The Relationship between the Interactive Computer Interview System and the *Praxis II* Principles of Learning and Teaching Test

By

Sharon M. Pruett

Submitted to the graduate degree program of Educational Leadership and Policy Studies and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Education.

_____
Chairperson, Dr. Howard Ebmeier

_____
Dr. Susan Gay

_____
Dr. Perry Perkins

_____
Dr. Suzanne Rice

_____
Dr. Argun Saatcioglu

Date Defended: February 28, 2012

The Dissertation Committee for Sharon M. Pruett
certifies that this is the approved version of the following dissertation:

The Relationship between the Interactive Computer Interview System and the *Praxis II* Principles of Learning and Teaching Test

_____

Chairperson Dr. Howard Ebmeier

Date approved: April 13, 2012

Abstract

The objective of this study was to compare the relationships between the subtests of the Interactive Computer Interview System and the ETS *Praxis II* Principles of Learning and Teaching examination. In particular, this study compares scores on the ICIS instrument subtests to those gathered from the same classroom teachers on the *Praxis II* PLT examination. Data collected from fifty-eight teachers, thirty-three secondary and twenty-five elementary teachers, was used in this study.

Results indicate the strongest correlations, $\alpha = .05$, exist between the *Praxis II* PLT exams and the ICIS Knowledge of Content and Knowledge of Teaching subtests for the secondary teachers and the ICIS Knowledge of Teaching for all teachers. Correlations also existed at the significance level, $\alpha = .10$, for all teachers for the ICIS subtests Knowledge of Content, Knowledge of Students, and Working with Others and for the secondary teachers for Knowledge of Students and Working with Others. A correlation was indicated for the elementary teachers at the significance level, $\alpha = .20$, between the ICIS Knowledge of Teaching and the *Praxis II* PLT exam.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction

Teacher selection provides the greatest impact on student performance. The employment screening instrument is second only to the application blank in frequency of use for schools to identify exemplary candidates to fill teaching positions (Schmidt, 1999). The interview questions provide a great deal of information for the selection committee or person responsible for the hiring process. Indeed, recent research indicates that employment interviews have higher validity for predicting job performance than previously believed (McDaniel, Whetzel, Schmidt, & Maurer, 1994). In the history of published research dating back to Scott in 1915, few conclusions have been more widely supported than the idea that structuring the interview enhances reliability and validity of the selection process. "Using the same questions may be the most basic way to convert the interview from a conversation into a scientific measurement" which helps allow appropriate comparisons between interviewees (Campion, Palmer, & Campion, 1997, p. 663).

During the interview process, it has been found that nonverbal signals are sometimes more powerful than the spoken word (Ambady & Rosenthal, 1992). Some people naturally smile more, maintain more eye contact, adopt different postures and body orientations, and some people are more physically attractive (Burnett & Motowidlo, 1998). Argyle (1988) argued that nonverbal behavior is less controllable than verbal behavior; it is unintended and unconscious. Birdwhistell (1970) suggested that only 30 – 35% of the social meaning of a conversation is communicated through the actual words spoken. Communication experts have long supported the idea that the use of nonverbal behavior as a second channel of communication, in addition to verbal content, enhances the complexity of the message (Burnett & Motowidlo, 1998). The structured interview provides performance ratings based on the content of answers helping to eliminate biases that may result from nonverbal characteristics.

Structured interviews have been found to have higher levels of inter-rater reliability than unstructured interviews (Conway, Jako, & Goodman, 1995; McDaniel et al., 1994). Structured interviews provide consistency during the interview process; consequently, the greater ability to legally defend these interviews makes them an attractive alternative to unstructured interviews (Williamson, Campion, Malos, Roehling, & Campion, 1997). Jose Cortina, et al. (2000) reports some researchers have explained the effectiveness of structured interviews in terms of both method and content including standardization, multiple interviewers, content validity, and training of interviewers.

In today's era of high stakes testing and teacher accountability, it is increasingly important for the success of all school districts that they hire only the teachers most knowledgeable in their subject fields. Indeed, the requirements of the federal mandates of No Child Left Behind include having "highly qualified teachers" in all classrooms. A teacher selection interview process that helps identify the most highly qualified teachers is critical to this goal. This study evaluates the relationship between the Interactive Computer Interview System, a commercial, structured interview protocol, used during the teacher selection process and scores on the *Praxis II* Principles of Learning and Teaching test that is required for state licensure. A number of different validity studies have been completed on the Interactive Computer Interview System, known as the ICIS instrument, but none have been identified that verify the validity of the instrument in relationship to external instruments such as the *Praxis II* Principles of Learning and Teaching, known as the PLT, examination.

This study compares data obtained from interviews using the ICIS instrument with the scores earned on the *Praxis II* PLT examination for these same individuals. When a correlation is shown between the ICIS subtests and the *Praxis II* PLT scores achieved, the information

gained during the ICIS interview can be viewed as a validation of the relationship. As school districts strive to hire the best teachers available, a correlation between the ICIS interview and *Praxis II* PLT scores will validate the relationship between these two instruments and help identify the most highly qualified teachers. *Praxis II* PLT scores are typically not available for those applicants being interviewed. Yet successful completion of that test is an indication of teaching ability and is required for state licensure. A relationship between the ICIS subtests and the *Praxis II* PLT tests will validate the use of the Interactive Computer Interview System as an effective interview instrument.

Chapter Two: Review of the Literature

2.1 Components of the Structured Interview Defined

Michael Campion, David Palmer, and James Campion (1997) have reviewed, evaluated, and summarized the many ways interviews can be structured in "A Review of Structure in the Selection Interview." They define structure "as any enhancement of the interview that is intended to increase psychometric properties by increasing standardization or otherwise assisting the interviewer in determining what questions to ask or how to evaluate responses" (Campion, Palmer, & Campion, 1997, p. 656). They have defined 15 components of structure that are divided into two categories: those that influence the content of the interview or the nature of the information elicited, and those that influence the evaluation process or the judgment of the information elicited. When defining the structured interview, the 15 components of structure identified include the following:

1. *Base questions on a job analysis*. Integration of critical incidents is the most common of the job analysis methods used to develop structured interviews. These provide ideas for interesting and job-related questions. The development of questions is part of the art of structured interviewing.
2. *Ask exact same questions of each candidate*. The most basic component of structure is standardization of questioning. The highest level of maintaining this component requires that the identical question be asked of each candidate in the same order.
3. *Limit prompting, follow-up questioning, and elaboration on questions.* The use of prompts and follow-up questions is a primary means by which interviewers might bias information gathering.

4. *Use better types of questions*. Question type can refer to either how the question is asked or its content. Situational questions are used to gain responses from candidates regarding what they would do in hypothetical situations. Past behavior questions focus on past jobs as those relate to the position they are seeking. Background questions are asked to gain knowledge on work experience, education, and other qualifications. Job knowledge questions seek information to document or describe job knowledge.
5. *Use longer interview or larger number of questions.* Within reasonable limits, longer interviews are more structured because they obtain a larger amount of information.
6. *Control ancillary information.* Uncontrolled ancillary information, including the use of application forms, resumes, test scores, recommendations, previous interviews, or transcripts can confound the interpretation of the value of the interview if not available for all candidates.
7. *Do not allow questions from candidates until after the interview*. Uncontrolled questions from candidates reduce standardization by changing the interview content in unpredictable ways. Time can be allowed for questions outside the interview.
8. *Rate each answer or use multiple scales*. Rating each answer provides the highest level of structure because judgments are linked to specific responses with scales tailored to each question. Multiple ratings at the end are less structured because they are based on answers to multiple questions or on the entire interview.
9. *Use detailed anchored rating scales*. Anchored rating scales use behavioral examples to illustrate scale points in order to reduce ambiguity and semantic differences possible with adjective anchors and are presumed to enhance objectivity.

10. *Take detailed notes*. Note-taking may enhance structure because it reduces memory decay.
11. *Use multiple interviewers*. Sharing perceptions may help interviewers become aware of irrelevant inferences that are not job related, and multiple interviewers may reduce the impact of biases.
12. *Use the same interviewer(s) across all candidates*. Diverse interviewers may evaluate answers differently reducing the reliability of the interview process.
13. *Do not discuss candidates or answers between interviews*. Discussing candidates may lead to irrelevant information entering the evaluation process, as well as instrumentation effects such as changing standards between interviews.
14. *Provide extensive interviewing training*. Trained interviewers should be able to elicit content and evaluate consistently, thus improving interview reliability.
15. *Use statistical rather than clinical prediction*. Most interviews combine ratings into total scores based on statistical rules because different interviewers weigh information differently.

Campion, Palmer, and Campion (1997) found that based upon the supporting evidence, good rationale for using completely unstructured interviews does not exist. While this review identified 15 components of the structured interview, there may be more that were not identified. Any interview may be enhanced by using at least some of these components. These authors recommend that interviews be structured in all possible ways within the limitations that may exist.

2.2 The Need for the Structured Interview

At a fundamental level, the interview is a social interaction between the interviewer and the applicant (Posthuma, Morgeson, & Campion, 2002). Applicant fit refers to how it is perceived that the applicant's characteristics will match the job or organization. Individuals may fit because they are exactly the same as other members of the work group, or they may possess the needed skills and abilities that a work group is lacking. When basing hiring decisions on how it is perceived that the candidate will fit the organization, unconscious bias can affect the selection process. Bias causes interviewers to look at a situation from a specific and sometimes limiting perspective and is based on past personal experience. We all possess some forms of bias, which is part of normal thought processes. We are unaware of possessing unconscious bias, so its manifestation is unintentional but can impact important hiring decisions. Project Implicit—a study by Harvard University, the University of Washington, and the University of Virginia—found that "our brains make associations based on experiences and information we receive, whether we consciously agree with those associations or not" (Savini, 2010, p. 66). Prompts and follow-up questions are primary means by which interviewers can bias information gathering (Dipboye, 1994). A structured interview protocol provides consistent questioning techniques and evaluation procedures that help avoid the impact of unconscious bias. Structured interviews seem to be less vulnerable to interviewer biases (Dipboye & Gaugler, 1993) and enhance reliability and validity (Campion, Palmer, & Campion, 1997).

The most important property of a personnel assessment method is predictive validity: the ability to predict future job performance, job-related learning, and other criteria (Schmidt & Hunter, 1998). The benefits of using hiring methods with increased predictive validity include substantial increases in employee performance as measured in percentage increases in output and

increased learning of job-related skills. The most well-known finding from 85 years of research is that for hiring employees without previous experience in the job, the most valid predictor of future performance is general mental ability (GMA), intelligence or general cognitive ability. This research encompassed unskilled and semi-skilled workers through skilled workers and managerial and professional employees (Schmidt & Hunter, 1998).

