required many more different types of assessments be created compared to that required by the self-contained model of instruction at the elementary level. Another manner in which the elementary teachers' experience in creating SGAs differs from the middle school teachers' involve the size of the collaborative teams. An elementary level team creating an SGA may include between eight to 12 members while that team at the middle school level may be a single individual as in the case of Encore, two to four teachers (a math team), or six members (an ELA team).

The overall purpose of the group interviews was to investigate if the teacher-created SGAs positively impacted teacher efficacy. All three interviews were recorded to ensure response accuracy.

Data Analysis Techniques

Survey Analysis

The survey was sent to 166 teachers who qualified for participation due to both their involvement in the creation of SGAs and their having student growth as a component of their teacher evaluation. The survey was open for two weeks at the end of the first school year in which SGAs were created and administered. One hundred ten teachers responded.

A data review of the survey was organized on each of the three District 32 campuses in the month of October. Teacher input into the data analysis occurred at these three data reviews. Feedback from this analysis will be used to craft future surveys and it informed group interviews.

Group Interview Analysis

Group interview participants consisted of volunteers from the survey notification and data analysis sessions. Group interview data were analyzed by interview question topics: reflection, collaboration, and assessment literacy. Data analysis occurred during the interviews in order to make adjustments to questions to increase qualitative responses. After interviews, data were organized according to common themes apparent from the coding.

Coding and Themes

Data were analyzed and coded for emergent themes. According to Patton (2008), "Striving for simplicity means making the data understandable, but balance and fairness need not be sacrificed in the name of simplicity" (p. 481). Qualitative data from group

interviews were sorted and simplified to provide insight as to where the SGA program is most beneficial to teachers and areas where it could improve.

SECTION FOUR: FINDINGS AND INTERPRETATIONS

Introduction

This section explores the various data collected regarding teacher-created SGAs. Quantitative data were collected through one 10-item survey, which provided participants an opportunity to offer additional data through comment boxes for three of the items. Qualitative data were collected through three interviews consisting of kindergarten through eighth-grade teachers who were members of the District 32 SGA Inter-rater Reliability Team. Another interview with kindergarten through fifth-grade teachers representing both elementary campuses in the district was a source of qualitative data. A third interview took place with teachers from the district's one middle school with sixth- through eighth-grade teachers. The focus of all data collection revolved around the primary question: What is the effect of teacher-created SGAs on teacher efficacy in the classroom?

This section begins with a review of the survey data and ends with an examination of the qualitative data gathered in the three interviews. Possible interpretations of the data are presented as they relate to increasing student learning.

Survey Results

The following survey was made available at the end of teachers' first year of creating these assessments. The survey was administered to all teachers who created SGAs. Sixty-six percent of the teachers who created SGAs in this first year provided feedback in this survey.

Table 1 illustrates the number of years of teaching experience for all teachers in District 32 teachers. This table also demonstrates a strong correlation between the years

of teaching experience between all teachers in District 32 and those who participated in the survey.

Table 1

District 32 teachers' years of teaching experience.

Years experience	**N**	**%**
1-5 years	65	39.7%
6-10 years	62	37.0%
11-15 years	19	11.3%
Over 16 years	22	13.1%

These data are an accurate representation of teachers in the four different categories defined in the survey. Results from the first survey item in Table 2 indicate a majority of the respondents to the survey were teachers with 10 or fewer years of teaching experience in District 32. Teachers participating in the survey with this degree of experience totaled 75.5%. The total number of teachers in District 32 with 10 or less years of experience totaled 76.7%. This well-balanced array of survey participants gives me confidence that the data collected will provide accurate perceptions of the teachers who created SGAs. A certain degree of credibility has been established in the survey data based on responses to item one.

Table 2

Survey item 1: I have been a teacher in District 32 for _____ years.

Years experience	**N**	**%**
1-5 years	41	37.27%
6-10 years	42	38.18%
11-15 years	14	12.73%
Over 16 years	13	11.82%

The surprising figure in Table 2 is the high quantity of participants in the 16 years or more teaching experience category. This demonstrates that the many seasoned teachers in District 32 are still willing to provide feedback on new educational initiatives. It is also important to note that teachers newer to the district with one to five years of experience felt comfortable voicing an opinion as well.

Table 3

Survey item 2: Being a part of creating PERA assessments has been a _____ experience for me as an educator.

Answer choice	**N**	**%**
Very Positive	10	9.4%
Positive	65	79.8%
Negative	27	25.4%
Very Negative	4	3.8%

The data from item two yields results that suggest District 32 teachers find creating SGAs a positive experience with the total combined positive amount of "very

positive" and "positive" responses registering 89.2%. Though this is a significant quantity of teachers favoring the creation of these assessments, the second highest response indicates 25.4% of teachers found the experience of creating SGAs to be negative. The combined negative total of "negative" and "very negative" responses combined to equal 29.2%. This is a significant quantity that signifies discontent that must be addressed through suggestions for improving this program. The nature of the dissatisfaction with this program reported by a quarter of the teachers participating in this survey has not yet been defined. Further questions for my research include exploring the value of the scope of this work as it relates to instruction versus the procedures and protocols for completing this work.

Table 4

Survey item 3: Being a part of creating PERA assessments has _____ impacted my teaching practice.

Answer choice	**N**	**%**
Highly	45	41.3%
Somewhat	54	49.5%
Not	10	9.2%

A majority of District 32 teachers who created SGAs indicated that doing so impacted their teaching practice to some degree. The combined total for those who indicated the process "highly" impacted and "somewhat" impacted their practice was 90.8%. This overwhelming number suggests creating SGAs had a significant influence on teachers and their daily instruction. This data implies that the impact of collaborating to create these assessments is not limited to merely the assessment-creation process. The

number of teachers indicating the creation of SGAs highly impacted their teaching practice was 41.3%, a surprisingly high amount. This elevated number suggests that this work is deemed a valuable process and effective means of professional development for a significant number of the teachers taking part in this survey. Such a high number of teachers reporting this impact suggests this work may lead to an increase in effective instructional practice for teachers.

Table 5

Survey item 4: Being a part of creating PERA assessments has _____ my level of collaboration in my teaching practice.

Answer choice	**N**	**%**
Greatly increased	16	15.4%
Increased	83	79.8%
Decreased	5	4.8%
Greatly decreased	0	00.0%

Results from the item regarding teacher collaboration with respect to creating SGAs produced interesting results. The combined positive "increased" and "greatly increased" responses totaled 95.2%. A decidedly low number of teachers, 4.8%, reported creating SGAs decreased collaboration in their work. These data are encouraging for the future of the teacher-created SGA program as it is built on teachers' ability and willingness to collaborate. These data suggest that teachers' efforts did involve some level of cooperative work and that this work provided an opportunity for teachers to interact productively with each other. Item four on the survey had a comment box to gather additional data. Fourteen responses were recorded in these comment boxes. Some

responses generated in the comment box included positive elaborations regarding collaboration, while some detailed the stress related to creating SGAs, or the inequitable distribution of work in this process.

Numerous comments provided more detailed support to the combined positive response of 95.2%. One such comment from a teacher with six to 10 years experience in District 32 stated:

> The PERA requires colleagues to get around the table and work together to create an assessment with integrity. In my experience, there was effective collaboration among department members. Since this was new territory for all of us, we really listened to each other. It was helpful to have so many people creating the assessment because everyone brought their own ideas and perspective. It was a very positive learning experience for us as colleagues.

This comment suggests that teachers used each other effectively to overcome the challenges of taking on this new work. This teacher found value in hearing different perspectives and contributing to the collective work as well.

Another teacher with one to five years of experience in District 32 expressed, "I feel I am already very collaborative but feel I have been able to collaborate and share student learning at a deeper level using PERA and PERA formatives [assessments]." This response indicates that teachers collaboratively creating SGAs produced an opportunity to build upon, for this individual, an already established degree of collaboration. A veteran teacher with 16 or more years experience in District 32 shared, "Creating the PERA assessments has led to many discussions of the CCSS, rubrics, assessments, grade-level alignment of standards including vertical alignment and expectations of students

and their growth in learning." This seasoned teacher's comments broaden the scope of the positive aspects of teachers creating their own SGAs to include how this work increased understanding and discussion about teaching standards and assessment literacy.

Other opinions gathered in the comment boxes for item four suggested the creation of SGAs elevated the level of stress put on the teachers. Numerous teachers expressed this concern in their comment box. One teacher with six to 10 years experience in District 32 concluded:

> I feel it has helped me get to know my grade-level peers, but it has caused stress and tension as well. I understand the need to write it, but it was very time-consuming and frustrating with all of the revisions we were asked to make. I feel we started with our best and then learned more and more about what it should have looked like and had to adapt then.

This perspective indicates the complexity of the process of SGA writing. This teacher is recognizing the challenges of dedicating significant time to this new process and needing multiple revisions yet in the end sees that the team is eventually meeting the rigorous expectations of writing these new assessments.

Another teacher of six to 10 years experience in District 32 documented concern regarding the time commitment related to creating SGAs. According to this teacher, "We have spent so much time working on the PERA assessment that we haven't had very much time to do any collaboration lately." This statement expresses the impact of the SGA creation process on teachers' conversations. This type of response suggests that teachers, even if they value the creation of these assessments, believe the assessments need to be less dominant in their collaborative conversations. Another teacher with six to

10 years of teaching experience in District 32 shared, “It has increased the collaboration with some teachers. However it seems as though the same teachers on each team are doing the majority of the work and others are not.” This remark acknowledges teacher collaboration has increased due to this program, but it also suggests a lack of equity in the degree of influence teachers have in these conversations. This comment also indicates that some teachers’ voices may be dominating conversations or that some teachers may be doing more of the work than others due perhaps to their own desire to control the project even out of necessity if the collaborative group lacks the ability to persevere in this new work.

Though an overwhelming amount of teachers felt that creating SGAs was in fact a collaborative process, other teachers have specific experiences that signify areas for improvement in the program.

