

Putting it to the Test: Examining the Use of Standardized Test Scores to Select Teacher Candidates into Preparation Programs

Abstract

This pilot study analyzed the long term predictive ability of standardized test scores to select teacher candidates that will go on to become effective teachers. Participants were tracked from their teacher training program into full-time teaching positions in order to determine if there was a connection between a high score on the PRAXIS II standardized test and novice teacher effectiveness as measured by the Classroom Assessment Scoring System (CLASS) instrument. Novice teachers were observed at three times throughout the school year. Findings suggest that standardized test score are not entirely effective in predicting which teacher candidates will go on to become effective teachers.

Keywords: teacher education, teacher candidate selection criteria, teacher education reform

Introduction

With the ever increasing accountability placed on schools to generate high student achievement, teacher preparation programs worldwide are feeling pressure to produce teacher candidates that go on to develop into highly effective teachers. Government mandates and legislation at both federal and state levels have once again initiated inquiry into the ways that teachers are trained and more specifically into the ways in which teacher candidates are selected into teacher preparation programs. The reality of what teacher candidates face upon graduation is daunting, making the decisions about who to select into teacher education programs increasingly significant and critical. Those that are selected into teacher preparation programs need to be able to handle challenging circumstances and contexts far beyond what they will receive in their training (Feiman-Nemser, 2003). When teacher preparation programs make decisions about who to accept into their program, they are essentially making predictions and gambling on who they believe will grow and develop into highly effective teachers *after* they have left the program.

To make these far reaching decisions, teacher preparation programs have used a variety of admittance criteria to select teacher candidates into their program. Petersen and Speaker (1996) surveyed 50 teacher education programs across the United States and found that 97% of the teacher education programs reported using the GPA to admit students, 67% reported using standardized admittance tests such as the PRAXIS I/Pre-Professional Skills Test (PPST), 44% reported using letters of recommendation, and 38% reported utilizing individual interviews. More recently, Casey and Childs (2007) determined that teacher preparation programs continue to use grade point average (GPA), standardized test scores, written profiles, letters of reference, and individual and group interview performance to make determinations about which students to select into their teacher education programs. Though most teacher preparation programs utilize

multiple forms of information to make these far reaching decisions, these findings demonstrate the heavy reliance and emphasis teacher preparation programs continue to place on teacher candidate knowledge, as indicated by standardized test scores or GPA, in selecting teacher candidates into teacher education programs. Teacher candidate knowledge, as indicated by GPA and standardized test scores, may be important to higher education institutions comfortable with an emphasis on grades and test scores, but does this type of data provide necessary knowledge about the type of teacher an individual will become?

The purpose of this study was two-fold. First, this study sought to extend the research literature by tracking teacher candidates into the classroom once they have secured employment to determine if standardized test scores (such as PRAXIS II) are effective in predicting not only student teaching success but classroom teacher effectiveness as well. Most of the studies analyzing teacher candidate selection practices have examined the correlation between selection criteria and student teaching performance and success. Little research has been done examining the connection between selection criteria and the performance of teacher candidates *after* they have acquired a teaching position to determine the long term significance of established teacher candidate selection criteria.

Second, this study introduced the use of the Classroom Assessment Scoring System (CLASS; Pianta, LaParo, & Hamre, 2008), as a research tool to measure long term effectiveness of teacher candidate selection criteria through classroom observations analyzing degrees of emotional support, classroom organization, and instructional support as evidenced in the classroom of these teacher candidates after they have graduated from teacher preparation programs. Too frequently, the research on teacher candidate selection criteria utilizes student ability and success as measured by cooperating teacher and university supervisor feedback or

grades. These data are fraught with issues of poor reliability and validity. Educational researchers have struggled to develop valid and meaningful measures that can predict teacher candidates that will go on to display highly effective teaching practices (Byrnes, Kiger, & Shechtman, 2003). This study sought to answer this need. The research questions providing the framework for this pilot study was as follows: (1) Do teacher candidates that perform well on standardized admittance tests, such as the PRAXIS II, go on to interact effectively with students and create high quality classrooms? and (2) Does the CLASS instrument serve as valid and reliable way to determine the effectiveness of teacher candidates after they have begun teaching in a classroom?