2.3 Structured versus Unstructured Interviews

McDaniel, Whetzel, Schmidt, and Maurer (1994) report in their meta-analytic review, *The Validity of Employment Interviews: A Comprehensive Review and Meta-Analysis*, that unstructured interviews gather applicant information in a less systematic manner than do structured interviews. Unstructured interviews often do not specify the questions in advance of the interview and do not have formalized scoring guides to judge the responses. Persons interviewing for the same position are not asked the same questions. The evaluations of the interviews can be quite different due to the lack of a standard interview protocol, and consequently are less reliable than structured interviews. These authors hypothesized higher validities for structured interviews and believe they are better for obtaining a wide range of applicants' information (McDaniel, Whetzel, Schmidt, & Maurer, 1994).

Unstructured interviews have a tendency to be conversational with both parties asking questions. Reducing the conversational nature of an interview will reduce the ongoing interactions between the interviewer and the candidate. This helps prevent the interviewer from using the candidate's questions to make judgments about the candidate, and it helps prevent candidates from asking questions and using the information to shape their answers (Beatty, 1986). Reducing this conversational nature may decrease contamination of the interview, but it has to be acknowledged that relevant information may not emerge from the interview. When

prohibiting the unstructured interview negative user reactions may result because it restricts freedom and may even lead candidates to feel less positive about the interview experience (Campion, Palmer, & Campion, 1997).

2.4 The Impact of Cognitive Ability on Teacher Success

Research has shown that cognitive ability is among the components that best predict job performance (Barrick & Mount, 1991; Huffcutt & Arthur, 1994; Hunter & Hunter, 1984). Further research has led to the widespread belief that cognitive ability is the single best predictor of many aspects of job and training performance for a variety of jobs (Hunter & Hunter, 1984; Ree, Earles, & Teachout, 1994; Reilly & Chao, 1982; Schmidt, Ones, & Hunter, 1992). Research and review has shown that interview scores are related to cognitive ability test scores also. Campion and his colleagues (Campion, Campion, & Hudson, 1994; Campion, Pursell, & Brown, 1988) found consistent relationships between scores on structured interviews and scores on cognitive ability tests. Indeed, some researchers suggest that structured interviews are merely poor measures of cognitive ability (Campion et al, 1988; Wright et al, 1989).

Cortina, Goldstein, Payne, Davison, and Gilliland (2000) found from the results of their meta-analysis of the relationships between cognitive ability and interview scores that there is a weak relationship between scores on unstructured interviews and cognitive ability. The relationship between scores on structured interviews and cognitive ability was considerably higher (p = .253) than those involving interviews with less structure. Scores resulting from greater structure in the interview process when compared to measures of cognitive ability reveal that this relationship (p = .27) is even stronger than those involving interviews with less structure.

Cognitive ability and personality traits are the two most commonly studied individual difference predictors of job performance. Cognitive ability best predicts task performance (Borman & Motowidlo, 1993; Organ & Ryan, 1995). Motowidlo et al. (1997) found that cognitive ability leads to task habits, skills, and knowledge, which influence task performance. Cognitive ability has been linked to performance in all jobs (Ghiselli, 1973; Schmidt et al., 1981) and to managerial performance (Borman et al., 1991; Hunter, 1983; Gallatly et al., 1991) in particular. Bergman, Donovan, Drasgow, Overton, and Henning (2008) hypothesized that cognitive ability is a direct antecedent of leadership knowledge in their "Test of Motowidlo et al.'s (1997) Theory of Individual Differences in Task and Contextual Performance." Situational Judgment Tests were utilized and showed that cognitive ability was directly and indirectly (through leadership knowledge) related to leadership performance. Hunter (1986) provided a framework based on learning theory to explain why assessments of general cognitive ability predict task performance. General cognitive ability is needed for workers to learn how to do their jobs. Job knowledge allows workers to respond to situations involving job-related tasks appropriately. Cognitive ability should be related to job performance through this effect of job knowledge.

One major finding of Hunter and Schmidt (1992) shows that general mental ability, another term for cognitive ability, is a good predictor of job performance based upon the acquisition of job knowledge. The main reason more intelligent people have higher job performance is that they acquire job knowledge more rapidly and acquire more of it. Mental ability has its most important effect on job performance through job knowledge. "GMA tests and work sample measures have high validity. Of the predictors examined, two stand out as being both practical to use for most hiring and as having high composite validity: the

combination of a GMA test and an integrity test (composite validity of .65) and the combination of a GMA test and a structured interview (composite validity of .63)" (Schmidt & Hunter, 1998, p. 272). An integrity test uses questions to assess previous training and experience to measure behavioral consistency realizing "the best predictor of future performance is past performance" (Schmidt & Hunter, 1998, p. 268). Both combinations predict performance in job training programs effectively, as well as performance on the job, and both are less expensive to use than many others. Thus, both combinations are excellent choices. Employers must make hiring decisions, but they can choose methods that are deemed most effective to assist in this hiring process.

Chan and Schmitt (2002) found that cognitive ability best predicts task performance. Data from 160 civil service employees demonstrate a relationship between situational judgment ability and job performance, especially relative to other individual difference predictors including cognitive ability. Situational judgment tests can be used to predict job performance in human resource functions such as personnel selection. Many authors have maintained that personality variables may be the best predictors of contextual performance, whereas cognitive ability best predicts task performance (e.g., Borman & Motowidlo, 1993; Organ & Ryan, 1995). Consistent with previous research (Schmidt & Hunter, 1998), the authors found cognitive ability to be a valid predictor of core technical proficiency, or task performance ($r = .25$, $p < .05$) (Chan & Schmitt, 2002, p. 242).

Mount, Witt, and Barrick (2000) studied 376 employees in clerical jobs and reported that individuals with higher intelligence are more likely to produce higher quantity and quality of output. Job knowledge is required for effective performance because people who understand the job better are more effective at solving job-related problems. General mental ability predicts

three aspects of performance: task, contextual, and adaptive. The participants in this study were given the Wonderlic Personnel Test (WPT) which is psychologically equivalent to other known measures of cognitive ability. The correlation between the WPT and the Wechsler Adult Intelligence Scale (WAIS), a well-established measure of intelligence for adolescents and adults, is .93. This comparison confirms that the WPT is a construct valid measure of cognitive ability.

A number of authors have suggested that interviews and GMA are correlated (Harris, 1998; Hunter & Hirsch, 1987; Schmidt, 1988). Both primary studies as well as meta-analyses have shown a relationship between GMA and interviews. Delery, Wright, McArthur, and Andersen (1994) showed a large correlation between GMA and a situational interview wherein respondents must exercise judgment when responding to situations that occur in work settings. Campion, et al. (1994) found a similar result for both situational questions and behavior description questions that focus on past or future behaviors. Pulakos and Schmitt (1995) found a small correlation between GMA and a behavior description interview. A comprehensive meta-analysis of the relation between interview and GMA was conducted by Huffcutt, Roth, and McDaniel (1996); using interview content as a moderator, they found that the situational interview and behavior description interview correlated less with GMA than the conventional interview in which questions focus on a variety of constructs. The true average correlation for behavioral interviews was .24 and for conventional interviews was .41. The average value found by Huffcutt, et al. for low and medium structure level was .43 based on a total sample of 7,121 individuals and 27 primary studies. In the meta-analysis conducted by Salgado and Moscoso (2002), 17,971 individuals from 53 studies were analyzed. Salgodo's and Moscoso's findings support Huffcutt's findings, and they conclude that conventional interviews partially assess GMA. A possible explanation for this result is that individuals scoring higher on GMA tests

demonstrate greater ability for answering more effectively in conventional interviews that assess social skills, GMA, emotional stability, and extraversion. Conventional interviews also assess job experience, openness, agreeableness, and conscientiousness to a slightly lower degree (Salgado & Moscoso, 2002).

2.5 Commercial Teacher Interviews

"It is the function of a pre-employment assessment system to efficiently select teachers who have a high probability of success in the classroom. Since there are generally many more applicants than the district is able to hire, efficiency of selection is paramount" (Webster, 1988, p. 252). Nearly 2,000 school districts in the United States, or about 15%, use a commercially-designed teacher interview in their hiring decisions (Delli, 2001). Scott Metzger (2003) describes commercial teacher interviews as "proprietary, marketed instruments that schools can use to screen and hire teachers who demonstrate particular personal traits or affective qualities." These interviews "are scripted, structured, and criterion-scored—giving the impression of being more systematic and objective than unstructured interviews" (Metzger & Wu, 2003. p. 1). Unstructured interviews, on the other hand, are guided at the discretion of the interviewer. The two main commercial teacher interviews are Martin Haberman's Urban Teacher Selection ("Star" Teacher) Interview and Gallup's Teacher Perceiver Interview (TPI).

The Urban Teacher Selection Interview ("Star" Teacher) was created by Martin Haberman of the University of Wisconsin – Milwaukee (Metzer & Wu, 2003). This online tool asks candidates to answer 50 multiple-choice items with three possible responses, allowing 45 seconds per question. A summary sheet provides the candidate with a score and a High/Average/Low rating for each category rated. Haberman's structured interview was developed to select "successful" beginning teachers. His early publications referred to a different

number of themes that over time evolved into the 10 used in the online version. He originally identified 14 "functions" of successful teachers of the urban poor that represent "teaching behavior," beliefs or commitments that predispose teachers to act, and these have remained stable over time (Metzger & Wu, 2003).

The Gallup Organization's Teacher Perceiver Interview is the most widely used commercial teacher interview. Over 1,200 school districts have employed the TPI (Young & Delli, 2002). Selection Research, Inc. (SRI) produced the TPI in the 1970's, and in 1988, acquired the Gallup Organization and adopted its corporate name. The TPI is a face-to-face interview in which the administrator asks the applicant 63 open-ended prompts—60 items related to TPI's 12 themes, with 3 follow-up questions pertaining to research. Administrators must complete training and are approved to administer the TPI only after achieving at least 85% inter-rater reliability in item-by-item scoring with Gallup trainers (Metzger & Wu, 2003). TPI largely ignores subject matter knowledge, which is a serious problem for commercial interviews. In some ways TPI even penalizes candidates who talk about subject knowledge when giving their responses. Only six questions focus on subject matter and a point is given only if the candidate focuses on student growth, student feelings, or the importance of teaching. Content knowledge is not awarded with any points. The TPI does not demonstrate that a teacher possesses sufficient content knowledge to help students learn complex subject matter. While educational researchers may not agree that teacher subject-matter knowledge is linked to student achievement, a strong body of evidence exists suggesting that it is at least an important influence.

Gallup introduced the TeacherInsight Interview in 2002, a Web-based talent assessment system that asks applicants to respond to a series of statements using a 5-point Likert scale, to a range of multiple-choice questions, and to a number of open-ended questions (Metzger & Wu,

2003). Gallup shifted to the TeacherInsight system principally after 2005, and it is based upon more than 30 years of Gallup research into selecting and developing teachers. TeacherInsight takes about 40 minutes to complete online, and the program generates a report for the school that provides the applicant's percentile rankings of his/her predicted potential for teaching success. The Gallup website offers TeacherInsight as the next generation product to help schools hire the best teachers available. The system is marketed as a means to attract a larger pool of candidates like their best teachers using a centralized approach that requires less staff time. Gallup provides feedback to school personnel via their website, sorting the candidates into priority levels for possible subsequent interviews with either the longer TeacherInsight or the actual face-to-face TPI.