Table 6

Survey item 5: Being a part of creating PERA assessments has _____ my knowledge of the Common Core State Standards.

Answer choice	**N**	**%**
Greatly increased	20	19.6%
Increased	78	76.5%
Decreased	4	3.9%
Greatly decreased	0	00.0%

A clear majority of District 32 teachers felt their knowledge of the CCSS had increased due to their involvement in creating SGAs. The combined positive response of “increased” and “greatly increased” responses totaled 96.1% of the participants. This

excessively high number in the survey suggests that teachers working with learning standards in their SGAs produced a learning opportunity that increased their understanding of the CCSS. Increased knowledge of the CCSS provides value to teachers, serves as an important component of quality instruction, and can lead to increased student learning. Darling-Hammond (2013) stated, "These standards [the CCSS] are intended to provide guidance for understanding how students learn in progressive fashion along skill strands as well as what should be taught to enable them to be both college and career-ready by the end of high school" (p. 17). A teacher increasing their understanding of the CCSS is valuable, and these data suggest teachers do increase their understanding of the CCSS, which may provide for greater student understanding in the classroom.

Table 7

Survey item 6: Being a part of creating PERA assessments has _____ my knowledge of unit and curriculum development.

Answer choice	**N**	**%**
Greatly increased	12	11.9%
Increased	78	77.2%
Decreased	11	10.9%
Greatly decreased	0	00.0%

With respect to increased knowledge of curriculum development, 89.1% of responding teachers found the creation of SGAs to be beneficial. This amount demonstrates that the creation of these common assessments reached beyond the intended arena of creating assessments for measuring student growth in teacher evaluation. These

data suggest collaborative creation of SGAs produced a broad degree of positive effects that could benefit educators. A teacher increasing their understanding of curriculum development as a result of teachers creating their own SGAs suggests further value in this program. Instructional best practice suggests rigorous curriculum aligned to standards that contain sound assessments might positively impact student learning. Wiggins and McTighe (2005) asserted, "[Curriculum] is a map for how to achieve the outputs of desired student performance, in which appropriate learning activities and assessments are suggested to make it more likely that students achieve desired results" (p. 6). These data in item six suggest teachers may increase their ability to plan more targeted daily lessons, which can result in greater student understanding which parallels the work of McTighe and Emberger (2006) and Dunn et al. (2000). Though 10.9% of the response to this item is a low amount, it is worth noting that this is one of the highest negative responses registered by almost a two-to-one margin compared to other survey items. These negative data could be an area of focus for recommendations for changes the teacher-created SGA program.

Table 8

Survey item 7: Being a part of creating PERA assessments has

_____ my knowledge of formative assessment.

Answer choice	**N**	**%**
Greatly increased	17	16.7%
Increased	79	77.5%
Decreased	6	5.9%
Greatly decreased	0	00.0%

Clearly District 32 teachers felt the creation of SGAs increased their understanding of formative assessment, which mirrors the work of Dunn et al. (2000), Psenik and Balwin (2012), and Darling-Hammond and Falk (2013). The combined-positive total of "increased" and "greatly increased" knowledge in this area was 94.2% of the response. The use of formative assessment in instruction has greatly increased in District 32 over the past four years. Chappuis (2015) informed us that "formative" does not denote an assessment as much as a practice of data collecting that alters instruction. Chappuis defined formative assessment to be, "formal and informal processes teachers and students use to gather evidence for the purpose of informing next steps in learning" (p. 3). Teachers' increased knowledge in this area makes their instruction data-based and flexible, meeting the needs of students. Increases in this capacity stand to make teachers better informed educators, focused on teaching to students' needs, which could bring about greater gains in students' learning.

Table 9

Survey item 8: Being a part of creating PERA assessments has _____ my knowledge of summative assessment.

Answer choice	**N**	**%**
Greatly increased	16	16.7%
Increased	74	77.1%
Decreased	6	6.3%
Greatly decreased	0	00.0%

The combined positive response to this item totaled 93.8% of all participants. This is somewhat surprising as summative assessment is a traditional assessment perspective

and one I suspect that most teachers would feel sufficiently knowledgeable about in District 32. Chappuis (2015) defined summative assessment to be "assessments that provide evidence of student achievement for the purpose of making a judgment about student competence or program effectiveness" (p. 4). This high number favoring increased knowledge in summative assessment suggests teachers may now have a more solid foundation of how quality assessments are constructed, a development that could affect their daily assessments and instruction which parallels Good (2012). Though summative assessments are not new to District 32 teachers, it could be the collaborative nature of the summative assessment creation that provided teachers the additional learning and insight reported in this combined positive number. An increase in teachers' knowledge regarding summative assessment could benefit teachers and lead to greater student learning.

Table 10

Survey item 9: The challenges from being a part of creating PERA assessments were _____ significant.

Answer choice	**N**	**%**
Highly	46	44.7%
Somewhat	51	49.5%
Not	6	5.8%

Results from item nine indicate that a very large quantity of participants in the survey felt some level of challenge with respect to the creation of SGAs. The combined total of "highly" and "somewhat" responses equaled 94.2%. Being an administrator in a building where SGAs were created, I did not find this combined number surprising. What

was surprising is that more responses did not reflect the "highly" challenging option. What these data suggest is that this type of new work is formidable for a little less than half of the respondents but perhaps a manageable duty for almost 50% of those who participated. These are encouraging results as we as a district look for recommendations for improving this program and work within our means to reduce the challenges.

Item nine included a comment box for additional insights regarding how creating SGAs brought about challenge. Twenty-seven out of the 166 teachers chose to provide further data by offering comments. Some of the comments reflected a positive perspective about the challenges of creating SGAs. One such comment by a teacher with six to 10 years experience in the district stated:

> I would definitely agree that there were significant challenges in creating this assessment largely because we did not pilot this assessment with students first. After having students take the assessment, we noticed flaws in our questions as well as our rubrics. However, we were able to correct these flaws for next year.

The comments of this teacher suggest the challenges related to this work stem from its newness and that the assessments had not been fully vetted with students prior to administering. The fact that this teacher and her colleagues were able to pinpoint areas of weakness in the assessment and make adjusts indicates that teacher learning had occurred. This homegrown professional development that produces opportunities for teachers to take part in instructional problem solving stands to increase teacher efficacy, which might result in greater student learning. A teacher of one to five years teaching experience in District 32 shared a similar thought about the work being challenging but beneficial. This teacher stated:

> There were times when we were ahead of the game in finalizing our PERA, then all of a sudden we would be behind and rushing to meet deadlines. This was STRESSFUL!!! I for the most part did enjoy being able to create the assessments. Being involved in other teams if [*sic*] curriculum writing this year was a bit difficult. At the end of the day it ALL pays off! The kids benefit from well-written assessments and we benefit as well.

This teacher expressed the challenges related to the amount of time that was required to invest in this work. The confusion regarding being on track and then behind in the assessment creation timeline suggests a breakdown in communication and understanding regarding the expectations for this program. In the end, faced with these challenges, this teacher saw value in teachers creating these assessments—namely, the benefits for students**.** A teacher of 16 or more years in the district provided, "Yes the challenges were still there but, once overcome, they were worth it." This again suggests that this new work was viewed as difficult and challenging for District 32 teachers, but some teachers were still able to see the value in this novel exercise.

Some negative responses related to the challenges of creating SGAs were shared in the comment box on item nine as well. A teacher of six to 10 years experience stated, "Most of the challenges faced in writing the PERA assessments stem from the amount of time that is put into designing a valid assessment as well as finding model items rigorous enough to guide us in the PERA writing." This response echoes previous comments about the time commitment this work entails but also brings to light the lack of exemplars for teachers to use in guiding their assessment writing. A teacher of 11 to 15 years experience in the district stated, "Too much too fast." This comment again indicates that

time was a factor and a challenge for teachers as they created SGAs. Not only does this individual think more time could have reduced the challenges related to this work, but the comment suggests the amount of new thinking was arduous for teachers. A teacher with over 16 years experience in the district shared, “Way too much time was spent on revisions.” This response suggests that the meticulous nature of this work was a challenge for some teachers. This type of remark also implies the perception that a lack of communication may have occurred between the administration and teachers regarding the level of rigor and substance required for such assessments or perhaps a breakdown on the receiving end of assessment expectations. A teacher of six to 10 years experience in District 32 shared, “Different administrators giving conflicting feedback. Making changes according to one person’s opinion and then changing it back for another person’s was frustrating.” This again reinforces the perception that communication issues may have played a part in teachers feeling challenged while creating these assessments.

Further feedback gathered in the comment boxes for item nine referenced collaboration and how it posed a challenge when creating SGAs. “In our department, we all had very different ideas and ways of collaborating. Disagreeing in our department usually ended up creating personal and professional barriers that have not yet been addressed,” stated a teacher with 11 to 15 years of experience in the district. This comment suggests that teachers’ differing abilities in collaboration played a role in this program being a challenge. A newer teacher with one to five years experience in District 32 shared:

> It’s hard to create an assessment with 10 people and trying to have all members agree on the different parts of the test. Also, we felt that the test should be scored

> by unbiased eyes in the future. Teachers should not be grading their own students' test. A grading committee should be formed for more uniform scoring throughout the district.

This comment reiterated the difficulty teachers faced when asked to reach consensus on their assessments. This teacher also questioned the validity of these assessments and how they were scored, particularly when teachers score their own assessments that affect their evaluation.

The comment boxes for item nine revealed additional data regarding assessment literacy and teacher integrity. One teacher of six to 10 years experience in District 32 shared:

> It's a wonderful opportunity to create our own assessments, but at the same time how research based could our own assessments be? There is a question of validity and reliability. Not that I think a company on such short notice would do better. These could be important factors since the dual role of PERA assessments is to monitor teacher as well as student growth.

These comments suggest teachers possess a sense of insecurity when it comes to their abilities to design instruments to detect student understanding. This most definitely needs to be addressed and given support because this is the very foundation of a teacher's purpose.