Theoretical and Empirical Implications

Defining and Measuring Highly Effective Teachers and Classrooms

The first hurdle to surmount when attempting to determine the effectiveness of using standardized test scores to select teacher candidates is to have a clear definition of what it means to be a highly effective teacher. Unfortunately, clearly defining a highly effective teacher may be as elusive as the process used to train one. There is little agreement on the characteristics needed to be a good teacher (Squires, 1999; Stronge, 2002; Turner-Bisset, 2001). The No Child Left Behind Act (2001) issued the challenge of requiring that all classrooms have a “highly qualified teacher.” But just what is a highly qualified teacher?

Within some states and jurisdictions, standards have been established in an attempt to clearly articulate the knowledge, skills, and attitudes that teacher candidates need to have in order to be considered effective in the classroom. Casey and Childs (2007) explained that in Canada, the *Standards of Practice for the Teaching Profession* is used to outline the standards expected and required of teachers in Ontario. These areas include commitment to students and

student learning, professional knowledge, teaching practice, leadership and community, and ongoing, professional learning. In the United States where this study was situated, the standards established by the Interstate New Teacher Assessment and Support Consortium (INTASC; 2009) have been adopted as a means to define the knowledge and skills required of and demonstrated by highly effective teachers. These standards include knowledge and skills encompassing content pedagogy, student development, diverse learners, multiple instructional strategies, motivation and management, communication and technology, planning, assessment, reflective practice and professional development, and school and community involvement. Though these standards provide a framework for discussions regarding ways to train teacher candidates, they do not provide a valid and reliable method to measure classroom or teacher effectiveness.

For the purposes of this study, the CLASS observation instrument was utilized to make valid and reliable judgments about classroom effectiveness demonstrated by the teaching abilities of participants in this sample. The CLASS observation instrument was designed to assess the teacher-student interactions as experienced by the average student in the classroom. The CLASS observation instrument is described more fully in the methodology section.

Teacher Candidate Knowledge and Teaching Effectiveness

Confidence in teacher preparation programs to produce effective teachers has ebbed and flowed through the years and more recently these programs have been cast within a shadow of doubt. Enveloped within the demands for highly effective teachers and higher student achievement is the reality of how elusive the task of preparing today's teachers can be.

Historically, there have always been those critical of how teachers are trained, but reports issued during the 1980s (Carnegie Task Force on Teaching as a Profession, 1986; Holmes Group, 1986; National Commission for Excellence in Teacher Education, 1985; National Consortium for

Educational Excellence, 1985) were replete with changes and recommendations. As a result of these reports, teacher education programs emphasized the importance of teacher candidate knowledge and began requiring teacher candidates to demonstrate mastery of subject content knowledge with a passing score on admittance tests such as the PPST (Byrnes, Kiger, & Shechtman, 2003).

By 1996, the National Commission on Teaching and America's Future (NCTAF) issued the report entitled *What Matters Most* highlighting the perpetual problems that plague teacher education programs by stating that "Key elements of teacher learning are disconnected from each other. Coursework is separate from practice teaching; professional skills are segmented into separate courses; faculties in the arts and sciences are insulated from education professors. Would be teachers are left to their own devices to put it all together" (p. 32). This report suggested there was much more involved in training a teacher than just ensuring that teacher candidates had mastery of subject content knowledge. Shulman (1986) maintained that there is an expert teacher knowledge base beyond content knowledge that all preservice teachers must acquire. He explained it as follows:

The goal of teacher education is not to indoctrinate or train teachers to behave in prescribed ways, but to educate teachers to reason soundly about their teaching as well as to perform skillfully. Teaching is both effective and normative; it is concerned with both means and ends. Processes of reasoning underlie both. The knowledge base must therefore deal with the purposes of education as well as the methods and strategies of teaching (p. 13).

In 2003, Feiman-Nemser further explained the issues, concerns, and problems related to training teachers and teacher knowledge:

“Before novices begin teaching, they go through an initial phase of learning. In a preservice program, they can acquire subject-matter knowledge, study the learning process and students’ cultural backgrounds, and acquire a beginning repertoire of approaches to planning, instruction, and assessment. But we misrepresent the process of learning to teach when we consider new teachers as finished products...” (p. 26).

The truth of the matter is that when teachers leave teacher preparation programs, regardless of how much subject content knowledge they can demonstrate, they are far from ready to teach in ways that are expected and/or required of them (Feiman-Nemser, 2003).