Gallup also began marketing the Urban TPI, perhaps in response to Haberman's Urban Teacher Selection Interview, in the late 1990s (Metzger & Wu, 2003). The Urban TPI is based on 11 themes that define who successful urban teachers are and how they act. The Urban TPI claims to identify the best urban teachers based upon a common series of consistently recurring patterns of thought, feeling, and behavior.

This research focuses on a third structured interview instrument, the Interactive Computer Interview System (ICIS), developed by Dr. Howard Ebmeier from the University of Kansas (AASPA, 2003). This instrument is designed for use in K – 12 schools and follows an extensive review of literature on teacher effectiveness, discussions with practitioners and their professional associates, and an examination of national commission reports. The ICIS instrument allows the interviewer to focus on evaluating the candidate's responses while the computer program tracks response patterns, provides questions based on these patterns, and constructs detailed summaries to report on various aspects of the interview.

To ensure content validity, questions used in the ICIS instrument were derived from two documents from national studies—"Teacher of the Future," the result of work of a national commission of school personnel officers and the "*Praxis III:* Classroom Performance Assessments," created by the Educational Testing Service (Ebmeier, Dillon, & Ng, 2007). These two national commissions had identified each concept measured as critical to teaching excellence. The questions in the Interactive Computer Interview System are grouped into four clusters—Working with Others, Knowledge of Teaching, Knowledge of the Content, and Knowledge of Students. Scoring rubrics for each question were established and an overall framework for effectiveness in each of the four clusters was constructed. A computer program was written to manage the questions and responses. A unique feature of this structured interview protocol is the program's ability to remember how the candidate answered previous questions and use that information to select future questions from the bank. A minimum number of questions is answered and continue until the applicant demonstrates stability in the area, then the program moves to questions in the next cluster. Reports are generated during the interview process and presented visually on the computer monitor. These reports are written to a file on the computer's hard drive accessible through Microsoft Word. One report displays the number of questions asked, the mean, and the standard deviation for each theme and a weighted average and standard deviation for all questions asked during the interview. The interviewer has the option of choosing the desired version of the interview—short, normal, or long.

When utilizing a structured interview, training and preparation of the interviewers are essential according to the American Association of School Personnel Administrators (2003). If several interviewers observe the same interview but score it differently, the results have little value. To have confidence in the results obtained during the ICIS screening, administrators must

know that the information is reliable and independent of the person who conducted the interview. The ICIS training program provides 140 video clips of teachers responding in varying degrees to the interview questions. The learner classifies the responses with the results being compared to the correct answers until a high level of mastery is obtained.

Questions contained in commercial, structured interviews are carefully researched for validity before selected. Districts can expect higher validity because only questions found to correlate with job performance are retained for inclusion in the interview protocol. The process of question selection may be susceptible to sampling errors; thus the possibility exists to cancel out the expectation for higher validity. The interview questions strive to obtain the candidate's first reaction and not a response that would be produced after reflection and thought. The applicant's initial reaction is hypothesized to be more valid; the absence of reflection does not allow applicants to try to guess the "correct" answer (Schmidt & Rader, 1999).

2.6 Testing Required for Teacher Licensure

Dating back to the 19th century, teachers were tested in basic subjects in many states (Angrist & Guryan, 2007). In most states the tests were graded and certificates issued at the county level. Following World War II the supply of needed teachers decreased. Because of this shortage of teachers, most states discontinued the use of required teacher testing and even accepted teachers with alternative certification. Some states resumed testing prospective teachers beginning in the 1960s to ensure that teachers met minimum standards for basic skills and subject knowledge.

The Educational Testing Service's National Teachers Examination has been widely used since 1998 to certify teachers. This series has become known as the *Praxis* tests and includes the *Praxis I*, also known as the Pre-Professional Skills test (PPST), used to screen applicants to

Schools of Education; the *Praxis II*, used to verify a required level of knowledge based upon each state's credentialing requirements; and the *Praxis III*, a series of classroom performance assessments. By 1999, 41 states required teacher licensure applicants to pass a standardized certification test (Angrist & Guryan, 2004). Although some cross-state reciprocity in the form of probationary and provisional licensing exists, most states relying on tests for certification require newly employed teachers to pass their own tests even when licensed in other states.

A 1986 report of the Carnegie Task Force on Education and a follow-up report released in 1996 called for the introduction of more centralized systems of certification for public school teachers (Angrist & Guryan, 2007). A policy of stricter and more centralized teacher licensing has also received support from the National Education Association and groups promoting education reform (Ballou & Podgursky, 2000). Occupational licensing may provide a signal of worker quality and helps to maintain quality standards when information about quality is imperfect. This is the stated rationale for government-imposed licensing requirements (Angrist & Guryan, 2007). Importantly, a study by Goldhaber and Brewer (2000) links student achievement with state teacher licensing and testing requirements.

Teacher quality is widely recognized as influencing student achievement and success in school (Blanton, Sindelar, & Correa, 2006). Based upon the analysis of a subset of teacher quality studies in which student outcomes were used as a dependent variable (Hess, 2001; Walsh, 2001), the No Child Left Behind Act of 2001 challenged traditional concepts of good teachers by emphasizing content mastery and verbal ability and downplaying the importance of pedagogy. Leading researchers (Berliner, 2005; Fenstermacher & Richardson, 2005) have focused on teacher quality and have defined it as encompassing two parts—(1) good teaching, meaning that the teacher meets the expectations using age-appropriate methods, upholds the standards of a

field of study, holds degrees, and possesses other attributes, and (2) effective or successful teaching, meaning the results of the teacher's actions on student learning and achievement.

As research addresses the complexities of teaching, instruction and understanding teacher quality continues to expand and change as well as focus on the effective dimensions of teacher quality. Accountability and performance standards are dominating the teacher-quality agenda, with accompanying changes in teacher education accreditation and teacher licensure, which are the major quality control mechanisms for the profession (Blanton, Sindelar, & Correa, 2006).

Gitomer and Latham (2000) sought to enrich the knowledge of teacher supply and the effects of teacher licensure testing. They analyzed academic and demographic information taken from nearly 200,000 candidates who had taken at least one teacher test from ETS's *Praxis Series* between 1994 and 1997. In their study, they used candidates' SAT and ACT scores as a proxy for their academic ability. Gitomer and Latham noted that if schooling is to be considered an academic enterprise, it is logical that teachers be drawn from among the more academically talented; academic ability is clearly a desirable trait in teachers. There is growing evidence that SAT and ACT tests demonstrate strong verbal ability of teachers. This evidence suggests a positive relationship between these tests and verbal ability (Ehrenberg & Brewer, 1995; Ferguson, 1998).

Gitomer and Latham found that those who were successful on *Praxis I* and *II* tests had substantially higher average SAT scores than those who failed. The higher the *Praxis* passing score, the higher the average SAT scores were of the passing population. Those who pass teacher testing requirements, on average, appear to be more academically able than those who do not. While admissions and licensure testing may be feared because of limiting the supply of teachers, testing does not appear to serve as a barrier to most teacher candidates. A relatively

high success rate exists with 3 out of 4 candidates seeking entrance to a college of education passing *Praxis I*. Nearly 9 out of 10 seeking teacher licensure pass the *Praxis II* PLT. Some organizations such as the Education Trust (1999) have proposed that states "raise the bar" by requiring teachers to meet more stringent testing guidelines than are currently in place in order to earn a license to teach. While higher passing standards would raise the academic profile of teachers, it would also reduce the pool of teacher candidates and limit the supply of teachers (Gitomer & Latham, 2000).

Those who qualify for licensure in an academic subject have comparable academic skills to other college graduates. However, representatives from the Education Trust believe that math teachers, for example, should be compared with professional mathematicians, not college graduates from other fields (Blair, 1999). The elite in any subject area are limited in number, and teaching demands a large supply of teachers within a given field. It seems far more realistic to seek a teaching force with reasonable academic skills and help them develop into excellent teachers who can prepare students for more in-depth pursuit in particular disciplines than it is to insist that only the academically elite be admitted into teaching (Gitomer & Latham, 2000).

The challenge of improving teaching is far more complex than simply adjusting licensing requirements. The study conducted by Gitomer and Latham (2000) suggests that within academic fields, licensed teachers as a whole have adequate educational skills given the limitation of only using standardized admissions test scores. Thus, the policy focus should be on finding ways to hire and retain these teachers in large enough numbers to supply the teachers needed for our schools' classrooms. Licensure is only one step in the quality control process. This study found that the areas of elementary, special education, and physical education appear to be drawing disproportionately high numbers of candidates with lesser academic ability.

Schools are communities of learning, and it is important that teachers working with students of all ages and in all subject areas have reasonably strong academic backgrounds. Elementary teachers who cannot share an appreciation of mathematics, science, or social studies with their students, and who cannot explain to students the importance of what they study will not be able to inspire excitement and motivation on the part of their students.

The data in the study by Gitomer and Latham (2000) provides compelling evidence that high standards will lead to a teacher population with higher mean academic ability. Teacher testing that includes high passing scores has a major payoff in terms of the mean academic skill of the prospective teacher pool. When adopting higher standards, aggressive efforts to support and enhance all candidates' knowledge and abilities must be included so they have a better chance of meeting those high standards. Data suggests that the act of raising passing standards will not be a "silver bullet solution for improving teacher quality" (Gitomer & Latham, 2000, p. 10). Testing with higher standards is a wave of the future for ensuring that teachers are academically able to meet the rising demands of the profession. The policy objective that motivates teacher testing is to identify and hire those most qualified to teach.

Laura Goe, a principal investigator for the federally-financed National Comprehensive Center for Teacher Quality, states, "A standardized test is very useful to take a snapshot of the inventory of knowledge and skills that a student has, but it's not as useful to change teacher practice and improve strategies and determine what needs to happen next" (Sawchuk, 2011, p. 1). Teachers look at student learning every day, but we are now doing this to look at teachers' impacts on student learning, which raises the stakes on the use of an assessment instrument. Highly qualified teachers are necessary to gain the desired performance by their students.

Studies have shown a correlation between teachers' crafting of rigorous learning objectives and higher performance by their students (Sawchuk, 2011).

2.7 Specific Testing and Licensure Requirements in the State of Kansas

For initial licensure in the State of Kansas, all candidates must complete the following: (1) earn a baccalaureate degree; (2) complete an approved program of teacher education; (3) pass the *Praxis Series* tests for their licensure area; and (4) submit the appropriate licensure form and fees. Prior to entering a teacher education program at a college or university, students may be required to take the *Praxis I* tests to demonstrate their qualifications. In Kansas, teachers must also take the appropriate *Praxis II* tests for each area of licensure being sought. Candidates must receive a passing score from one of the four Principles of Learning and Teaching (PLT) assessments for the initial educator license:

(1) Principles of Learning and Teaching: Early Childhood;

(2) Principles of Learning and Teaching: Grades K – 6;

(3) Principles of Learning and Teaching: Grades 5 – 9;

(4) Principles of Learning and Teaching: Grades 7 – 12.