Another teacher with six to 10 years teaching experience in District 32 expressed the view:

> At the end of the year, teachers were discussing how to "grade" the assessments more than how to improve the questions. Comments I heard were "If we grade

harder in the fall and easier the second time…." I guess too much is at stake for the teachers to actually use it to inform instruction.

This comment raises serious concerns. If this program is going to be successful, it is imperative that these sentiments are addressed and challenged. As a district, plans must be put into place to avoid unethical scheming and as a profession, standards must be established collectively in order to empower teachers to maintain a moral compass in this work.

Table 11

Survey item 10: Being a part of creating PERA assessments could further influence my teaching practice if changes were implemented.

Answer choice	**N**	**%**
Strongly Agree	10	10.6%
Agree	65	69.2%
Disagree	18	19.2%
Strongly Disagree	1	1.1%

A significant amount, 79.8%, of surveyed teachers agreed to some degree that changes could benefit how teachers create SGAs. It is surprising that only approximately 11% of these teachers "strongly agree" there is a need for change in our SGA creation procedures considering how many teachers commented on how the process was challenging. It is interesting that this item offers the highest number of opposing views to the majority opinion expressed. The combined-negative total of "disagree" and "strongly disagree" responses equaled 20.3%. This is a surprisingly high quantity considering how

demanding this program was reported to be through survey comments and how few resources were available for teachers to reference.

Item 10 had a comment box that collected 11 responses. Some of the data collected in comment boxes ranged from “no need for change” to “change is an inescapable occurrence.” One teacher with one to five years experience shared, “Change is inevitable. I think I have learned that in education change always occurs. No matter what change, in my opinion is ALWAYS good!” This comment suggests education is an ever-changing field that requires a constant need to adapt. A teacher with six to 10 years teaching experience in the district who is supportive of the status quo with respect to creating SGAs stated:

> I think the experience of creating the assessment has already influenced my teaching practice greatly. I think the current process of creating these assessments definitely requires some trial and error. This is what helps us learn and grow as educators. I don’t think changing the process of creating these assessments is necessary.

The words of this teacher establish content and satisfaction with the current conditions under which people design, implement, and score SGAs.

Some teachers suggested changes could be made to help improve how SGAs are created while some voiced concerns that should be addressed to improve this program. A teacher with one to five years teaching experience in District 32 said, “The theory of and reasoning for PERA assessments are excellent. However, the implementation and strict time limits (testing windows, and amount of time between pre- and post-tests) renders

them less effective and frustrating for teachers." A teacher with six to 10 years teaching experience in District 32 expressed a concern about how SGAs affect instruction:

> I think the PERA assessment should be aligned more to what we actually teach in class. The standards were attached to lesson plans, but we made very little effort to actually teach the standards. We just gave the same type of assignments over and over and hoped the students would perform better without teaching them how to perform better.

This view suggests the connection between SGAs and daily instruction is not occurring. This disconnect needs to be shored up as this is the very essence of why SGAs are proposed—to be a more authentic measure of student growth than a standardized test.

A small number of teachers found item 10 to be unclear. A teacher with 16 years or more of teaching experience in the district stated, "I do not understand this question." A teacher with six to 10 years experience in District 32 shared, "Not sure what kind of 'changes' might be made. Changes with the law? Changes with a content area PERAs? Specific PERA task changes?" It is possible that more data regarding changes to the teacher-created SGA program could have been gathered had this item been more clear to teachers.

Summary of Survey Findings

Findings from the survey and feedback from survey comment boxes are shared below.

A majority of District 32 teachers believe creating SGAs:

- Serves as a valuable exercise that provides benefit to educators in areas of collaboration, knowledge of the CCSS, curriculum development, and assessment literacy
- Produces challenges for educators
- Requires some changes to benefit the program

Some District 32 teachers expressed concern regarding:

- The amount of time involved in creating SGAs
- The validity and benefit of the teacher-created SGAs
- The communication of expectations with respect to teachers creating SGAs

Additional information regarding individual teacher opinions regarding the SGA creation program will be shared in the next segment of this section through a review of group interview results.

Teacher Interview Participants

Three separate interviews took place in order to gather data on teachers creating their own SGAs. It was decided that the gender of each teacher in these interviews would be depicted as female though in reality some of the teachers interviewed were male. The purpose of this decision was to maintain the anonymity of all interview participants. The first interview involved six teachers from the district's Inter-rater Reliability Team which cross-reference teachers' scoring of SGAs. This team meets periodically to sample teachers' scoring and provide feedback to teachers about the validity of these assessments. It was decided to interview these teachers separately as they may have different perceptions about teachers creating SGAs. In addition, teachers not on this team may feel more comfortable speaking about SGAs without the Inter-rater Reliability Team

in their interview. The following table reflects the teachers involved in the interview with members of the Inter-rater Reliability Team.

Table 12

Group interview: members of the Inter-rater Reliability Team

Teacher	**School**	**Grade/Subject**	**Years experience in D32**	**Total years experience**
Teacher 1	Prairie	Science	38	38
Teacher 2	Prairie	Math	24	25
Teacher 3	Greenview	1st grade	9	9
Teacher 4	Williams	4th grade	2	5
Teacher 5	Greenview	4th grade	2	2
Teacher 6	Prairie	Encore	16	16

Another interview was conducted with middle school teachers of grades six through eight. It was decided to keep this conversation solely among middle school teachers as they are involved in a different, specific context for the creation of SGAs. An additional difference in the middle school teachers' experience with respect to creating SGAs includes their daily common planning period and the fact that all middle school teachers work within the same building. The following table reflects the teachers involved in the middle school interview.

Table 13

Group interview: middle school teachers

Teacher	School	Subject	Years experience in D32	Total years experience
Teacher 7	Prairie	Encore	13	17
Teacher 8	Prairie	Social science	22	24
Teacher 9	Prairie	English Language Arts	8	9
Teacher 10	Prairie	English Language Arts	10	12
Teacher 11	Prairie	Science	10	10

A third interview was conducted with only elementary school teachers who teach grades kindergarten through fifth grade. It was decided to interview these teachers as their own groups because they create English Language Arts SGAs only, have common planning periods once a week, and work between two buildings within the district. The following table reflects the teachers involved in the elementary school interview.

Table 14

Group interview: elementary school teachers

Teacher	School	Grade/Subject	Years experience in D32	Total years experience
Teacher 12	Williams	2nd Grade	4	4
Teacher 13	Williams	Interventionist	3	11
Teacher 14	Greenview	2nd Grade	8	8

The comment below illustrates one elementary teacher's view regarding the opportunity to create her own SGA.

Teacher 13: I feel like [creating SGAs] gives us a certain dignity…people don't really hear that it gives us a certain level of respect for ourselves and for our profession that we are trusted enough to create an assessment that's used in our own evaluation. I don't think people really value that or just understand how cool it is.

Teacher Interview Themes

Over the course of the three interviews conducted, the following four themes emerged from these conversations.

- Theme 1: Teachers creating SGAs created changes in their daily instructional practice
- Theme 2: Teachers learned from each other when creating SGAs
- Theme 3: Teachers creating SGAs produced varied degrees of collaboration
- Theme 4: Teachers creating SGAs presented an opportunity for teachers to examine the manner and purpose in which they evaluate students
- Theme 5: Changes could be implemented to increase the effectiveness of the SGA creation process

Theme 1: Changes in Teachers' Daily Instructional Practice

This theme emerged in all three of the interviews, which reflects that changes are occurring in each of the three District 32 schools. The following conversation with members of the Inter-rater Reliability Team suggests instructional practice has benefitted from teachers participating in the creation of SGAs.

Researcher: I'm interested in knowing to what level this type of work, creating these assessments, has changed your perspectives about assessment or changed your approach to daily instruction. Has it affected you, nudged you in different ways, or opened up new ways of thinking about things, or reaffirmed things in your classroom on a daily basis?

Teacher 1: I think it has made me really stop and think about everything I do and how it all ties together. And there is much more of a connection to the lessons I

am planning and the formative assessments I am giving... I can see the connection and that is the biggest change that I've seen in what I do.

Teacher 2: I think the type of assessments I am giving are different. I mean before I focused more on skills and what they could do at a lower level, and now I'm trying to test what their conceptual understanding is a lot more. Do they really understand the standard or the objective that we are teaching? And how can I craft our questions so they are a little bit more rigorous…I think my students are achieving more. It's not just about their grade but about their learning, what are they understanding.

Teacher 3: One of the things I find myself doing is really a lot more feedback or exemplars that we can talk about and say "Okay this is how this student answered this, or this is how…can we make this better?" It's having these conversations with students, what they did right, what we can do better. I think that is very big.

Teacher 4: Talking about feedback, being very specific to [*sic*] feedback that's what I find myself doing a lot with a small groups of kids. Like you all didn't do this correctly so let's go back and how can we specifically fix this one area that you need to work on.

These comments made by members of the Inter-rater Reliability Team, representing all three District 32 schools, suggest teachers creating SGAs is affecting daily instruction in a positive fashion. Teachers are able to take what they have learned from this process and apply these new perspectives toward other aspects of their teaching. These teachers' comments suggest an increase in teacher efficacy for participants in this program. These findings correlate with those of Darling- Hammond and Richardson (2009) that supported the notion that efficacious teachers provide increased quality learning experiences for their students as well as Moolrenarr et. al. (2010) who asserted that an increase in teacher collaboration produced an increase in teacher efficacy. Additionally Hargreaves and Fullan (2012) and Ross and Bruce (2012) who purported collaborative school cultures cultivated increased confidence in teachers. Teachers who collaboratively create their

own SGAs view their curriculum as more connected, their assessments as more targeted, and their success criteria and feedback to students more explicit and detailed.

A further example of such thinking is reinforced by a comment from a middle school teacher:

> Teacher 7: I think I am a better teacher in regard to the standards…I believe I understand them a little bit better. I believe I can create curriculum and formative assessments that show both students, myself, and parents what they can do. So I think yes, it's made me a better teacher. I'm not sure I like the format or all the other parts, but I think that it's helped refine my skills.