Traditionally, novice teachers are given the heaviest workloads and the most challenging classes, receive minimal support from administration, often struggle to deal with student discipline problems, feel a lack of influence on school policy, and receive low compensation. Moreover, Darling-Hammond and Cobb (1996) emphasized the fact that preservice teachers are expected to perform responsibilities that were previously held by experts and specialists leading to higher levels of mastery than ever before for teacher candidates today further complicating the types of training teacher candidates must receive and the types of knowledge they must have to be effective in today’s classroom.

Standardized Test Scores as a Screening Device

To date, there is insufficient evidence to support the notion that standardized tests are a suitable screening device when making determinations about which teacher candidates are capable of becoming highly effective teachers. Cobb, Shaw, Millard, and Bomotti (1999) charged that the “relevance-utility of [these basic skills teacher competency tests] ...has nothing to do with the rational screening-out function it purports to have and much more to do with the love affair and comfort the public seems to have with ... tests, no matter how functional they

are” (p. 172). Gitomer, Latham, and Ziomek (1999) added that “passing a Praxis test does not guarantee that an individual will become a satisfactory teacher” (p. 13).

In actuality, previous studies have demonstrated that there is not a strong correlation between standardized test scores and effective teaching as measured by student teaching grades and supervisor feedback (Riggs & Riggs, 1990; Baskin, Ross, & Smith, 1996). Furthermore, Dybdahl, Shaw, and Edwards (1997) found no significant relationships between basic competency tests and teaching success and suggested that “the arguments for testing that assumed increased teacher quality need to be revisited” (p. 252).

In 2001, Lawrence and Crehan studied the correlation between ACT scores, undergraduate grade point average (GPA), and PPST scores and found there to be none. Additionally, Byrnes, Kiger, and Shechtman (2003) analyzed the GPA of 68 student teachers as well as their ACT scores (another indicator of teacher candidate knowledge) to determine their predictive validity of high student teaching scores. Interestingly, Byrnes and colleagues found that the GPA did not predict high university supervisor and cooperating teacher ratings but that high ACT scores actually had a negative association. Moreover, Casey (2005) noted that a student’s incoming GPA, a student profile, or a combination of the two, led to no significant correlation with student teaching success. Current evidence suggests that there is little to no correlation between teacher candidate knowledge and student teaching success.

Tracking Teacher Candidates into the Classroom

Much of the research reviewing the practices of how teacher preparation programs make teacher candidate selections has been limited because the analysis has been situated around the teacher candidate experiences within the teacher preparation program. No studies were located that followed teacher candidates into the classroom to determine the accuracy of the selection

criteria in predicting effective teachers. Anders, Hoffman, and Duffy (2000) concluded that “more longitudinal studies of program effectiveness” (p. 278) are needed. Tracking preservice teachers as they enter the teaching field is complex and arduous, but this type of information is needed in order to determine if current selection criteria used by teacher education programs is sufficient. The real test of the skills and abilities of teacher candidates is not at the completion of their training, but when these teachers begin to teach varied and diverse learners within unanticipated circumstances and contexts. This study was designed to track these individuals as they begin teaching in a classroom of their own in order to make more knowledgeable decisions about which teacher candidate selection criteria to use.

Research Methodology

Research Design and Participant Selection

This study compared each participant’s standardized test score on the PRAXIS II standardized test with their performance on the CLASS observation instrument to explore the long term effectiveness of using PRAXIS II as teacher candidate selection criteria. As this was a blind study, the researcher did not know the PRAXIS II test scores of participants. Therefore two lists containing names of novice teachers who had just recently graduated from a teacher preparation program in a mid-sized western public university was given to the researcher. The researcher contacted two teachers from each list until a total of four teachers agreed to participate in the study. Two teachers received low, but passing scores on the PRAXIS II while the other two received high passing scores on the PRAXIS II. The researcher did not know which list contained low scoring teachers or high scoring teachers. Contact with teacher candidates was made over the phone inviting their participation in the study.

Sample

The four participants in this pilot study sample were novice teachers that had just received their teaching license to teach elementary school. Three of the participants completed a traditional teacher preparation program with one of the participants completing an alternate route to license issued by the state office of education in the state in which this study was situated. All participants had received at least a minimum passing score of 160 on the PRAXIS II standardized test as required by the state office of education. All four novice teachers were teaching in schools that were located no more than 70 miles from each other within the same state. Interestingly enough, all of the teachers either completed their student teaching assignment in the same school or school district where they eventually received employment. This suggests that teachers were familiar with the school culture and student population prior to their employment.