Each test area requires a qualifying score of 161.

No Child Left Behind (NCLB), Title II, Part A ensures that all K – 12 students have teachers with subject matter knowledge and teaching skills necessary to help all students achieve high academic standards regardless of their individual learning styles or needs. Kansas requires that all public elementary and secondary teachers employed to teach core academic subjects be designated "highly qualified." All general education core content teachers must be "highly qualified." All teachers must have demonstrated subject-matter competency in each of the core academic subjects they are assigned to teach. To add additional endorsements to their Kansas

licensure, teacher candidates must pass the *Praxis II* licensure examinations for each additional content area.

Chapter Three: Data and Methods

3.1 Research Question / Goal

The goal of this study was to examine the relationship between the scores generated by the Interactive Computer Interview System and the scores on the *Praxis II* Principles of Learning and Teaching assessment.

3.2 Validity of the Interactive Computer Interview System

Validity refers to the extent to which a test measures what it is designed to determine. To ensure content validity, questions selected for use in the ICIS instrument were required to measure constructs important to both the *Praxis III* and *Teacher of the Future* documents. Using these two documents assured two national commissions had identified each concept measured by the interview instrument as critical to teaching excellence. Upon purchasing the ICIS program to be used by a district, administrators must be trained in its use ensuring the results are reliable. A scoring rubric increases internal reliability by consistently measuring answers given by the interviewees.

The *Praxis II* PLT test includes multiple-choice questions, essays or constructed-response questions, or a combination of the latter question types. Educators, faculty of teacher preparation programs, and disciplinary specialists prepare *Praxis* test questions. The developers and reviewers follow rigorous, standardized procedures to ensure that test material accurately reflects the skills being tested. High quality assessment standards result in each question being reviewed by ETS experts as well as content advisory groups, individuals who are also engaged in establishing guidelines and standards for what the tests should measure. After the test questions have been reviewed and revised, they are administered in trial situations and assembled into the assessments. Tests are then reviewed according to ETS procedures to ensure that all questions

are free of cultural bias. Statistical analyses of individual questions ensure that all items provide appropriate measurement information.

The *Praxis II*, also referred to as the Principles of Learning and Teaching (PLT) Test, measures the general pedagogical knowledge of potential new teachers at one of the four grade level ranges: Early Childhood, K – 6, 5 – 9, and 7 – 12. The grade level tests are based upon the same general knowledge and contain largely questions that are universal for all grade levels. However, some of the questions are directed toward a specific grade level. Neither a specific number of questions nor an established percentage of questions will appear on the *Praxis II* PLT relevant to the particular grade level. In fact, because this is a computer-generated test, there contains some variation from test to test within the same grade level. When testing multiple examinees, no two tests are identical. Questions will appear in different orders, and the same content question may be stated in a slightly different manner without altering the meaning of the question. For the purpose of this study, it is important to note that the content of the *Praxis II* PLT contains minimal variations between elementary and secondary levels; however, slight differences do exist in the questions that are grade-level specific.

Representative descriptions of topics covered in each category of the *Praxis II* Principals of Learning and Teaching and the approximate percentage of the examination are provided below. A detailed list of these topics is found in Appendix D. These categories are the same for both the elementary and the secondary examinations.

Seventy multiple-choice questions are given based upon the following content categories:

1. Students as Learners (22.5%)
    a. Student Development and the Learning Process
    b. Students as Diverse Learners

c. Student Motivation and the Learning Environment

2. Instructional Process (22.5%)
    a. Planning Instruction
    b. Instructional Strategies
    c. Questioning Techniques
    d. Communication Techniques
3. Assessment (15%)
    a. Assessment and Evaluation Strategies
    b. Assessment Tools
4. Professional Development, Leadership, and Community (15%)

Four constructed-response questions are given and the scoring guide follows:

5. Analysis of Instructional Scenarios (25%)
    a. Score of 2
        - Demonstrates a thorough understanding of the aspects of the case that are relevant to the question;
        - Responds appropriately to all parts of the question;
        - If an explanation is required, provides a strong explanation that is well supported by relevant evidence;
        - Demonstrates a strong knowledge of pedagogical concepts, theories, facts, procedures, or methodologies relevant to the question.
    b. Score of 1
        - Demonstrates a basic understanding of the aspects of the case that are relevant to the question;

- Responds appropriately to one portion of the question;
- If an explanation is required, provides a weak explanation that is supported by relevant evidence;
- Demonstrates some knowledge of pedagogical concepts, theories, facts, procedures, or methodologies relevant to the question.

c. Score of 0

- Demonstrates misunderstanding of the aspects of the case that are relevant to the question;
- Fails to respond appropriately to the question;
- Is not supported by relevant evidence;
- Demonstrates little knowledge of pedagogical concepts, theories, facts, procedures, or methodologies relevant to the question.

3.3 Data

For this study, the scores from the Interactive Computer Interview System subtests were used to compare the scores from the *Praxis II* Principles of Learning and Teaching examination to discover if the two scores have a correlation. Records from 129 actual interviews held by an administrator in a school district in the state of Kansas using the Interactive Computer Interview System were provided to Dr. Howard Ebmeier for research purposes. The interviewer was trained in the administration of the interview tool and used it regularly. The request for data, the *Praxis II* PLT scores, was approved by the research department of the Kansas State Department of Education. Researchers were asked to match the individuals' scores from the *Praxis II* PLT with the scores from the 129 ICIS interviews submitted. From the 129 ICIS interview scores submitted, 58 were successfully matched by KSDE researchers, and *Praxis II* PLT scores were

provided for those interviewees. Because of the location of the school district, it is possible that many of the candidates interviewed were from states other than Kansas and thus had no *Praxis* scores on file with the state department.

Subjects for this study included 58 teachers, 33 secondary teachers and 25 elementary teachers. The interviews were conducted for the purpose of filling teaching vacancies in the district. Questions were asked in each of the areas of the Interactive Computer Interview System until the minimum number of questions had been answered and stability had been achieved. If the candidate's answers were still varied after the minimum number of questions had been answered, the program continued to ask questions in that same theme until stability was attained or the maximum number of questions had been reached. For these interviews the short version was used. The questions available from the bank for each of the four subtests are listed in Appendix C.

Table 3.1: Table of the Descriptive Statistics of the Sample

Description of Participants	Number of Interviewees
Total Number of Interviewees	58
Secondary Interviewees	33
Elementary Interviewees	25

To gain insight into the relationship of the *Praxis II* PLT scores and the subtests calculated by the Interactive Computer Interview System received by these same individual teachers, scatter plots, or scatter diagrams, were used. In a scatter plot, every experimental subject or unit in the study is represented by a point in two-dimensional space. For the purpose of this study, the *Praxis II* PLT score represents the independent variable (x) because it is expected to result in variation of the ICIS subtest score. The dependent variable (y) is the

Interactive Computer Interview System subtest, Knowledge of Content, Knowledge of Teaching, Knowledge of Students, or Working with Others, score for each teacher. A regression line drawn through the data points for each scatter plot, *Praxis II* PLT scores and ICIS scores, shows the slope of the regression. The degree to which the points cluster around the regression line is related to the correlation (*r)* between the two variables. The Pearson product-moment correlation coefficient (*r*) was calculated to measure the degree of relationship between these two variables. Correlation measures both the strength and direction of the relationship between *x* and *y*. Values of *r* range between -1.0 and +1.0. A correlation coefficient close to 0 is evidence that there is no relationship between *x* and *y*. The strength of the relationship between *x* and *y* is measured by how close the correlation is to +1.0 or -1.0. The actual correlation coefficient was calculated using the following equation:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

A hypothesis test was used to determine whether the population correlation coefficient, *p*, is significantly different from 0 based on the value of the calculated correlation coefficient, *r*. The hypotheses are stated as follows:

$H_0: p \leq 0$

$H_0: p \geq 0.$

This statement tests whether a positive correlation exists between *x* and *y*. The test statistic for the correlation coefficient uses the t-distribution as follows:

$$t = \frac{r}{\sqrt{\frac{1 - r^2}{n - 2}}}$$

for which:

r = the calculated correlation coefficient from the ordered pairs

n = the number of ordered pairs

The critical t-statistic is based on $d.f. = n - 2$, and the level of significance for the two-tailed test is charted at $\alpha = 0.05$, $\alpha = 0.10$, and $\alpha = 0.20$, and the value of critical t-statistic is taken from the Statistical Table (Howell, 2008). The data collected for each interviewee is presented in Appendix A.

Chapter Four: Results

4.1 Analyses of Correlations between the Interactive Computer Interview System Subtests and the Praxis II PLT Exams

From the data collected from 58 licensed teachers (two interviewees were given the normal ICIS interview and the short version was used for all others) the Pearson Product-moment, or correlation coefficient, was utilized to measure the strength of the linear relationship between the ratings of the *Praxis II* PLT and the subtests of the Interactive Computer Interview System. Two-tail *t*-tests were performed to determine whether the population correlation coefficient, *p*, is significantly different from 0 based on the value of the calculated correlation coefficient, *r*. The correlation coefficients are listed in Table 4.1.

Table 4.1: Correlation Coefficient, *r*, between the ICIS Subtests and the *Praxis II* PLT Scores

Interviewees	Knowledge of Content	Knowledge of Teaching	Knowledge of Students	Working with Others
All Interviewees	0.175**	0.351*	0.217**	0.189**
Elementary	-0.011	0.231***	0.204	0.140
Secondary	0.487*	0.434*	0.257**	0.269**

* $\alpha = .05$, correlations that allow for the rejection of the H_0.
**$\alpha = .10$, correlations that allow for the rejection of the H_0.
***$\alpha = .20$, correlations that allow for the rejection of the H_0.

Scatter plots were also used to visualize the correlations between the ICIS subtests and the *Praxis II* PLT scores for the interviewees. These scatter plots follow.

Figure 1: Line of Regression for the ICIS Knowledge of Content Subtest and *Praxis II* PLT for all interviewees.