This comment continues to reinforce the generalizable potential of the skills practiced when teachers create SGAs, particularly as it relates to learning standards. This parallels the findings of Darling-Hammond and Falk (2013) who found:

> Examining and assessing students' work helps teachers learn more about what their students know and can do, as well as what they think. Doing this in the context of standards and well-designed performance tasks stimulates teachers to consider their own curriculum and teaching. (p. 6)

The interesting revelation with this previous teacher's comment involves her increased ability to communicate students' understandings to the student and their parents. Most surprising is the teacher's reflection regarding her own ability to evaluate her students in a more meaningful and valid way since participating in writing SGAs with her colleagues. This teacher's comment suggests that she has increased her ability to assess her students, which could potentially lead to greater student understanding.

The following comment illustrates how teachers feel the need to incorporate a particular quantity of instruction within the testing window to ensure students demonstrate growth.

Teacher 12: I think for me it makes me more conscious to the fact that they need to get to that certain point. You want to make sure that you teach everything that you're supposed to be teaching, but then at the end of the day, you're like "Oh wait…I haven't started doing this or doing that," and it's like well, "You need to get going on that because by this month they need to be able to be secure with it." It does affect [instruction] definitely, and it does make me nervous sometimes because sometimes, I'm like "Did I do enough?"

This comment suggests that when teachers are working with SGAs, they are hyper-aware of where they need to be in the curriculum in order for students to succeed. This comment also implies a certain degree of stress that accompanies teachers who work with SGAs. This teacher indicates a shift in the lexicon used to express student achievement as being "secure" with a standard. A middle school math teacher espoused the benefits of this program and how it increased her level of assessment literacy.

Teacher 2: We're becoming more educated on what makes a good assessment—not just a PERA assessment but an assessment overall. You know, we are better educated on formative and summative and what it looks like—how to craft the questions maybe to meet the standards. I think that all comes out of the PERA assessment because we had this expectation and people are learning from the revisions even though we may not always like to be told to do it again or to be shown a better way to do it.

This statement regarding assessment literacy echoes the work of Stiggins (2013) in particular. These comments made by teachers suggest that teachers creating their own SGAs have provided the opportunity for greater understanding of assessment, curriculum, and accuracy of communication with students through feedback and use of success criteria such as exemplars.

Theme 2: Teachers Learning from Each Other

Common to all three interview settings was the fact that teachers reported that they learned from working with other teachers in the collaborative meetings when

creating and scoring SGAs. The conversation with the Inter-rater Reliability Team is captured below.

Researcher: Did you learn from each other?

Teachers 1: Yes.

Teachers 2: Yes.

Teachers 6: Yes. I thought I was saying that. I learned far more than I ever would have learned from a book or a lecture. From writing my own [SGA] as I did, I thought I'd learn by doing, no. I learned by talking. I learned by having a different perspective. I learned by rewriting these based on suggestions.

This exchange demonstrates the common belief that when given the opportunity to collaborate about authentic content, teachers can learn from these conversations. This data parallels the findings of Stiggins and DuFour (2009) as well as what Schmoker declared in DeFour, Eaker, and DeFour (2005), that teachers learn best from interacting with other teachers. A conversation with elementary teachers supported this thinking as well.

Researcher: As far as value of collaboration, do you feel you've learned things from your colleagues?

All: Yes.

Teacher 11: Totally.

Researcher: You see value in being given the opportunity to collaborate with your…

Teacher 11: I think even at the end of the day, if you are having like even those discussions that you've observed where maybe there is tension or whatever, no matter what, you're always going to learn something. Obviously you've heard what that person had to say and it's embedded in you and you're going to be thinking about it no matter what.

Teacher 13: Even if you don't agree with what that person is saying, you're measuring your belief or your perception against what they're saying.

This dialogue reinforces the concept that learning occurs when teachers discuss real-life

concepts that affect their daily instructional practice. The last comment by teacher 13 also suggests that these collaborative conversations create the opportunity to foster teachers' ability to be reflective of their instructional practice. A final example of such thinking came out of the middle school interview.

Researcher: Have you learned from working with your counterparts on the PERA assessment?

Teacher 8: We really work to bring out each other's strengths and really help where there might be weakness. So I really think [the SGA] has helped a lot.

Here again is a teacher emphasizing the professional benefits from conversations with colleagues. This finding coincides with the work of Darling-Hammond and Falk (2013) who found, "Involving teachers in scoring assessments is powerful professional development because it connects teacher learning directly to their examination of student learning, and gives them the opportunity to think together about how to improve that learning" (p. 5). The research collected in District 32 suggests that teachers from both elementary and middle school view collaborative conversations about instructional practice to be a form of professional development.

Theme 3: Varied Degrees of Collaboration

The theme regarding teacher collaboration when creating and scoring SGAs began as an exploration of possible challenges teachers experienced regarding collaboration. As the interviews evolved, the topic of collaboration naturally split into two sub-themes:

- Effectiveness of collaboration
- Teachers' ability to collaborate

Effectiveness of Collaboration.

The following conversation regarding effective collaboration took place between two elementary school teachers on the Inter-rater Reliability Team.

Researcher: Another thing I'm really interested in is how well teams collaborated and how many people had a stake in getting assessments out the door.

Teacher 4: When we score our assessments, we'll write down any question or kind of set aside any [student SGAs] we are unsure about and then when we come together as fourth grade at Williams, we'll go through and show examples to the team and kind of discuss what do we think this should be and come up with a score collaboratively…Then we'll bounce our ideas back and forth with Greenview and then share how we scored certain things so it was very collaborative in scoring them I feel.

Teacher 5: There are so many different, so many people between just fourth grade.

Researcher: So how many would be on your team?

Teacher 5: There are five on our team [Greenview].

Teacher 4: So ten people [District-wide fourth-grade team].

Teacher 4: So when we sit down around the table, to say let's change this, we do collaborate very well but sometime it takes a while to understand each other's ideas. It actually takes us quite a while. We do collaborate well, it is just time consuming with that many people to make actual changes to the assessment. It doesn't happen easily but is does happen. So that's good.

This conversation depicts a large group of teachers from more than one elementary school effectively collaborating to create and score SGAs. What is evident in these teachers' conversations is that more than anything, the one thing that school systems must provide is time. Teachers need time to meet and to hear each other's thoughts and perspectives which is also suggested in Stiggins (2013). The variable outside of the school system's control involves teachers having patience with one another and a growth mindset that allows for new ideas, approaches, and paradigms to be considered without feeling threatened or offended.

Effective teacher collaboration was reported within the middle school setting as well. Below, a teacher-leader in a fragmented Encore team consisting of music, art, STEM (science, technology, engineering, and math), math lab, math bonus, Spanish, and literacy/communication classes reported success with collaboration.

Teacher 7: Not to toot our own horn, but [the assistant principal] was like, "You guys are amazing. You collaborate." We are very unique, every single one of us, and we all have our own little thing that comes with it, but we have come together and created something. I think that's why we get along, because we're the oddballs in the building, or whatever it is, and that's what joins us together, and because of that we've found a way to collaborate. It doesn't mean we always love to do it, or there could be some heated debates, but at the end there is that utmost respect for one another, but I'm not sure how you teach that. I don't know if it's the leader. I came into a very dysfunctional department, and I'd like to say it's me, but I don't think that's what it was, because I've been with other groups, and there was an incredible leader, and they can't bring them together, so I'm not sure how you get to that collaboration piece, but there is something missing.

This teacher's remarks suggest many things that make quality collaboration occur. This team overcomes disagreements and the fact that they might not always love to collaborate, but they get the job done when necessary. When these teachers come together they have an elevated level of respect for each other that this teacher points to as the catalyst for effective collaboration. These findings are similar to those in Dunn et al. (2000) and Harrington (2009). One could make the argument that these teachers are successful at collaboration because they meet as a cross-curricular team. When they create SGAs, they often function as a department of only one teacher. This is the case because there is only one art teacher, one STEM teacher, and one Spanish teacher, etc. This arrangement allows for one-voice domination in many of the team member's instructional work such as creating SGAs. The collaboration that does take place occurs with other members of the team who do not have a stake in the success of the particular

SGA that they might be collaboratively discussing. This being the case, it is worth noting that teachers coming together to collaborate on things they are not vested in is no small feat either.

Responses in these interviews suggest collaboration isn't easy. It takes fortitude and a growth mindset. When a system expects collaboration that produces results such as a product like an SGA, the system must provide time to allow consensus to be reached. It is also apparent from the thoughts shared by teachers that the professional relationships amongst the team of teachers collaborating is of the utmost importance. Teachers must create relationships based on mutual trust and establish their own needs and conditions in order for collaboration to thrive.

Teachers' Ability to Collaborate.

One view that teachers from all three District 32 schools, teachers of all levels, from regular classroom teachers to members of the Inter-rater Reliability Team shared involved teachers' ability to collaborate. Below are several conversations that depict such an opinion.

In the following conversation with elementary school teachers, it is expressed that teachers lack an innate ability to collaborate.

Researcher: There seems to be an assumption that because you're a teacher and it's a people business, that you're naturally gifted and skilled in collaborating.

Teacher 13: I think that's the opposite. I think teachers are very territorial and especially what you were saying about if you're writing the SGA, if you're the primary author of it and people are offering criticism …

Teacher 11: Some people will accept it, some people will be like, "Don't even say anything because you didn't even pick up a pencil" or something. They'll feel that way. We're never talking to each other like that of course. If you look at it, if you're open minded about it then it's that constructive

criticism. You're being professional. You're discussing it. Not everybody is like that.

Teacher 13: Not everybody has the strongest understanding of how to communicate effectively and how to communicate criticism effectively, and how to evaluate their own criticisms. Sometimes I think the criticisms that come out are not really valid.