Teacher A secured employment in a school where the student population numbered 558 students. Of these students, 2% of the students were African American, 1% was American Indian, 1% was Asian, 8% were Hispanic, and 89% of the students were White. The percentage of students receiving free lunch was 17%, while 13% were receiving a reduced lunch. Teacher A was White, female, twenty-one years old, and completed her training at a traditional teacher preparation program.

Teacher B secured employment in a school where the student population numbered 528 students. Of these students, 1% of the students were African American, 1% was Asian, 6% were Hispanic, and 92% of the students were White. The percentage of students receiving free lunch was 19%, while 17% were receiving a reduced lunch. Teacher B was White, female, twenty-one years old, and completed her training at a traditional teacher preparation program.

Teacher C secured employment in a school where the student population numbered 321

students. Of these students, 1% was African American, 1% was American Indian, 2% were Asian, 33% were Hispanic, and 63% of the students were White. The percentage of students receiving free lunch was 55%, while 7% were receiving a reduced lunch. Teacher C was White, female, twenty-two years old, and completed her training at a traditional teacher preparation program.

Teacher D secured employment in a school where the student population numbered 338 students. Of these students, 2% were American Indian, 1% was Asian, 2% were Hispanic, and 95% of the students were White. The percentage of students receiving free lunch was 11%, while 9% were receiving a reduced lunch. Teacher D was White, female, thirty-six years old, and completed her training through an Alternate Route to Licensure offered through the state office and within the same state as the other novice teachers in this sample.

Instrumentation

PRAXIS II Test. According to the Educational Testing Service (ETS, 2009), the PRAXIS II test is a professional assessment for prospective teachers. It is designed to assess the qualifications of teacher candidates through assessment of professional knowledge of teaching and subject matter content in the areas in which teaching certificates are issued and administered. PRAXIS II attempts to measure content knowledge of subjects that future teachers will teach in the classroom. This subject assessment includes multiple-choice as well as constructed-response test questions. The PRAXIS II test is administered at testing centers throughout the United States and is administered with paper and No. 2 pencil. In the state where this study is situated, a passing score on the PRAXIS II is a score of 160.

CLASS Observation Instrument. For this pilot study, the CLASS (see Pianta, LaParo, & Hamre, 2008) instrument provided a means to determine the quality and effectiveness of

classrooms where sample participants were teaching. Teacher and student interactions were measured during multiple classroom observations. It is important to point out that classroom quality was measured through these observations and not teacher quality. As Graue, Rauscher, and Sherfinski (2009) explained, teacher quality is not measureable because of limitations such as time constraints, lack of adequate support staff, space limitations, and availability of funding and materials. An observation using the CLASS instrument focuses on the experiences of the average child in the classroom (Pianta, et al., 2008). The CLASS measures evidences of emotional support, classroom organization, and instructional support. La Paro, Pianta, and Stuhlman (2004) explained that “unlike previous research instrument, the CLASS provides a mechanism by which classroom practices can be gauged and improved systematically” (p. 420). Teachers receiving high ratings on the CLASS have been shown to improve social and academic outcomes for children and these findings have been replicated in numerous studies (Graue, et al., 2009).

La Paro, et al. (2004) explained that constructs forming the scales on the CLASS observation instrument were adapted from the National Institute of Child Health and Human Development, Early Child Care Research: Observational Record of the Caregiving Environment (NICHD, ECCRN, 1996), the Classroom Observation System (NICHD, ECCRN, 2002; Pianta, LaParo, Payne, Cox, & Bradley, 2002), as well as a review of the literature relating to topics such as teacher-student interactions, child care quality, and teacher education. The three main constructs of CLASS are emotional climate, management, and instructional support. The emotional climate domain attempts to measure the feeling tone residing in the classroom and the climate exhibited by both the teacher and students. Questions this domain seeks to answer are as follows: Are the needs of students being met? Is there flexibility to accommodate these needs?

What are the relationships like in the classroom? The dimensions included within this domain are positive climate, negative climate, regard for student perspectives, and teacher sensitivity.