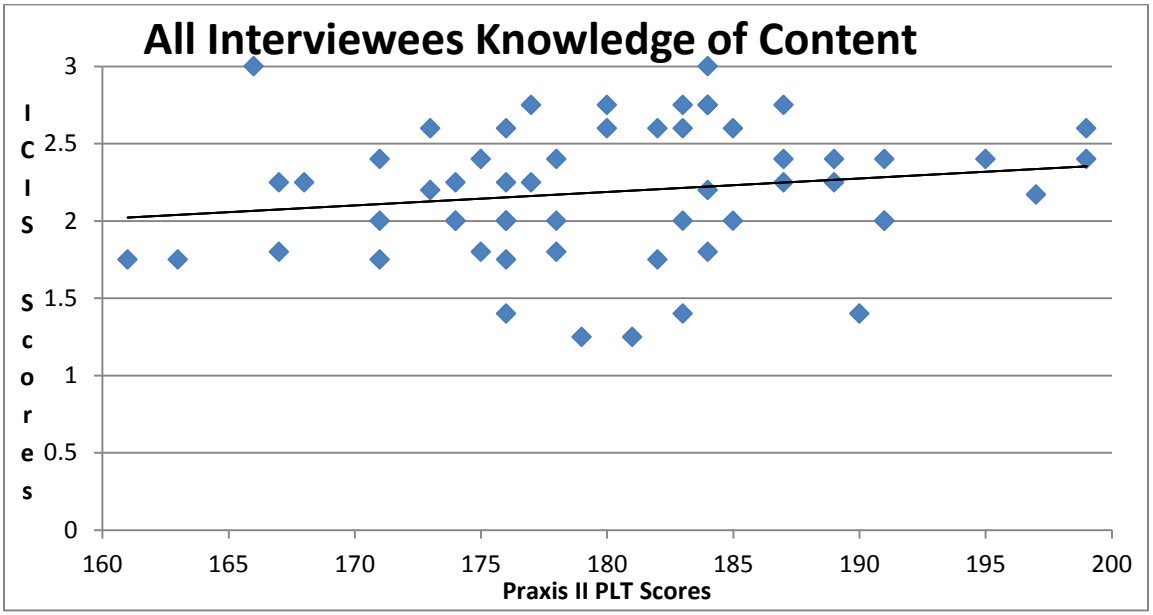


Figure 2: Line of Regression for the ICIS Knowledge of Content Subtest and *Praxis II* PLT for elementary interviewees.

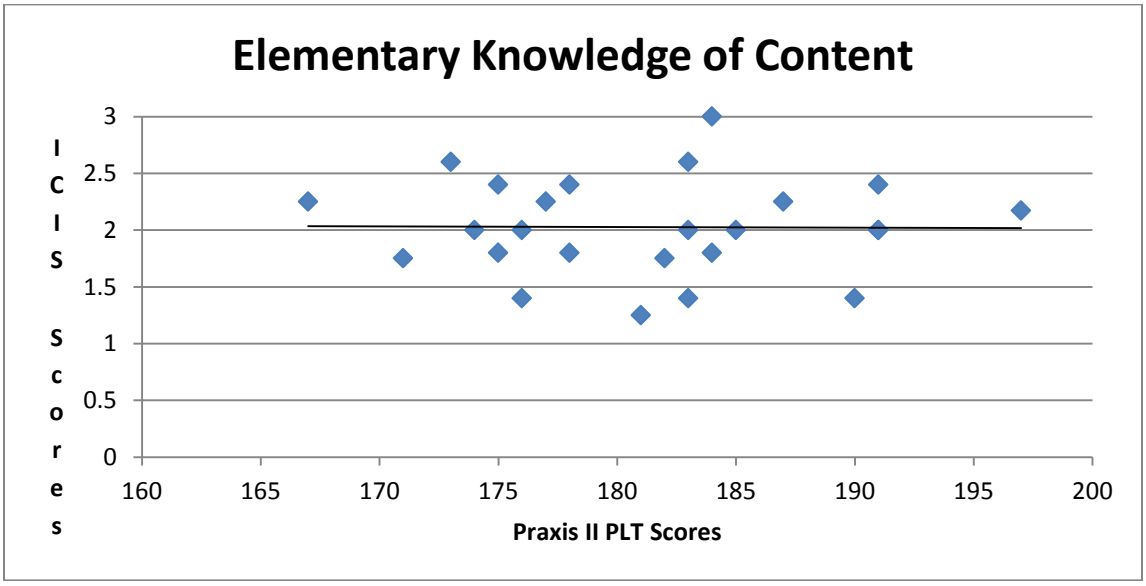


Figure 3: Line of Regression for the ICIS Knowledge of Content Subtest and *Praxis II* PLT for secondary interviewees.

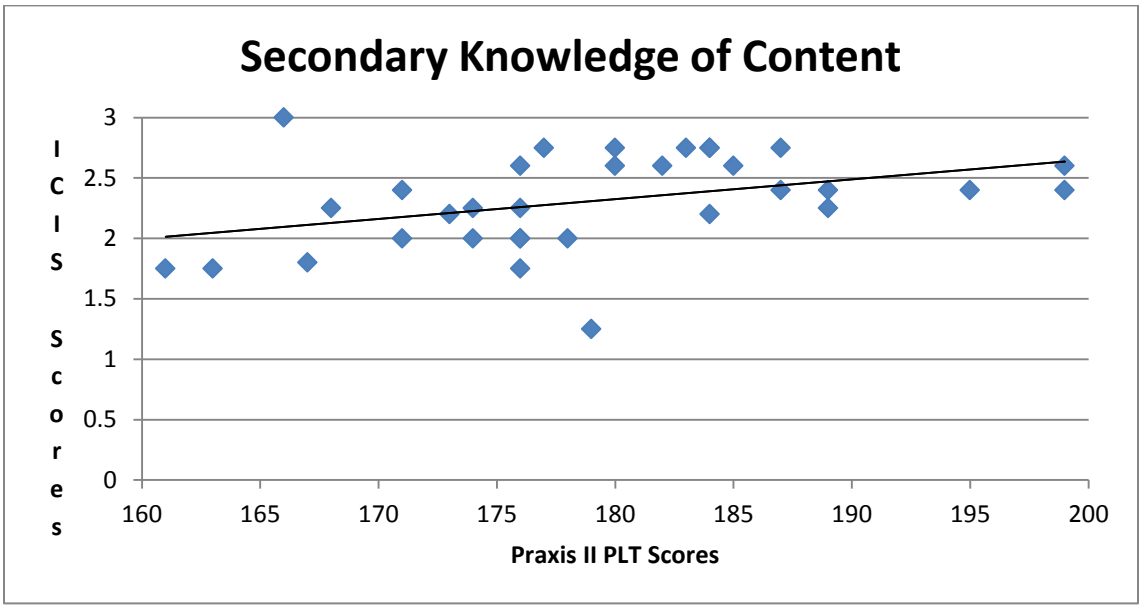


Figure 4: Line of Regression for the ICIS Knowledge of Teaching Subtest and *Praxis II* PLT for all interviewees.

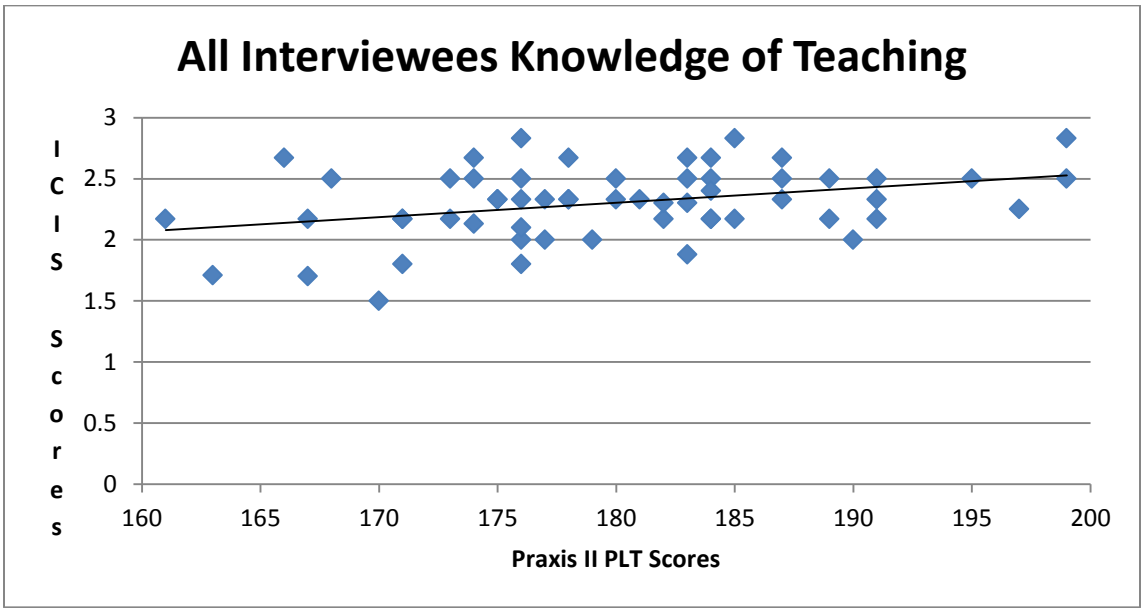


Figure 5: Line of Regression for the ICIS Knowledge of Teaching Subtest and *Praxis II* PLT for elementary interviewees.

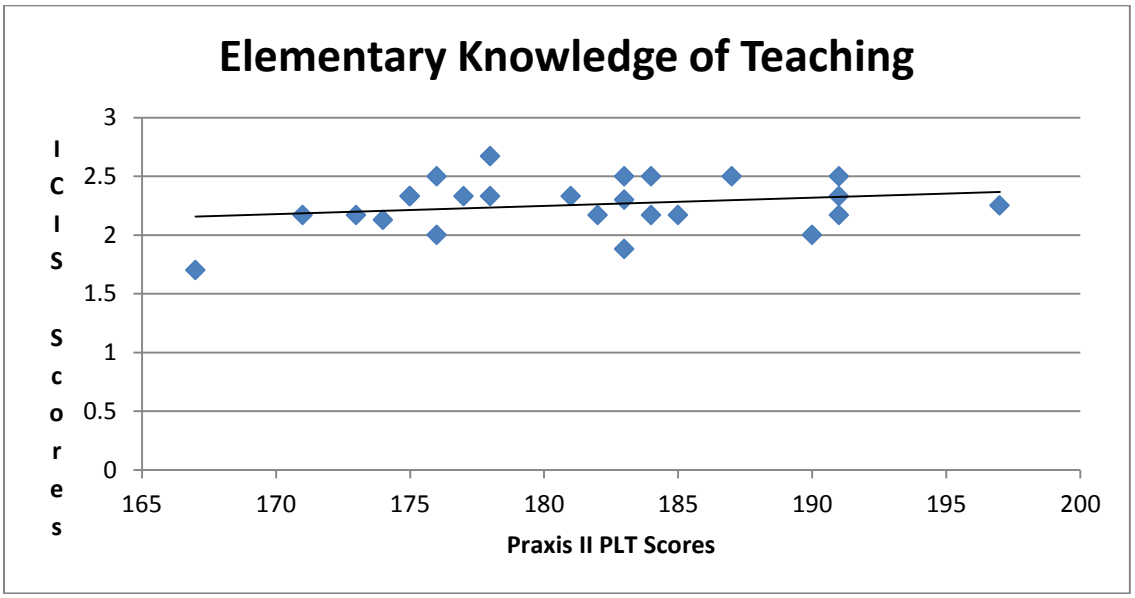


Figure 6: Line of Regression for the ICIS Knowledge of Teaching Subtest and *Praxis II* PLT for secondary interviewees.