In this discussion, teachers expressed their tendency to personalize their work to a great degree. It is also shared that teachers may not be equipped to accept constructive criticism in a way that is professional and promotes growth. As well, they may not be able to effectively communicate constructive feedback in a way that fosters increased learning and understanding for their colleagues. It is also interesting that while discussing teachers' inabilities to accept criticism, a thought was shared that perhaps not all feedback given to teachers is actually valuable. This comment suggests that even progressive-minded teachers hold critical feedback as suspect at times.

This same perspective regarding teachers' innate inability to collaborate was expressed in an interview with teachers from the middle school. The following interview reveals how middle school teachers feel about teachers' ability to collaborate.

Researcher: I'm wondering if it's accurate to say that it's a given that if you're a teacher you probably can collaborate.

Teacher 8: I don't think it's a given. I think it certainly is necessary, but I don't think everyone knows how to collaborate or how to collaborate effectively. I think because we are human beings and we all have feelings in this manner, I think sometimes that in itself can maybe get in the way…"

Teacher 11: I would say for the most part we're all capable of collaborating, but not necessarily willing to collaborate.

Teacher 10: I see a lot of collaboration. I think in ELA I see a large department that has sometimes a lot of cooks in the kitchen, and I think that at that point you start to lose either people's strengths or people's interest.

Again in this dialogue it is expressed that teachers are not necessarily skilled when it comes to collaboration. The interesting opinion stated by one of these teachers suggests teachers are equipped to collaborate; they just sometimes choose not to do so. That surprising revelation creates a more challenging issue to address, as it is not merely people not knowing how or having the resources to succeed in this enterprise. This thought suggests it is a conscious and deliberate act by some teachers to not collaborate with their colleagues simply due to lack of desire. This implies a more complicated problem that will require a more complex solution. If this is the case, teachers will not just need an increase in knowledge with respect to collaboration, they will need a greater understanding of its purpose and benefits.

Later in this interview, one middle school teacher shared:

Teacher 7: I don't think teachers are equipped to collaborate. I think it's a stereotype, just like not all teachers like to be up in front of a large group. That's a stereotype. I think that this is definitely a stereotype of teachers. I think there needs to be some training. Things need to be modeled, practiced, but you also can't teach people to get along either.

This teacher's insights address multiple stereotypes that people hold regarding teachers. This information illuminates the dangers in generalizing commonly held beliefs about teachers. Even if a school system provides opportunities for teachers to collaborate, that does not guarantee quality conversations that result in productive work opportunities for teachers will occur.

In a third interview involving members of the Inter-rater Reliability Team, the same sentiment was expressed regarding teachers' lack of natural ability when it comes to collaboration.

Researcher: I'm curious about educators in general, are they innately collaborative people?

Teacher 1: Well I think education has changed. It used to be where you didn't need to collaborate. You were in charge of your room. You closed your door, and you taught. And that is no longer the case so you need to collaborate.

Teacher 2: I agree, I was thinking the same thing. It has changed. And I think that is where some of the discomfort is coming in for some people. The game has changed a little bit, the rules are different, you have to rely on other people. You have to rely on others. You are expected to.

Teacher 1: And if you're playing by the old rules, you're going to get kicked out of the game.

Teacher 2: Yeah...yeah.

Teacher 1: And people don't innately know how to collaborate. I don't think so.

Teacher 2: I don't think so. But it's a large group of people when you are talking about educators, there will be all types of personalities.

Teacher 1: Because what seems like a successful collaboration could just be that you got two people where one person says okay to everything the other person does or says what they want and that is a great collaboration because there is such great harmony there. Then you have two people that have very strong opinions and neither of them are going to back down no matter what. So there's lots of personalities there. Though I think people think they know how to collaborate. But I don't think that is necessarily true. I think we're getting better because you have to be able to do it in order to survive.

Teacher 2: Yeah.

Researcher: Do you think it would be beneficial for some type of...do people need resources in order then to improve in this area or do you improve by doing it?

Teacher 2: I think you improve by doing it with guidance. Like I'm noticing a lot of the college students, just listening to them and the questions they ask us, and the things, I can tell there is more of a focus where they know they need to work with other people.

Teacher 1: Because they don't know any other way, because that is what they are teaching them. But I think it is what you said, they learn it by doing it but with the guidance. You can't take a bunch of people and suddenly, "We're going to give you the collaboration 101 lecture," because then you are going to turn all of those people off.

Teacher 2: It's exactly like with our students.

Teacher 1: You have to differentiate.

In this exchange between teachers, it is revealed that the current state of education requires teachers to collaborate if they wish to be effective in the classroom. These teachers reinforced too that collaboration is not necessarily an innate trait in all teachers. Both of these teachers feel the collaborative skills of teachers can be enhanced through guidance, support, and modeling an example of quality collaboration. They did caution others to be aware that the support offered to increase collaboration needs to be respectful and meet the needs of individual teachers if it is to be effective. Curry (2008) found value in the protocols of Critical Friends Groups that provide safe guidelines for teachers to conduct analytical dialogues regarding teachers' instruction.

Theme 4: Examination of Manner and Purpose for Evaluating Students

One repeated theme that reappeared in interviews involved the need to examine how students' understanding was evaluated. This idea of changing how students were scored or graded was very prevalent with middle school teachers in particular. A change in philosophy as well as teachers' vocabulary as it related to the evaluation of students became apparent in teachers of all grades. The below exchange between members of the Inter-rater Reliability Team suggests some of these changes.

Teacher 4: Now I'm giving more feedback and comments and how to change to improve.

Teacher 1: With us, you use to give an assignment, and the first thing said is, "Is this going to be for a grade?" And if it isn't for a grade, then they have a whole different mindset. But now they don't ask that. I don't think I've been asked that all year.

Teacher 5: Probably because we are giving so many formative assessments they know we are assessing all the time so they don't ask.

These comments suggest that when teachers provide students with consistent, specific feedback, a shift in student perceptions can occur moving away from the purpose of school as collecting points for a grade to that of learning. This change in mindset can bring about much more authentic learning opportunities for students as they turn from merely seeking a high grade to actually trying to learn and understand.

The following interview segment with middle school teachers relays this perception about how grading students aligns with the progressive instructional and assessment practices happening in the middle school.

Teacher 7: I'm struggling to come up with how the grade is reflective of their progress on the standards, and I have a sense of guilt with some of my students sitting and looking at my spreadsheet, "And this is what you scored on this and look at how far you've come on this," And how is that going to show in a grade book and what grade goes with that? Because there are some students that have shown a lot of growth. How do I go back and grade some of those things? Because to be very honest, I haven't been grading the same way all the time. I've been focused on a standard and I've been focused on four, three, two, one, and basing it off of my PERA rubric, so I'm not sure how to put a letter grade, and I'm finding that that's confusing to explain to students. They know what the standards were, they know where they're at with them, but I'm not sure what grade goes with that. I feel like it's a whole other puzzle that's come up for me.

Teacher 8: That's a whole other piece because of the grade.

Teacher 7: It's a good thing. I had a parent conference last week, and I felt probably the most prepared I've ever felt for the conference. I had many formative assessments. I had all these different ways to show what they were or weren't doing, but I couldn't have explained what that grade would have been.

Researcher: As an educator, do you feel that that's a necessary piece in the puzzle of explaining where a student is by giving them a letter grade, or do you think the con- ... It sounds like you had a conversation that didn't necessarily deal with a letter grade. Is that right?

Teacher 7: Yeah.

Researcher: Are you saying that was a good conversation?

Teacher 7: It was a great conversation. I think we got a lot out of it. I think that many parents are still looking for a grade though because they don't understand. I'm an educator, so as a parent I'm not looking for the grade. I don't want the grade. I actually want the other stuff that goes with it, but I find parents still want the grade, and I'm frustrated with then how to come up with that grade, because I feel like I have a lot of great stuff but I'm not sure that I have the time to show that on the report card through comments, and if they don't get to the conference, how could I explain that grade?

Researcher: Right.

Teacher 9: It does trip a lot of students, because I had a student that wrote me a sticky note on her midterm, and she said, "When you have a moment I'd really like to talk to you about my grade." We did have a lot of data, a lot of formative assessments. This is where you were and this is where you are now and this is what you need to do in order to demonstrate secure status, and you're not quite there. I think she's been somebody who has gotten all As and Bs, because she comes here every day. She participates really nicely, she asks good questions, but she's not secure with the standard, and as much as I admire her tenacity and her willingness to do well, she's not quite there…but that is a hard piece that we just haven't done yet. Although I will tell you, when I sat down with her and said, "Let's take a look at your data," she's like, "Okay." But I think it was like maybe talking with her family, and when her family sees that she's getting a B-, and a B- is meant to be celebrated. She's had a lot of secure hard work, but as far as other people are concerned, secure scores should be a strong A.

Teacher 10: I think with grading, that's the standards. It'll be helpful, but it's just that parents are going to be ... It takes us a few years to figure it out, and parents are use to a certain thing, but if we change…a painful year or two, and then they figure it out, but it is better. It is more of a score that reflects more where a child's at, so I feel better about that too if you're looking at the standard. You can show things.

The comments from these teachers suggest that instruction in the middle school has evolved to a new level yet the manner in which students are assessed outside of SGAs perhaps remains stagnant. These teachers are expressing a disconnect between how they view student progress and the language or system by which students are formally evaluated (i.e., grades). It is interesting how the teachers recognize the increased

effectiveness they have in student progress conversations when they borrow some of their practices and terminology from their scoring of SGAs.

The words of the middle school teacher below further illustrate how teachers regard grading at the middle school as still in flux.

> Teacher 2: I let the students know how they are succeeding when before it would have been with an A, a B, a C, or a D where now it is, "You have improved on this. This is where you need some…" You know even the assessment, the scoring is different. We've divided it into sections now; it's not an overall grade. It's more of, "On this I can add integers; this is how you did. I can subtract integers; this is how you scored." They are even giving the feedback of, "Oh I'm really good at this but this is where I need some help," and then we can target some extra instruction or time with them. And they are seeing that we're seeing that. So it' not just an overall, "You're an A student or you're a B student." So it's kind of changed how we even talk to the students.