The classroom management domain reflects the manner in which the classroom is organized, how learning takes place, and the amount of time spent on activities. Questions this domain seeks to answer are as follows: Is learning time maximized? Are a variety of modalities represented? Is behavior management proactive? Are their clear behavioral expectations? The dimensions of this domain include behavior management, productivity, and instructional learning formats.

The instructional support domain refers to the types of questioning and feedback provided to the student. Questions this domain seeks to answer are as follows: How is feedback provided? Is there repetition and extension? Is scaffolding evident? Are higher order thinking skills incorporated? The dimensions of this domain include concept development, quality of feedback, and language modeling. More information and descriptions of the domains and dimensions included within the CLASS are available in the CLASS Manual (Pianta, et al., 2008).

The CLASS was selected for this study for two main reasons. First, it has demonstrated a strong empirical record and has been used in over 3,000 classrooms (La Paro, et al., 2004); and second, it has strong psychometric properties. The validity of CLASS was determined by comparing ratings of CLASS with two other observational instruments (La Paro, et al., 2004) and through the use of confirmatory factor analysis. A confirmatory factor analysis determined three main domains being measured by CLASS (Graue, et al., 2009). Graue, et al. (2009) explain that reliability for the CLASS instrument was assessed across numerous observations cycles with reliability ratings in the moderate to high range ($r = .68 - .97$). The evidence of its reliability and validity has been detailed in the CLASS manual (Pianta, et al., 2008).

Observers rate classroom experiences on a Likert scale ranging from 1 to 7. A score of 1 or 2 is considered a low rating. A score of 3, 4, or 5 is considered a moderate rating, and a score of 6 or 7 is considered a high rating. In order for a rater to be considered reliable on the CLASS instrument, he or she must be within one point of the standard on 80% of the responses. The researcher in this study completed three days of training on the CLASS instrument and was deemed reliable upon completion.

Data Collection and Analysis

To evaluate its predictive validity, the PRAXIS II test was administered to all participants prior to acceptance into a teacher preparation program or teacher licensure. Upon completion of the teacher training, participants were tracked into the classroom and observed on three separate occasions – once at the beginning of the school, once at the middle of the school, and one towards the end of the school year. Three 30-minute cycles of observation/recording took place during each of the three visits totaling one and a half to two hour visits in each of the four classrooms. Observational visits were scheduled to take place during instructional time so as to capture the bulk of instruction each morning.

Results

PRAXIS II. As this was a blind study, the researcher did not know the PRAXIS II scores of any of the participants. After observing in the classroom over the course of a school year, the results of the PRAXIS II scores were made available for comparison. It was revealed at the conclusion of the study that Teacher A and Teacher B had received a low, but passing score on the required PRAXIS II test. Teachers C and D had received a high passing score on the required PRAXIS II test. These findings are important when analyzing the data collected from the CLASS observation instrument.

CLASS Observational Data. A mean score averaging all cycle observations for each classroom visit is reported for each participant in Table 1. Teachers A and B reported the highest overall CLASS scores while Teachers C and D reported the lowest mean scores on the CLASS instrument. Results indicated that the two teachers receiving the higher scores on the Praxis II test received the lowest scores on the CLASS instrument while the two teachers receiving the lower scores on the PRAXIS II test received the highest scores on the CLASS instrument.

Discussion

The first research question asked, “Do teacher candidates that perform well on standardized admittance tests, such as the PRAXIS II, go on to interact effectively with students and create high quality classrooms? The findings in this study suggest that standardized test scores may not be the most accurate prediction of teacher candidates that will go on to interact effectively with students to increase academic outcomes. Though there were no studies located in the research that analyzed the predictive value of standardized test scores to determine effectiveness of teacher candidates after they had already secured employment teaching, similar findings were found in studies analyzing the predictive value of standardized test scores or other measures of knowledge (such as GPA) on teacher candidates at the conclusion of their training program. It was determined in these studies that standardized test scores or GPA may not be as effective at predicting effective teachers when compared to other data such as leadership, interpersonal, and communication skills (e.g. see Byrnes, Kiger, & Shechtman, 2003; Casey, 2005; Lawrence & Crehan, 2001). These findings suggest a need for teacher preparation programs to reconsider the selection criteria they utilize to admit teacher candidates into their programs. Though teacher candidate knowledge, as indicated by GPA and test scores, may be important, its long term significance may be limited.