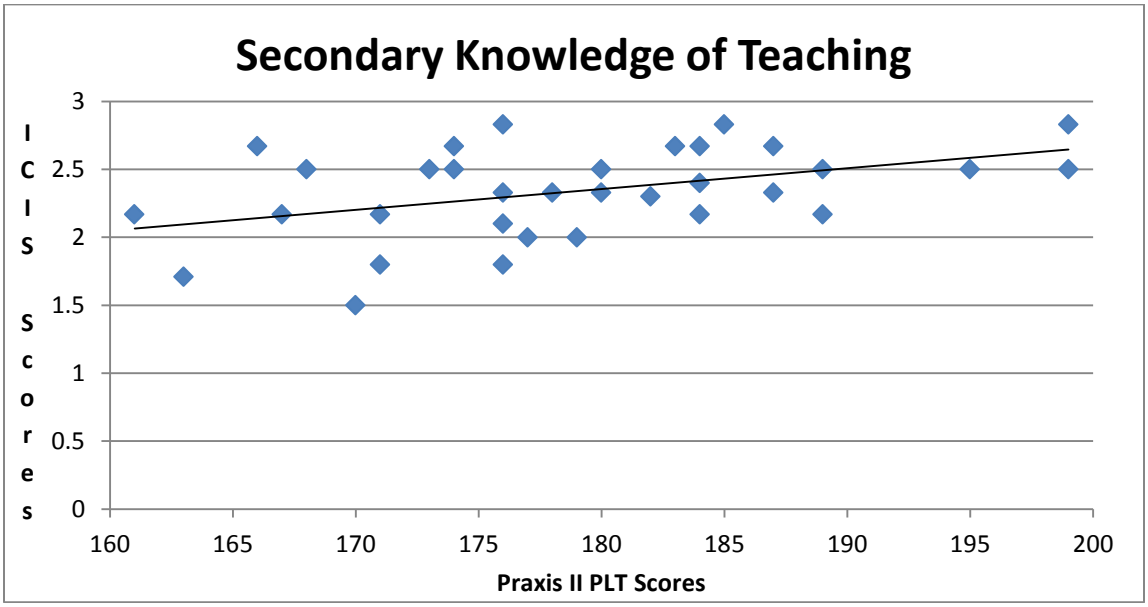


Figure 7: Line of Regression for the ICIS Knowledge of Students Subtest and *Praxis II* PLT for all interviewees.

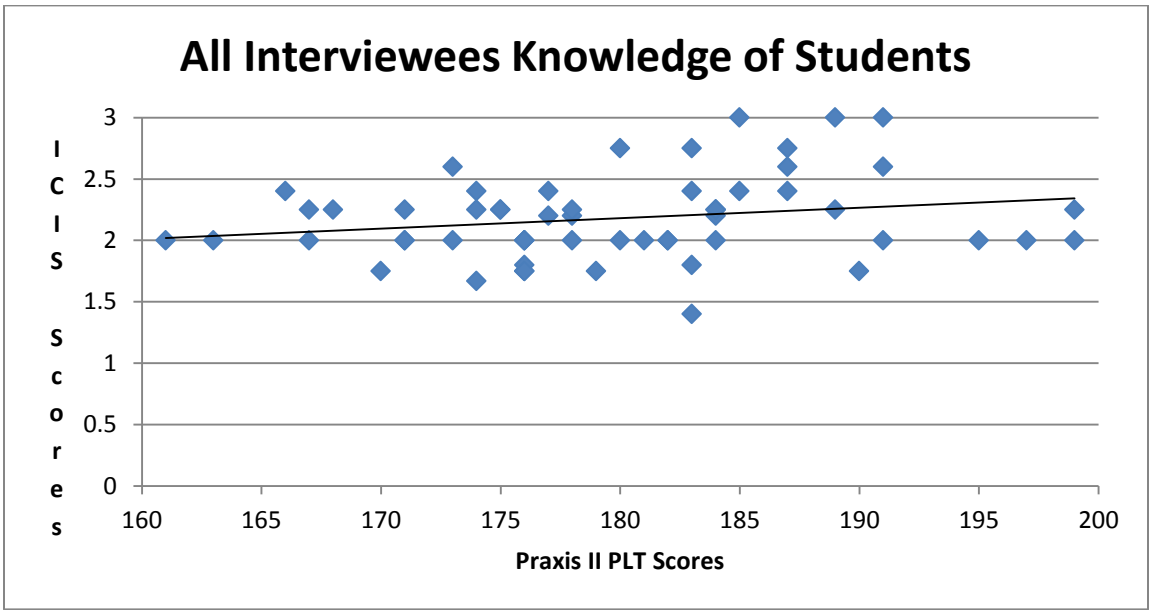


Figure 8: Line of Regression for the ICIS Knowledge of Students Subtest and *Praxis II* PLT for elementary interviewees.

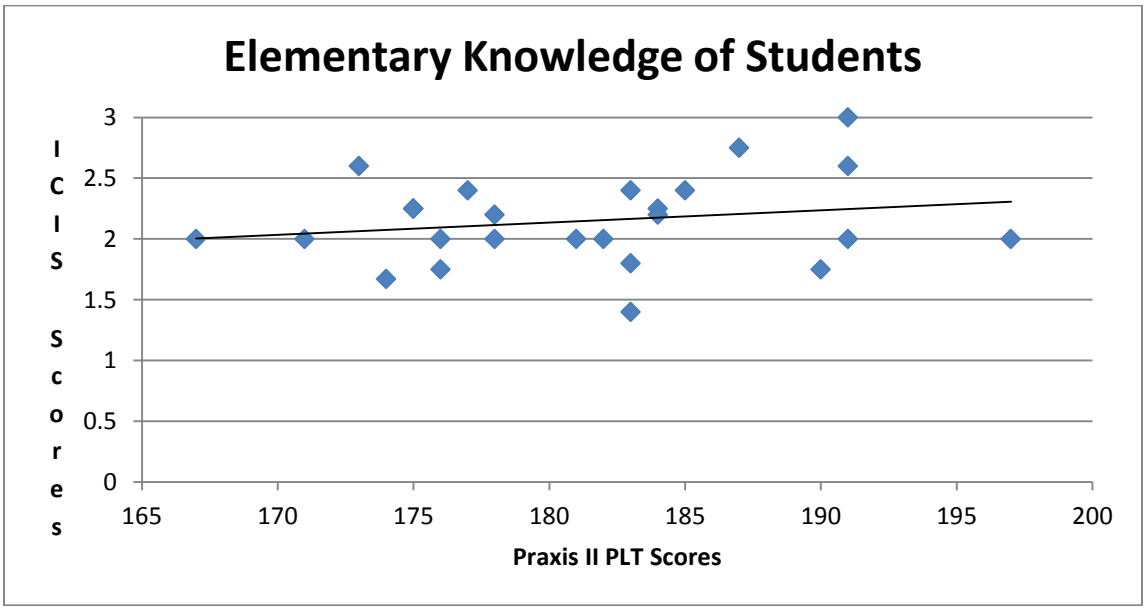


Figure 9: Line of Regression for the ICIS Knowledge of Students Subtest and *Praxis II* PLT for secondary interviewees.

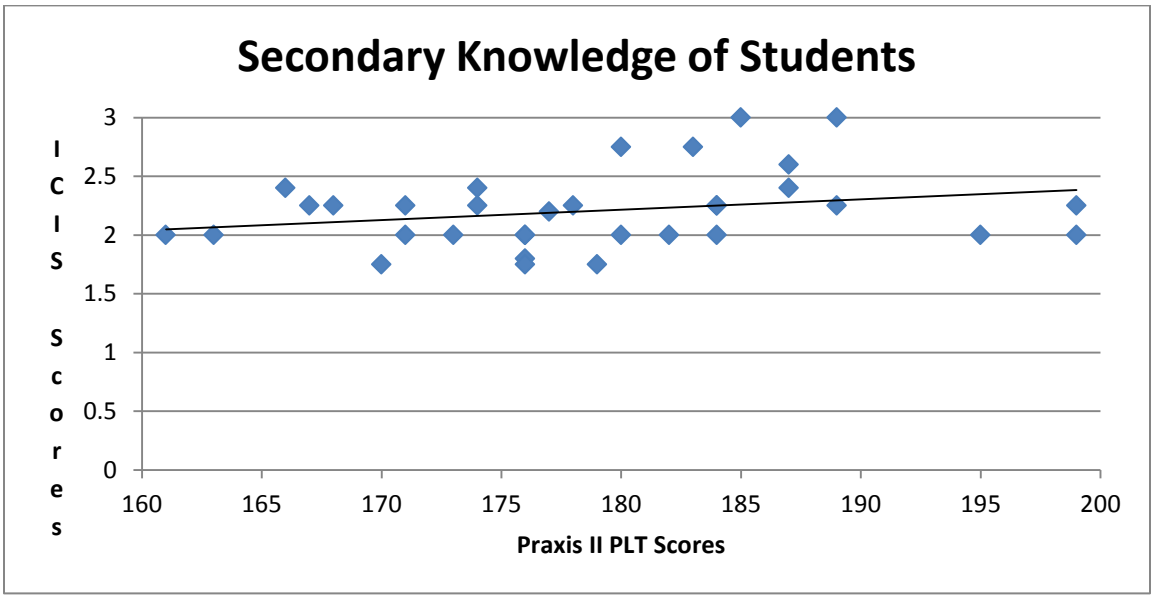


Figure 10: Line of Regression for the ICIS Working with Others Subtest and *Praxis II* PLT for all interviewees.

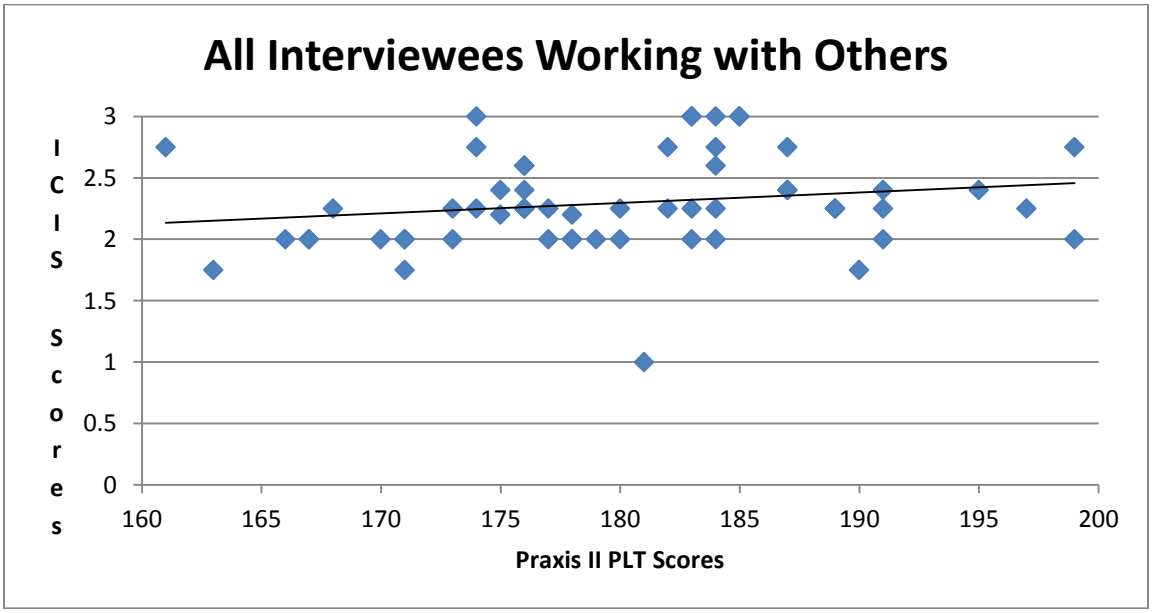


Figure 11: Line of Regression for the ICIS Working with Others Subtest and *Praxis II* PLT for elementary interviewees.