The comments of the teacher above suggest that the assessment and instructional practices involved with the SGAs have provided a new foundation with which to consider all student work and evaluations. It is apparent from the reflections shared by these teachers that their efficacy has increased with respect to their instruction. These findings correlate with Bandura (1993), who stated, "Teachers' beliefs in their personal efficacy to motivate and promote learning affect the types of learning environments they create and the level of academic progress their students make" (p. 117). These teachers' comments also reveal that students are beginning to adopt more progressive ways of viewing their own academic learning and progress.

Theme 5: Changes to Program

Collaborative creation of SGAs is a new program to the administrators and teachers of District 32. Because of that, it is understandable that there would be room for improvement through changes. Particularly with the level of challenge reported by

teachers in the survey, feedback from teachers regarding changes is important to the future of this program. The following conversation took place between elementary level teachers. These teachers examined how teams of teachers might better be suited to work together collaboratively in the future.

Teacher 13: I think that there needs to be more team building activities. There needs to be professional development experiences that allow people to develop trust and team building.

Teacher 14: I know they sent teachers to the "seven habits" training years ago. Just applying those habits, maybe even at a staff meeting, giving everyone that book, read a chapter each month, we'll talk about it and jigsaw it in groups or however it might be. It's just finding a program to help them be more collaborative.

Teacher 13: I feel too, from what I hear people saying in chitter chat that a lot of people feel that administration does not select people for leadership positions that other people feel comfortable working with. Do you know what I'm saying?

Researcher: I know what you're saying.

Teacher 13: I think if people had more of a voice in who was taking leadership positions, maybe they would be more responsive to participating in certain things. I don't know. To get people to have a growth mindset, they have to feel like it's safe to take a risk and there's not somebody sitting keeping score. For some reason, that perception is extremely strong here. It's thick.

Researcher: You're not talking building wide, you're talking district wide?

Teacher 13: I think it's a little bit of both. I think it's particular to each team.

These comments suggest teachers in District 32 are in need of some level of professional development with respect to collaboration. This reinforces the previous collection of comments regarding teachers not being innately equipped to collaborate effectively. The interesting revelation from this exchange involves the judgment regarding teacher-leaders. Teacher 13's comments suggest that the teacher-leader sets the tone for collaboration on teams and if the leader is not creating a safe environment for everyone,

then productive, collegial interactions will not occur. Most disturbing in this exchange is the concept of a "score" being kept that might be shared with others outside of meetings. This perception implies the climate and culture in the district is in serious need of healing. Whether this is merely a perception or fact, a trusting, safe environment is doomed if even a small quantity of teachers feel this is reality.

A suggestion regarding changes to the administering the SGA arose from this same conversation with elementary teachers. These teachers recognized issues with the testing window for the baseline SGA occurring just weeks into the school year.

Teacher 12: One problem we came across this year was we couldn't give the SGA right away. We had wonderful teachers coming around and helping us with assessments, our Type 3, but the problem with that was that we needed them to focus on accommodations. They couldn't with their schedules so we had to wait until three weeks in order to give the [Type 2] SGA and then all of our units call for …

Teacher 11: Teaching it.

Teacher 12: Teaching the elements of a story or character traits before we start the unit, and we can't do that because we would …

Teacher 11: You're teaching to the test.

Teacher 12: We want to get a good snapshot of where they're at in the beginning of the year. That was hard this year. It wasn't last year.

This conversation recognized that teachers need the baseline assessment to occur as early as possible because they are holding off on providing instruction until after the baseline data are collected. These teachers shared that they found themselves potentially wasting instructional time because they didn't want to skew the data by pre-teaching anything on the SGA baseline. The concern these teachers expressed suggests the elevated value they place on instructional time and how they wish to maximize instructional time regardless of their SGA schedule.

The following conversation considered changes that might benefit the validity of SGAs by allowing for two separate assessments used for the baseline and the outcome assessments.

Teacher 12: One advantage of having a different assessment, I didn't attend the B.U. [Broadridge University professional development opportunity] but I talked to some of my teammates who did. [The Assistant Superintendent for Learning] had said, "Wouldn't it be great to go over their pre-assessment with the class and say 'Look at what you did.' Give them the feedback because they're not getting them [again]." They take this for three days and that's it. It goes in a box. How great for the kids and for us to give it to them.

Teacher 11: I agree with that because for the book that we picked, obviously the copies were ordered but I have the big book and there's [*sic*] so many things that I can do with it. Every time I run into one, I'm like I can't grab it because I need to give it to them for the post- and they can't do it, and it stinks because it would be a learning opportunity to go back and say, "Well, you only supported this character trait by saying this, but look there's pictures, there's text here, there's all of this." That's where our conversation that we're having now, that's the way we looked at it with a positive mindset but not everybody will. Some people are going to be like "Are you kidding me? Again?"

These teachers bring up an interesting point regarding validity in assessment. Can a test be truly valid without the same assessment being used for the baseline and outcome assessments? Or can an assessment only be truly valid if two tests of similar rigor are administered because the use of the same assessment allows students a second attempt at the same test items? These comments suggest a desire to use two assessments because that allows teachers to use the baseline assessment as a teachable resource. When a single assessment is used, that test is not revisited in order to avoid students being taught how to grow on a specific assessment as opposed to the standard-based skills. These teachers are savvy enough to recognize that not all teachers would welcome creating a second

assessment as the development of a second test might be more than a challenged group of teachers could withstand.

Further discussion regarding changing the teacher created SGA program took place with the members of the district Inter-rater Reliability Team. This conversation suggests improvements to the assessment scoring process.

Researcher: What do you think are the strengths of having teachers do this whole process that we do? What are things that have helped, what are strengths and what are changes that would make it better?

Teacher 4: One thing that I have always thought that would make it more valid maybe is having some, I know the Inter-rater Team pulls some of the assessments and rescores them but if there was a way to maybe exchange grade levels, have at least two sets of eyes grading each assessment, would probably, would just...to have multiple people looking at each assessment making sure the scores are you know, valid.

Researcher: As opposed to pulling and sampling?

Teacher 1: We scored eight separate assessments of the PERA IIs and we had disagreed, not the overall but the individual tasks, we disagreed with three of the eight, then what the teachers gave for one of the tasks. So if that is with these eight and you multiply that by 200, that's quite a few scores where there is a disagreement with and I think if more people were looking at them then there would be more validity like you were saying.

Teacher 4: Right, even if it's not just the Inter-rater [Team] that would be huge, if you had to score every assessment but if you had to score a different grade level's. I also thought it was really interesting this being my first year on the team, even just looking at a different grade-level's assessments is really interesting and seeing how their rubric is worded and just to get another perspective. But I think if there were at least maybe two people scoring each assessment.

These comments suggest a perceived need for an increase in the level of validity in SGA scoring. Teacher 4 stated that the additional scoring did not necessarily need to be done by the Inter-rater Reliability Team but might be done by all teachers. This teacher also shared how grading assessments other than their own creates a professional development

opportunity for teachers with respect to assessment literacy. Teachers viewing and scoring other grade level's assessments would promote articulation amongst grade levels. This teacher conversation suggests that the additional work involved in creating two separate assessments for baseline and outcome assessments might not be welcomed by some teachers. Teachers did not mention the extra work of additional scoring duties and how some teachers may not view scoring two separate assessments favorably. This could be because this suggestion came from teachers on the Inter-rater Reliability Team who may have an elevated sense of duty and dedication to their work.

The below conversation between middle school teachers addresses the challenge of providing differentiated instruction and assessments throughout the school year and how these modifications are not made available to students with SGAs.

Teacher 9: I think a lot of the work that we've done in the district has been about differentiation and getting kids to work at different levels, and I think that that can be challenging then, because if you know you have children in your room that are reading at a fourth- and a fifth- and a sixth-grade level, but you're a seventh-grade teacher, and during the day you know that you need to differentiate and provide scaffolded instruction to support those students, but then we're giving them all a test that we know that all the kids can't read or be successful at. I know that there's a lot of differentiation that goes on in my classroom that won't ever be reflected in the student growth. I can show you data when this child was able to read a fourth-grade passage or a fifth-grade passage with a standard at their level. They were able to be successful, but that's not going to count for me for my student growth.

Teacher 7: I was thinking the same thing. We have a lot of students that we're now giving the iPads to record their responses, because their written expression is poor, but we're crippling them for some of our PERAs, and we weren't realizing that until we started to really do our data review that, "How do you differentiate at the same time as providing them with that same test?" We've seen that as well.

The sentiments shared by these teachers suggest that there are areas in which students are growing academically that are not being captured in the data from SGAs. These remarks

imply that the philosophy of the district's instruction, which is to meet students where they are developmentally, is not being adhered to in the SGA process. These comments suggest that individual student growth, not just student growth with respect to grade level CCSS is valuable and valid and worthy of consideration in teacher evaluations.

Teacher 9: Having been part of the Inter-rater Team and the Assessment Design Team, sometimes I felt like the purpose and the outcomes were not always ... and in having conversations with other colleagues in the building who were struggling with PERAs that not all of the "knowns" were shared at the start. It could be just because they weren't known to those people either, but I think that was for some people ... You spend time and you spend energy in creating, and then the rules change, and you're like, "Well that was great for the sake of collaboration, but my time is more important."

Teacher 10: Right. "Next time I'm not going to do it." Even that it opens itself to a situation where not everybody can participate in it or be in a situation where "Hey, I did this last year. I'm not going to do this again. I spent last summer doing it." It opens itself up to a situation of like I said before, not welcoming maybe a collaborative environment or not even having everybody be a part of it.

As shared repeatedly in comment boxes from the survey, teachers feel the manner in which they received information and direction regarding SGAs did not alleviate stress or confusion. Teacher 9 in this interview rightfully assumed that district-level administrators did not always have much information to share at the start of this process as virtually no other district in the state had taken on the task of creating SGAs. As direction and feedback were made available to district administrators, that information was relayed to teacher teams and departments who were creating SGAs. Even with that being the case, a lack of appreciation seems to exist between teachers and the district office.