One important reason for reconsidering the current teacher candidate selection criteria is the consideration of the teacher candidates that are not selected as potential candidates or admitted into teacher preparation programs. Who are these individuals? What do they have to offer that candidates traditionally accepted into teacher preparation programs do not? It should be noted that minority teacher candidates are often denied admittance into teacher preparation programs based on standardized test scores such as the PRAXIS I/PPST, when there is sufficient evidence suggesting that the selection criteria is not adequate in determining which teacher candidates will go on to teach effectively (Bennett, McWhorter, & Kuykendall, 2006).

Garcia (1986) pointed out the need to reconsider the selection process for teacher candidates to ensure that minority teacher candidates have equitable chances of being accepted. With the documented achievement gap, more needs to be done to attract teacher candidates that fit minority categories in order for the teaching population to more closely represent the student population. Garcia further suggested that decisions regarding teacher candidate selection should consider additional skills and abilities of teacher candidates such as language and cultural knowledge along with other nontraditional considerations. Garcia (1986) quoted G. P. Smith as saying, “Clearly any professional practice that excludes disproportionate numbers of minorities represents neither excellence nor equity. If this nation is considered at risk now, a decade of willful elimination of minority teachers will result in a nation lost” (p. 352-353). More than two decades have passed since this statement, and the profession may still be considered neglectful if it refuses to consider the implications of current teacher candidate selection criteria. Understanding more about the long term significance of teacher candidate selection criteria may be one way to improve teacher quality and student achievement generally.

The second research question asked, “Does the CLASS instrument serve as valid and reliable way to determine teacher effectiveness after teacher candidates have begun teaching in a classroom?” Previous studies have relied on cooperating teacher and/or university supervisor grades as well as comments to determine teacher effectiveness. Though these methods may hold value, they are limited in their ability to extensively determine teacher effectiveness beyond the experiences within the teacher preparation program. This study demonstrates how valuable reliable and valid classroom observations can be in determining teacher effectiveness. The results of this study recommend the use of the CLASS observation instrument in determining novice teacher effectiveness and classroom quality to aid in making decisions about future teacher candidates.

Research Limitations

As there are indeed limitations to the findings of this study, a few cautions are in order. The first limitation of this research study is the sample size. As it is a pilot study, the use of four participants does not lend itself to generalizability of the quantitative data to a larger population of teacher candidates and novice teachers. However, the information gathered from this pilot study would be useful in the design of a larger scale study implementing a larger sample size.

Secondly, this study does not incorporate student data to provide further insight into the effectiveness of teachers and classrooms. Student achievement data would provide additional and meaningful information about teacher effectiveness. Additional insight from students may also prove to be noteworthy in gathering data into how teachers are doing. For example, Manning and Payne (1994) determined that students in elementary grades are able to determine differences between effective and ineffective teachers. Additional information, including student data and

student insight could provide further understanding of how teachers selected into teacher training programs actually perform once hired to teach full time.

References

- Anders, P. L., Hoffman, J. V., & Duffy, G. G. (2000). Teaching teachers to teach reading: Paradigm shifts, persistent problems, and challenges. In M. L. Kamil, P.B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 719-742). Mahwah, NJ: Erlbaum.
- Baskin, M. K., Ross, S. M., & Smith, D. L. (1996). Selecting successful teachers: The predictive validity of the urban teacher selection interview. *The Teacher Educator*, 32, 1-21.
- Bennett, C., McWhorter, L., & Kuykendall, J. (2006). Will I ever teach? Latino and African American students' perspectives on Praxis I. *American Educational Research Association Journal*, 43(3), 531-575.
- Byrnes, D., Kiger, G., & Shechtman, Z. (2003). Evaluating the use of group interviews to select students into teacher-education programs. *Journal of Teacher Education*, 54(2), 163-172.
- Carnegie Task Force on Teaching as a Profession (1986). *A nation prepared: Teachers for the 21st century*. Hyattsville, MD: Carnegie Forum on Education and Economy.
- Casey, C. E. (2005). *The relationships among teacher education admission criteria, practice teaching, and teacher candidate preparedness*. Unpublished dissertation, Ontario Institute for Studies in Education of the University of Toronto.
- Casey, C. E., & Childs, R. (2007). Teacher education program admissions criteria and what beginning teachers need to know to be successful teachers. *Canadian Journal of Educational Administration and Policy*, 67, 1-18.
- Cobb, B. R., Shaw, R., Millard, M., & Bomotti, S. (1999). An examination of Colorado's teaching licensure testing. *The Journal of Education Research*, 92, 165-175.
- Darling-Hammond, L., & Cobb, V. L. (1996). The changing context of teacher education. *The Teacher Educator's Handbook*, (pp.14-62). San Francisco, CA: Jossey-Bass.
- Dybdahl, C. S., Shaw, D. G., & Edwards, D. (1997). Teacher testing: Reason or rhetoric. *Journal of Research and Development in Education*, 20, 248-254.
- Educational Testing Service. (2009). *The Praxis series: Teacher licensure and certification*. Retrieved September 10, 2009, from www.ets.org/praxis/.
- Feiman-Nemser, S. (2003). What new teachers need to learn. *Educational Leadership*, 60(8), 25-29.
- Garcia, P. (1986). The impact of national testing on ethnic minorities: With proposed solutions. *Journal of Negro Education*, 55(3) 347-357.