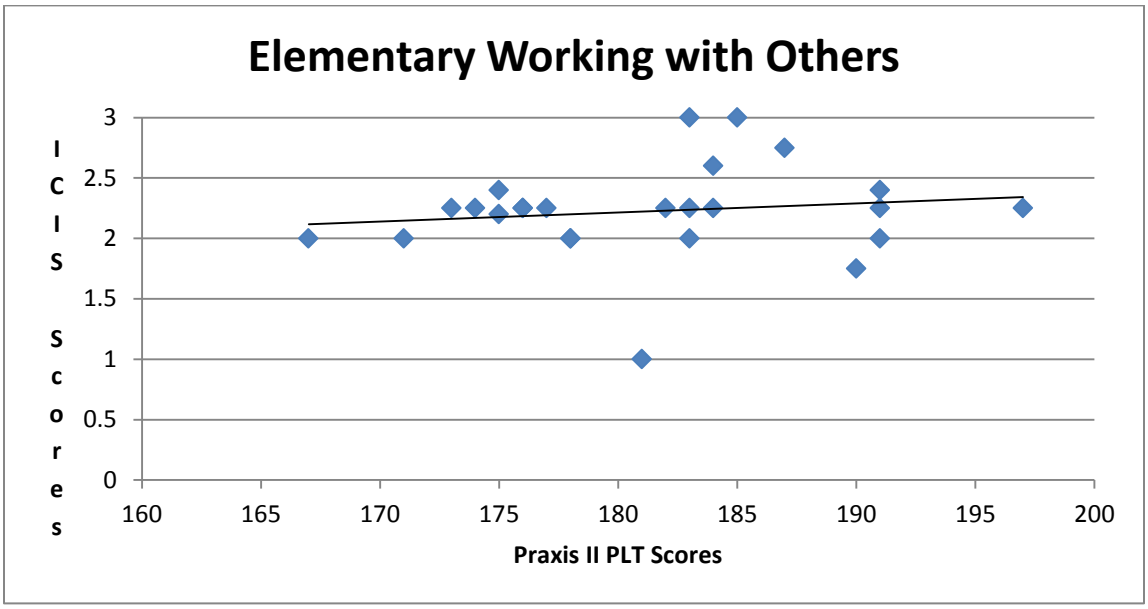
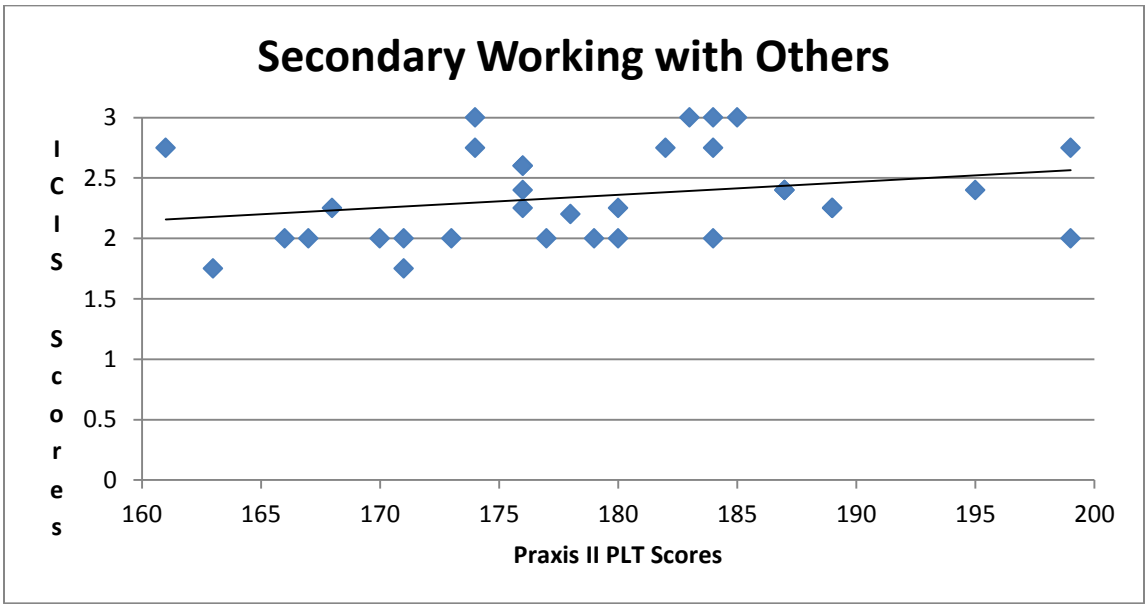


Figure 12: Line of Regression for the ICIS Working with Others Subtest and *Praxis II* PLT for secondary interviewees.



4.2 Analyses of Mean Scores from the Interactive Computer Interview System Subtests and the Praxis II PLT Scores

Examination of the means finds that the secondary interviewees have a slightly higher mean score on each ICIS subtest when compared to the elementary interviewees and all interviewees. However, the *Praxis II* PLT mean score is slightly lower for the secondary interviewees when compared to the other two groups examined. The mean scores on each subtest follow in Table 4.2.

Table 4.2: Mean Scores of the ICIS Subtests and the *Praxis II* PLT Tests

Interviewees	Knowledge of Content	Knowledge of Teaching	Knowledge of Students	Working with Others	*Praxis II* PLT
All Interviewees	2.19	2.30	2.18	2.29	180
Elementary	2.03	2.26	2.15	2.22	181
Secondary	2.31	2.34	2.21	2.35	179

Eighty-five percent of the secondary interviewees scored at two and above on the ICIS Knowledge of Content subtest. On the other hand, only 64% of elementary interviewees scored at 2 and above on this same subtest. On the ICIS Working with Others subtest 94% of secondary interviewees and 92% of elementary interviewees scored at 2 and above. Eighty-eight percent of secondary and ninety-two percent of elementary interviewees scored at two and above on the ICIS Knowledge of Teaching subtest. On the ICIS Knowledge of Students subtest, 88% of secondary and 80% of elementary interviewees scored at 2 and above. When examining the *Praxis II* PLT scores, 45% of secondary interviewees and 56% of elementary interviewees scored at or above the mean score of 180.

Chapter Five: Discussion

5.1 Conclusions

This study compares the Interactive Computer Interview System (ICIS), a structured interview protocol developed by Dr. Howard Ebmeier at the University of Kansas, and the *Praxis II* Principles of Learning and Teaching (PLT), the exam taken by teachers in many states to obtain state licensure. The relationships demonstrated between the ICIS subtests and the *Praxis II* PLT examinations were found to be greatest at the secondary level and for the Knowledge of Teaching scale.

This data analysis allows for the rejection of the H_0 when the value of the calculated *t* is greater than the value of the critical *t* and concludes that there is indeed a significant correlation between some of the ICIS subtest scores and the *Praxis II* PLT scores. Based upon the results of the Knowledge of Content subtests we are unable to reject the H_0 for the subgroup of elementary interviewees. However, for the secondary interviewees we can reject the H_0, $\alpha = .05$, and conclude that there is a significant correlation between the secondary teachers' Knowledge of Content scores and their *Praxis II* PLT scores. When $\alpha = .10$ we are also able to reject the H_0 when all interviewees are examined and conclude that a significant correlation is indicated for this alpha value.

When the correlations between the ICIS Knowledge of Teaching subtests and the *Praxis II* PLT scores are examined we are able to reject the H_0, $\alpha = .05$, for all interviewees and for the subgroup of secondary interviewees, concluding that a significant relationship does exist. We are able to reject the H_0 for the elementary interviewees when the significance level, α, reaches .20.

Upon examination of the ICIS Knowledge of Students and the Working with Others subtests and the *Praxis II* PLT scores, we are able to reject the H_0 when $\alpha = .10$ for all interviewees and for the subgroup of secondary interviewees. A significant relationship is indicated at this value for alpha for those two groups. However, no significant relationship is indicated between the scores for elementary teachers.

The scores of teachers licensed in elementary education only showed significance between the subtest Knowledge of Teaching and *Praxis II* PLT at $\alpha = .20$. The *Praxis II* PLT test is not a direct measure of cognitive ability; it is, however, reasonable to believe that cognitive ability allows a person testing to make more efficient judgments about the questions and answers and, consequently, perform better on the test. Likewise, while the Interactive Computer Interview System has a subtest labeled Knowledge of Content, it is not a direct measure of cognitive ability. Both are measuring teachers' abilities to apply pedagogical knowledge to given performance tasks. The results from secondary teachers are more likely to support the meta-analysis of the relationship between scores on structured interviews and cognitive ability completed by Cortina, Goldstein, Payne, Davison, and Gilliland (2000). This particular study showed a strong relationship between structured interviews and measures of cognitive ability.

As a result of these analyses of the data, this research indicates that there is a significant correlation between the *Praxis II* PLT assessment scores and the scores from Knowledge of Content and the Knowledge of Teaching themes from the Interactive Computer Interview System for teachers licensed at the secondary level. A significant correlation is also demonstrated for the Knowledge of Teaching when all interviewees are considered.

The *Praxis II* PLT exam is assessing pedagogical knowledge and the teacher's ability to apply this knowledge. When a teacher is being interviewed using the Interactive Computer Interview System he/she must use this same pedagogical knowledge to answer questions pertaining to each of the four subtests—Knowledge of Content, Working with Others, Knowledge of Teaching, and Knowledge of Students. While slight variations may occur within the questions at the two levels—elementary and secondary, the pedagogical knowledge is based upon the identical categories and topics. The *Praxis II* PLT is one measure of the knowledge, understanding, and application of educational practices foundational to beginning a career as an educator. This may explain why the Knowledge of Teaching subtest resulted in the most significant relationships among the groups examined.

This study found that the ICIS Knowledge of Teaching subtests and the *Praxis II* PLT scores demonstrated the most significant relationships. This study lends support and assists in the validation of the ICIS Knowledge of Teaching scale. All interviewees and the secondary subgroup were significant at the level $\alpha = .05$, and the elementary subgroup reached significance $\alpha = .20$. We may presume that with a larger sample size the elementary correlation would have been significant earlier as well.

5.2 Type I and Type II Errors

When a decision is reached with a statistical test, there is a chance that this decision may be the wrong one. The statistician makes this decision using a rational process and can specify the conditional probabilities of a decision being in error. Alpha (α) is the probability of a Type I error, which is the error of rejecting the H_0 when it is actually true. Type II error is the error of not rejecting the H_0 when it is indeed false. Because of the small sampling size, a type II error may have occurred during this study.