A final suggestion for change to the SGA program came from a middle school teacher on the Encore team. Encore teachers see students in nine-week rotations as

opposed to an entire year like core teachers of subjects such as math, English language arts, science, and social science.

Teacher 7: In a selfish regard, I feel that PERA needs to be revisited for a 30 to 40 day class rotation. We're evaluated the same way, the teacher is, that has months to do it with the same amount of standards, the same amount of tests, but less time to do it. I feel that a conversation needs to happen. We need to know what the law is, because it keeps changing, and I'm not sure that everybody knows what a PE or a health teacher does and what their timeframe is. I'm not sure they get what Encore is. There's always the flexibility to change and make a schedule fit a situation. If it's *that* important then I feel like that needs to be revisited.

This teacher, who has particular circumstances with her teaching assignment, seems to recognize challenges to the testing window. These comments suggest it is necessary to look at the one-size-fits-all limitations to the PERA law. Teachers with specific conditions related to the types of classes they teach may be penalized in their evaluation if alterations are not considered to meet the specific differences in theses teachers' duties.

Opinions expressed by teachers in all three interviews suggest that the teacher-created SGA program has influenced both teachers and students' perspectives on learning. The comment below suggests how this program has benefitted the stakeholders of District 32.

Teacher 1: The mindset is a thing, I think there is a definite different mindset than a year ago, or even from the end of last year, and it's in the kids too because the kids will come up and they will say, "I understand this, I'm ready to take my assessment again." And they'll ask you as opposed to coming in and hoping they guess better this time. They're actually, "I'm ready for it. I've done this, this, and this, and now I'm ready."

This paradigm shift has taken place for many teachers in District 32. Because of this, instruction and assessments have greatly improved, a development that can lead to greater student learning. The other revelation in this teacher's statement is that a paradigm shift has occurred with students as well. This is arguably a more important

finding as it suggests District 32 may be developing self-actualized learners, which may lead students to greater understandings.

Summary of Findings from Teacher Interviews

Data from group interviews revealed that teachers and students alike benefit from the implementation of this program. The following findings can be deduced from the four themes that emerged in this research.

Finding 1: Changes in Teachers' Daily Instructional Practice

A majority of the teachers interviewed reported changes in their daily instruction entailing the incorporation of specific elements practiced while creating their own SGAs. These findings occurred at all three District 32 campuses and thus both elementary and middle school settings. Additionally, teachers expressed that the changes to their instruction made them more effective at providing students with specific feedback, success criteria, and assessment creation and scoring. Members of the district's Inter-rater Reliability Team were able to report more elaborate and detailed changes to their instruction and knowledge regarding assessment literacy.

Finding 2: Teachers Learning from Each Other

Teachers from both elementary and middle school settings expressed the belief that collaborative conversations with their colleagues increased their knowledge regarding instructional practice. This consistent finding suggests that when teachers discuss and exchange ideas regarding instruction, it functions as a form of job-embedded professional development. Some teachers recognized that learning could occur when a teacher disagrees with a colleague's feedback and perspective. It was reported that this

type of conversation could lead to teachers reflecting on their own views, which in itself could be of value.

Finding 3: Varied Degrees of Collaboration

Teachers from all schools, both the elementary and middle school settings, reported positive occurrences of collaboration regarding the creation and scoring of SGAs. However, teachers at the elementary level faced with collaborating between two separate schools, reported a greater level of collaboration. Middle school teachers reported successful collaboration experiences in specific teams or departments depending on the size of the collaborative group and the personalities within the group. Smaller two-person teams were reported to be successful at both the elementary and middle school levels. Additional data gathered suggested that teachers from both the elementary and middle school levels believed the ability to collaborate is not necessarily a skill that all teachers naturally possess. One differing opinion from a middle school teacher expressed a disbelief that teachers are capable of collaborating but some teachers simply choose not to do so.

Finding 4: Examination of Manner and Purpose for Evaluating Students

A common struggle for teachers from both elementary and middle school settings entailed the challenge to produce a grade or "score" for students once instruction and assessment practices had changed because of teachers creating their own SGAs. This theme was particularly prevalent for both middle school teachers who were members of the Inter-rater Reliability Team and those who were not.

Finding 5: Changes to the Program

Teachers from all three District 32 schools and both school settings expressed a need for change in the teacher-created SGA program. Several teachers in all three interviews believed some degree of professional development regarding collaboration could benefit teachers. Additionally, the testing window in which the SGAs were administered was a suggested area for change. Teachers expressed challenges involving having to delay instruction regarding certain learning standards in order to not pre-teach content to be assessed on the baseline SGA.

A common area for change included the accuracy of communication from the district office regarding guidance for creating the SGAs. Other teachers proposed increasing the level of inter-rater reliability beyond that of the district-wide team dedicated to this task. Multiple teachers at the middle school level shared that the differentiated assessment practices that take place during the school year are not permitted with SGAs. Because of this, teachers are suggesting such accommodations and modifications for students be allowed to provide data based on student growth at a student's instructional not chronological grade level. This circumstance would be practiced when such conditions were happening on a daily basis for students.

Another proposed change involved implementing two separate assessments. This would allow teachers to use the baseline assessment as a teaching resource where that is not permitted when the same assessment is used for both the pre- and post-assessments. A perspective reported from a middle school teacher involved considering how teachers who do not instruct students for an entire year, working in timelines such as nine-week

rotations, might have different guidelines for their evaluations than a traditionally assigned teacher.

SECTION FIVE: JUDGMENT AND RECOMMENDATIONS

Introduction

The evaluation of this program is designed around the primary research question: What is the effect of teacher-created SGAs on teacher efficacy in the classroom? Judgments will be made regarding findings as they are deemed to be positive, negative, or unintended. Positive findings are defined as results that were perceived as beneficial to teachers with respect to their instructional practice. Negative findings are results found detrimental to teachers' instructional practice, those that hold no value to teachers' instructional practice, or those related to potential flaws in the research methodology. Unintended findings were surprise results that were not foreseen prior to research taking place. Following the judgments will come recommendations where suggestions, alterations, deletions, and adjustments to the programs will be made.

Judgment

The results of this research suggest that when teachers embark upon creating their own SGAs in a collaborative manner, their efficacy increases. Determining whether this was the case was the primary goal of this research. This judgment is supported by the results of the survey data that imply teachers' levels of competence in assessment literacy, as well as their knowledge of the CCSS and instructional lesson design increased from participating in this program. Improvements in such facets of instructional practice provide teachers with increased competencies in the classroom, which may lead to greater student learning. Results from the three group interviews provided additional insight that support the theory that teachers' efficacy increased from their involvement in creating and scoring their own SGAs.

The most positive result of this research could be that District 32 teachers recognize value in creating and scoring their own SGAs. Teachers revealed this insight in comment boxes on the survey and repeatedly in the group interviews. This positive perception by teachers is important as it offsets another common finding in this research, namely the challenges presented by having teachers create their own SGAs. In spite of the difficulties stated in section four, a majority of teachers provided positive feedback regarding this program even when suggesting potential changes. An additional positive finding in this research is that the teachers do not seem to feel the improvements to their instruction and assessments have reached its highest level. All dialogues with teachers indicated the trajectory for positive change in instructional practice will continue to increase for teachers.

Though most of the findings for this program could be classified as positive, there were some negative findings that need to be shared. Critical findings that surfaced from both the comment boxes on the survey and the interviews included the elevated level of frustration from teachers over the perceived changes to the rules and expectations when creating their own SGAs. Some of these criticisms may be considered within the control of district office while much of it was beyond the local control of anyone in District 32. Such examples include the number of tasks involved in each assessment and the number of standards that need to be included in each task. This groundbreaking work by teachers was so new, there was nothing to compare it to or refer to for exemplars or guidance. Much of this work was actually *establishing* the rules and guidelines, which created the unfortunate conditions that led to many adjustments and overhauls to procedures and expectations. Another critical finding from an interview involved the perception that

teachers are now "teaching to the SGA." This negative perception stems from a past practice of teaching *to* high-stakes tests that are not necessarily aligned to daily instruction. The difference with teaching to an SGA is it involves teaching to a standard, which is teaching to a skill. These are the learning standards that are to be taught each day in the teacher created curriculum. Following the guidance of Wiggins and McTighe's *Understanding by Design,* it is best practice to design the assessment prior to formulating the instruction that will lead to student success. An additional negative finding came from a middle school teacher who felt the content of her discipline was taking a backseat to the literacy standards she found herself teaching for her SGA. This teacher voiced a concern that partaking in creating SGAs made her more knowledgeable about the CCSS but made her a less effective in teaching her content.

Regarding the methodology of the survey, multiple teachers reported the survey could have benefitted from a neutral option. This choice was made by the researcher in order to force participants to choose a positive or negative response to items and avoid the potential of collecting a large quantity of uncommitted responses. An additional potential challenge to the survey included multiple teachers who reported they did not understand what item 10 was asking. This could have reduced the accuracy of the results for this item.

One unanticipated result of this research was the consistency in which teachers expressed challenges with grading students now that instruction and assessment practices had progressed. Some teachers discussed providing students with feedback based on their performance on standards while some actually referred to standards-based grading by

name. This was an unintended outcome from this research that will be mentioned in the recommendations.

Recommendations

The purpose of this evaluation of the teacher-created SGA program was to establish if teachers taking part in this process increased their sense of efficacy in the classroom, which might potentially increase student learning. In order to increase the likelihood of this occurrence, the following recommendations are made:

1. Differentiate professional development: The teachers of District 32 are involved in collaborative activities such as the creation and scoring of SGAs. Not all teachers may have knowledge of what effective, collaborative communication entails or how to execute such communication. It has also come to light that teachers may lack resources to assist them in this capacity. It is recommended that professional development and/or resources be offered at varying degrees to teachers in order to further skills and understanding of cooperative communication.
2. Create vertical articulation of SGAs: In order to increase the continuity of the district's SGAs, it is recommended that an iterative process or opportunity be created to foster vertical articulation of SGAs amongst grade levels. This could be done kindergarten through eighth grade for ELA. As numerous science, social science, and Encore SGAs often contain ELA standards in their assessments, it would benefit these teachers as well to be aware of the rigor being assessed at different grade levels. This recommendation is intended to move beyond providing all teachers access to SGAs through the District 32 intranet as is

currently the case. District 32 needs to move beyond "access" to "action" in this respect.