- Gitomer, D. H., Latham, A. S. & Ziomek, R. (1994). *The academic quality of prospective teachers: The impact of admissions and licensure testing*. Retrieved from the Educational Testing Service, The Teaching and Learning Division, Research and Data Analysis Group Web Site: [http:// www.teaching and learning.org](http://www.teachingandlearning.org).
- Graue, E. Rauscher, E. & Sherfinski, M. (2009). The synergy of class size reduction and classroom quality. *The Elementary School Journal*, 110(2) 178-201.
- Holmes Group. (1986). *Tomorrow's teachers: A report of the Holmes Group*. East Lansing, MI: Author.
- La Paro, K., Pianta, R., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal*, 104(5) 409-423.
- Lawrence, A. & Crehan, K. D. (2001, April). A study on the validity evidence of the Pre-Professional Skills Test (PPST) as a screening device for entrance into teacher education programs. Paper presented at the annual meeting of the National Council on measurement in Education, Seattle. (ERIC Document Reproduction Service No. ED452259).
- Manning, B. H., & Payne, B. D. (1984). Student teacher personality as a variable in teacher education. *The Teacher Educator*, 20(20), 2-12.
- National Commission for Excellence in Teacher Education. (1985). *A call for change in teacher education*. Washington, D.C.: Author.
- National Commission on Teaching and America's Future (NCTAF). (1996). *What matters most: Teaching for America's future*. New York: Author.
- National Consortium for Educational Excellence. (1985). *An agenda for educational research: A view from the firing line*. Nashville, TN: Vanderbilt University.
- National Institute of Child Health and Human Development, Early Child Care Research Network. (1996). Characteristics of infant child care: Factors contributing to positive care-giving. *Early Childhood Research Quarterly*, 11(3) 269-306.
- National Institute of Child Health and Human Development, Early Child Care Research Network. (2002). The relation of global first grade classroom environment to structural classroom features and teacher and student behaviors. *Elementary School Journal*, 102(5) 367-387.
- No Child Left Behind Act*, P. L. 107-110, 107th Congress. 1st Session, 2001.
- Petersen, G. J., & Speaker, K.M. (1996, February). *Bottom half of the pool: Who is admitted to teacher education?* A paper presented at the 19th Annual Meeting of the Eastern Educational Research Association, Boston, MA.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). Classroom Assessment Scoring System (CLASS). Baltimore, MD: Brookes.

Pianta, R. C., La Paro, K., M. Payne, C., Cox, M. J. & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school outcomes. *Elementary School Journal*, 102(3) 225-238.

Riggs, I. M., & Riggs, M. L. (1990, February). A test of the validity of selected predictors of student success in a teacher education program. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Chicago. (ERIC Document Reproduction Service No. ED324324).

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57 (1), 1-22.

Squires, G. (1999). *Teaching as professional discipline*. London: Falmer.

Stronge, J. H. (2002). *Qualities of effective teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.

INTASC Standards. (2009). *The INTASC standards*. Retrieved December 10, 2009 from <http://www.wresa.org/Pbl/The%20INTASC%20Standards%20overheads.htm>

Turner-Bisset, R. (2001). *Expert teaching*. London: David Fulton Publishers.