3. Provide information regarding teacher evaluation law: Because PERA, the Illinois law requiring student growth to be included in teacher evaluation is so new, it is recommended that extensive resources regarding the law and any subsequent revisions be made available to teachers. These resources will help inform teachers of the complexity involved in this law and help quell misinformation that is currently causing confusion and frustration with teachers.
4. Provide specific instructions for SGAs: In order to increase teachers' potential for success when creating SGAs, it is important that they are given explicit instructions regarding expectations for these assessments. This instruction for teachers should mirror that provided students: clearly defined, immobile learning targets; explicit success criteria; and exemplars modeling desired products. Now that District 32 has created SGAs for two years, it is recommended that distinguished-level assessments from District 32 teachers be used as these exemplars.
5. Explore standards-based grading: Because the work with SGAs has increased teachers' focus on learning standards in their instruction and assessment practices, it is recommended that District 32 begin exploring the implementation of standards-based grading.
6. Continue work to improve SGAs: A final recommendation would be for District 32 to continue to view these assessments as a new type of work and continue to

strive to improve the assessments each year. One middle school teacher summed up this recommendation during an interview:

> As exciting as it is to be one of the only schools, if not the only, that is created a PERA test for teacher evaluation, it could also be a negative if we don't continue to evolve and improve the assessment. My concern is that we will create a test whose purpose is good, and overall is reasonably effective, but then be too busy patting ourselves on the back and close ourselves off to improving the PERA and making it even better. Meanwhile other districts may take our model, and unlike us, build upon the PERA model we give them and create an even better assessment. It makes me think of the evolution of baseball stadiums. In the 1990's there was a need to replace baseball stadiums in many cities. Chicago was one of the first to take on this huge task. They went on to create what is now known as US Cellular Field. It's a nice ballpark, much better than the old Comiskey Park. But could it be better? Absolutely. No one talks about US Cellular Field anymore as a prototype. Other towns took the new model and evolved it into an even better new ballpark. State of the art technology, great seating, and a retro look that made people feel like they were enjoying baseball the same way it was viewed for the past 100 years. I don't want our PERA to be US Cellular Field. We shouldn't be a footnote that is barely remembered for all of our efforts.

The research in this study demonstrates that both the quantitative and qualitative data suggest a positive effect on teacher efficacy when teachers create and subsequently score their own SGAs. Because increased efficacy in teachers can lead to greater student understanding, other school districts are encouraged to adopt some type of program that allows teachers to create and score their own SGAs.

References

Adie, L. (2010). Developing consistency in teacher judgment formation through online meeting centres. *Journal of Learning Design, 3*(2), 14-26.

Bambino, D. (2002). Critical friends. *Educational Leadership, 59*(6), 25-27.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148.

Bensenville Elementary School District 2. (2014).

Calik, T., Sezgin, F., Kavgaci, H., & Kilinc, A. C. (2012). Examination of relationships between instructional leadership of school principals and self-efficacy of teachers and collective teacher efficacy. *Educational Science: Theory and Practice, 12*(4), 2498-2504.

Chappuis, J. (2015). *Seven strategies of assessment for learning.* Hoboken, NJ: Pearson Education.

Chappuis, J., & Chappuis, S. (2006). Understanding school assessment: A parent and community guide to helping students learn. Boston: MA: Pearson Education.

Cox, E. (2010). Critical friends groups: Learning experiences for teachers. *School Library Monthly, 27*(1), 32-34.

Curry, M. (2008). Critical friends groups: The possibilities and limitations embedded in Teacher professional communities aimed at instructional improvement and school reform. *Teachers College Record, 110*(4), 733-774.

Danielson, C. (2013). *The framework for teaching evaluation instrument*. Princeton, NJ: The Danielson Group.

Darling-Hammond, L. (2013). *Getting teacher evaluation right.* New York, NY: Teachers College Press, and Oxford, OH: Learning Forward.

Darling-Hammond, L., & Falk, B. (2013). *Teacher learning through assessment: How student-performance assessments can support teacher learning*. Retrieved from https://www.americanprogress.org/issues/education/report/2013/09/12/73978/teacher-learning-through-assessment/.

Darling-Hammond, L., & Richardson, N. (2009). Teacher learning: What matters? *Educational Leadership, 66*(5), 46-53.

Demir, M. K. (2013). Perception of classroom and branch teachers working in primary schools towards in-service education. *African Journal of Adult Education, 1*(4), 75-82.

DuFour, R., Eaker, R., & DuFour, R. (2005). *On common ground: The power of professional learning communities.* Bloomington, IN: Solution Tree Press.

Dunne, F., Nave, B., & Lewis, A. (2000). Critical friends groups: Teachers helping teachers to improve student learning. *Phi Delta Kappa International Research Bulletin, 28.* Retrieved from http://www.pdkintl.org/edres/resbul28.htm.

Good, J. (2012). Crossing the measurement and writing assessment divide: The practical implications of inter-rater reliability in faculty development. *The WAC Journal, 23*, 19-30.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school.* New York, NY: Teachers College Press, and Toronto, ON: Ontario Principal's Association.

Harrington, C. D. (2009). *Critical friends group: Effects on teacher practice and collaboration.* Retrieved from http://dl.uncw.edu/etd/2009-2/harringtonc/carolynharrington.pdf.

Illinois State Board of Education, (2014). Retrieved from http://www.isbe.net/rules/archive/pdfs/50ARK.pdf.

Illinois State Board of Education & Illinois Performance Evaluation Advisory Council (2011). Retrieved from www.isbe.net/peac/ppt/ISBE.

James, E. A., Milenkiewicz, M. T., & Bucknam, A. (2008). *Participatory action research for educational leadership*. Thousand Oaks, CA: Sage Publications.

McTighe, J., & Emberger, M. (2006). Teamwork on assessments creates powerful professional development. *National Staff Development Council, 27*(1), 38-44.

Moolenarr, N. M., Sleegers, P. J. C., & Daly, A. J. (2011). Teaming up: Linking collaboration networks, collective efficacy, and network achievement. *Teaching and Teacher Education, 28*(2), 251-262.

National School Reform Faculty. (2012). *Self-Guided Tour to NSRF Critical Friends Groups.* Retrieved from http://nsrf.iuserveit.org/sites/all/themes/NSRFtheme/assets/pdfs/Self-Guided%20Tour%201.0.pdf.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage Publications.

Powell-Moman, A. D., & Brown-Schild, V. B. (2011). The influence of a two-year professional development institute on teacher self-efficacy and use of inquiry-based instruction. *Science Educator, 20*(2), 47-53.

Psencik, K., & Baldwin, R. (2012). Link data to learning goals: Common district assessments connect teaching effectiveness to student performance *Journal of Staff Development, 33*(4), 30-33.

Ross, J. A, & Bruce, C. D. (2012). Evaluating the impact of collaborative actions research on teachers: A quantitative approach. *Teacher Development, 16*(4), 537-561.

Stiggins, R. (2013). *Defensible teacher evaluation.* Thousand Oaks, CA: Corwin.

Stiggins, R., & DuFour, R. (2009). Maximizing the power of formative assessments. *Phi Delta Kappan, 90*(9), 640-644.

Wiggins, G., & McTighe, J. (2005). *Understanding by design.* Alexandria, VA: ASCD.

Williams, M. L. (2010). *Teacher collaboration as professional development in a large, suburban high school.* (Unpublished doctoral dissertation). University of Nebraska, Lincoln, Nebraska.

Appendix A

Teacher Survey I: The effects of creating PERA assessments on professional practice. Administered at end of year one.

1. I have been a teacher in District 2 for _____ years

1-5 years
6-10 years
11-15 years
Over 16 years

2. Being a part of creating PERA assessments has been a _____ experience for me as an educator.

Very positive
Positive
Negative
Very negative

3. Being a part of creating PERA assessments has _____ impacted my teaching practice.

Highly
Somewhat
Not

4. Being a part of creating PERA assessments has _____ my level of collaboration in my teaching practice.

Greatly increased
Increased
Decreased
Greatly decreased

Comment box

5. Being a part of creating PERA assessments has _____ my knowledge of Common Core State Standards.

Greatly increased
Increased
Decreased
Greatly decreased

6. Being a part of creating PERA assessments has _____ my knowledge of unit and curriculum development.

Greatly increased
Increased
Decreased
Greatly decreased

7. Being a part of creating PERA assessments has _____ my knowledge of formative assessment.

Greatly increased
Increased
Decreased
Greatly decreased

8. Being a part of creating PERA assessments has _____ my knowledge of summative assessment.

Greatly increased
Increased
Decreased
Greatly decreased

9. The challenges from being a part of creating PERA assessments were _____ significant.

Highly
Somewhat
Not

Comment box

10. Being a part of creating PERA assessments could further influence my teaching practice if changes were implemented.

Strongly agree
Agree
Disagree
Strongly Disagree

Comment box

Appendix B

Group Interview Questions

Change in Practice

Did your involvement in creating SGAs affect your daily instruction? If so how and did it impact your practice?

Did your involvement in creating SGAs affect your daily assessment practice? If so how?

Did creating SGAs increase your understanding of formative assessment? How?

Collaboration

How was the level of collaboration when creating SGAs?

Were there any challenges collaborating with your team on the SGAs?

What are possible ways collaboration could improve when creating SGAs (if necessary)?

Does collaborating with others increase your abilities as an educator?

Did you learn from your colleagues while working on SGAs?

Did your colleagues learn from you while working on SGAs?

Suggestions for Change

What are the strengths of the SGA program?

What are possible changes that could make the SGA program more beneficial for you?