5.3 Limitations and Suggestions for Future Research

The *Praxis II* Principles of Learning and Teaching includes categories that overlap the Interactive Computer Interview System subtests. Questions relating to Students as Learners and Instructional Process will be found in the ICIS subtests Knowledge of Content, Knowledge of Teaching, and Knowledge of Students. Questions pertaining to Assessment are found in the Knowledge of Teaching subtest. Professional Development, Leadership, and Community questions are located in the ICIS subtests Knowledge of Content, Knowledge of Teaching, and Working with Others. Because distinctions do not exist between the *Praxis II* PLT categories and the ICIS subtests, a general ICIS score may be appropriate to properly determine the correlation between these two instruments. The *Praxis II* PLT provides one score that includes all assessment categories, whereas the ICIS instrument is divided into the four theme clusters. Examination of questions from both instruments finds that they are gathering some of the same information, yet each also includes data unique to the individual instruments.

This study helps to demonstrate the importance of including multiple measures in the employment process. Having the tools available to assist school administrators in the hiring process helps ensure teachers will continue to meet the increasing demands of the profession. The structured interview is one of the most effective tools to gather reliable information from interviewees, second only to the application blank. This study demonstrates relationships do exist between the Interactive Computer Interview System and the *Praxis II* Principles of Learning and Teaching test. Consequently, these may be two examples of the tools used in the hiring process. In addition to gathering and referring to multiple data points—the application, references, resume, interview scores, test scores, writing sample—multiple interviewers allow for consensus among a varied group when hiring new employees.

The direction of future inquiry can build upon this study's findings by including a larger sample size to explore the relationship between the use of the structured interview protocol of the ICIS and scores on the *Praxis II* PLT or other similar exams. It may be important to use only questions that are similar on both instruments to see if this would produce a higher correlation. There were undoubtedly some questions found on one or the other instrument that had no corresponding equivalent questions, and this may be an important factor that reduced the correlations found in this study.

References

American Association of School Personnel Administrators. *Interactive Computer Interview System.* Olathe, Kansas, 2003.

Angrist, J.D. & Guryan, J. (2004). Teacher Testing, Teacher Education, and Teacher Characteristics. *The American Economic Review* 94(2): 241 – 246.

Angrist, J.D. & Guryan, J. (2007). *Does teacher testing raise teacher quality? Evidence from state certification requirements.* A revised version of NBER Working Paper #9545 and IZA Working Paper 1500.

Bergman, M.E., Donovan, M.A., Drasgow, F., Overton, R.C., & Henning, J.B. (2008). Test of Motowidlo et al's (1997) theory of individual differences in task and contextual performance. *Human Performance* 21: 227 – 253.

Blanton, L.P., Sindelar, P.T., & Correa, V.I. (2006). Models and measures of beginning teacher quality. *The Journal of Special Education* 40(2): 115 – 127.

Burnett, J.R. & Motowidlo, S.J. (1998). Relations between different sources of information in the structured selection interview. *Personnel Psychology* 51: 963 – 983.

Campion, M.A., Palmer, D.K., & Campion, J.E. (1997). A review of structure in the selection interview. *Personnel Psychology* 50.

Chan, D. & Schmitt, N. (2002). Situational judgment and job performance. *Human Performance* 15(3): 233 – 254.

Clevenger, J. & Pereira, G.M. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology* 86(3): 410 – 417.

Cortina, J.M., Goldstein, N.B., Payne, S.C., Davison, H.K., & Gilliland, S.W. (2000). The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology* 53: 325 – 351.

Ebmeier, H., Dillon, A., & Ng, J. (2007). *Employment selection instruments—what we have learned from ten years of research.* A paper presented at Annual Conference of AASPA. Kansas City, KS.

Gitomer, D.H. & Latham, A.S. (2000). *Generalizations in teacher education: seductive and misleading.* A research memorandum providing preliminary and limited dissemination of ETS research prior to publication.

Kansas State Department of Education. *Highly Qualified Teacher Overview.* A Fact Sheet posted on www.ksde.org.

McDaniel, M.A., Whetzel, D.L., Schmidt, F.L., & Maurer, S.D. (1994). The validity of employment interviews: a comprehensive review and meta-analysis. *Journal of Applied Psychology* 79(4): 599 – 616.

McDaniel, M.A., Hartman, N.S., Whetzel, D.L., & Grubb III, W.L. (2007). Situational judgment tests, response instructions, and validity: a meta-analysis. *Personnel Psychology* 60: 63 – 91.

Metzger, S.A. & Wu, M.J. (2003). *Commercial teacher interviews and their problematic role as a teacher qualification*. A paper presented at the American Educational Research Association Annual Conference. Chicago, IL.

Metzger, S.A. & Wu, M.J. (2008). Commercial teacher selection instruments: the validity of selecting teachers through beliefs, attitudes, and values. *Review of Educational Research* 78(4): 921 – 940.

Mount, J.K., Witt, L.A., & Barrick, M.R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology* 53: 299 – 323.

O'Connell, M.S., Hartman, N.S., McDaniel, M.A., Grubb III, W.L., & Lawrence, A. (2007). Incremental validity of situational judgment tests for task and contextual job performance. *International Journal of Selection and Assessment* 15(1): 19 – 29.

Posthuma, R.A., Morgeson, F.P., & Campion, M.A. (2002). Beyond employment interview validity: a comprehensive narrative review of recent research and trends over time. *Personnel Psychology* 55.

Praxis Series, The. *Kansas Test Requirements.* A Fact Sheet posted on www.ets.org.

Praxis Series, The. Principles of Learning and Teaching: Grades K – 6 (0622). A Test at a Glance posted on www.ets.org.

Praxis Series, The. Principles of Learning and Teaching: Grades 7 – 12 (0624). A Test at a Glance posted on www.ets.org.

Ryan, A.M. & Tippins, N.T. (2004). Attracting and selecting: what psychological research tells us. *Human Resource Management* 43(4): 305 – 318.

Salgado, J.F. & Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology* 11(3): 299 – 324.

Savini, C. (2010). Bias among the well-intentioned: how it can affect the hiring process. *Independent School* 69(2): 60 – 67.

Sawchuk, S. (2011). Wanted: ways to measure most teachers. *Education Week* 30(19): 1, 15.

Schmidt, F.L. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin* 124(2): 262 – 274.

Schmidt, F.L. & Rader, M. (1999). Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology* 52: 445 – 464.

Webster, W.J. (1988). Selecting Effective Teachers. *The Journal of Educational Research* 81(4): 245 – 253.

Young, I.P. & Delli, D.A. (2002). The validity of the teacher perceiver interview for predicting performance of classroom teachers. *Educational Administration Quarterly* 38(5): 586 – 612.

Zigo, D. & Moore, M.T. (2002). Accountability and teacher education: just like everyone else—teaching to the test? *English Education* 34(2): 137 – 155.

Appendix

Appendix A: ICIS Subtests and *Praxis II* PLT Scores

Table A.1: Data from the Interactive Computer Interview System and *Praxis II* Principles of Learning and Teaching Scores

Level	ICIS Working w/ Others	S.D.	ICIS Knowledge of Content	S.D.	ICIS Knowledge of Teaching	S.D.	ICIS Knowledge of Students	S.D.	*PRAXIS II* PLT
Sec.	2.40	.55	1.75	.50	1.80	.63	1.80	.84	176
Elem.	2.75	.50	2.25	.50	2.50	.55	2.75	.50	187
Elem.	2.00	.63	2.00	0	2.50	.53	2.00	0	191
Sec.	2.00	0	2.75	.50	2.17	.41	2.00	.71	184
Sec.	3.00	0	2.75	.50	2.67	.52	2.75	.50	183
Elem.	1.75	.50	1.40	.55	2.00	0	1.75	.50	190
Sec.	2.25	.50	2.25	.50	2.50	.55	3.00	0	189
Sec.	2.00	0	2.40	.55	2.83	.41	2.00	0	199
Sec.	2.00	0	2.20	.84	2.50	.55	2.00	1.00	173
Elem.	2.40	.55	2.40	.55	2.33	.52	2.25	.50	175
Sec.	2.25	.50	2.75	.50	2.50	.55	2.75	.50	180
Sec.	2.00	.63	1.25	.50	2.00	0	1.75	.50	179
Elem.	2.00	0	2.25	.50	1.70	.67	2.00	0	167
Elem.	2.25	.50	1.75	.50	2.17	.41	2.00	.71	182
Sec.	2.00	0	3.00	0	2.67	.52	2.40	.55	166
Sec.	2.75	.50	2.20	.84	2.40	.84	2.25	.50	184
Elem.	2.25	.50	2.60	.55	2.17	.41	2.60	.55	173
Elem.	2.00	0	1.25	.50	2.33	.52	2.00	0	181
Sec.	2.25	.50	2.25	.50	2.33	.52	2.00	0	176
Elem.	2.60	.55	3.00	0	2.50	.55	2.25	.50	184
Sec.	2.40	.55	2.75	.50	2.33	.52	2.60	.55	187
Sec.	2.75	.50	2.60	.55	2.50	.55	2.25	.50	199
Elem.	2.25	.50	1.40	.55	1.88	.35	1.40	.55	183
Sec.	2.25	.50	2.25	.50	2.50	.53	2.25	.50	168
Sec.	2.00	.63	2.00	.89	1.50	.53	1.75	.50	170
Elem.	2.20	.84	1.80	.84	2.33	.52	2.25	.50	175
Elem.	2.00	0	2.40	.55	2.33	.52	2.00	0	178
Elem.	2.25	.50	1.80	.84	2.17	.41	2.20	.84	184
Sec.	2.00	0	2.60	.55	2.33	.52	2.00	0	180
Sec.	2.20	.84	2.00	.71	2.33	.52	2.25	.50	178
Sec.	2.75	.50	1.75	.50	2.17	.41	2.00	0	161
Sec.	3.00	0	2.00	0	2.67	.52	2.40	.55	174
Elem.	2.25	.50	2.25	.50	2.33	.52	2.40	.55	177
Sec.	2.00	0	2.75	.50	2.00	0	2.20	.84	177
Elem.	2.00	0	1.80	.84	2.67	.52	2.20	.84	178
Elem.	2.25	.50	2.17	.75	2.25	.46	2.00	.63	197
Sec.	2.75	.50	2.25	.50	2.50	.55	2.25	.50	174
Sec.	2.75	.50	2.60	.55	2.30	.82	2.00	0	182

Sec.	2.00	0	2.40	.55	2.17	.41	2.00	.71	171
Elem.	2.40	.55	2.40	.55	2.17	.41	3.00	0	191
Elem.	2.00	.71	1.75	.50	2.17	.41	2.00	0	171
Elem.	3.00	0	2.00	0	2.17	.41	2.40	.55	185
Sec.	1.75	.50	2.00	0	1.80	.63	2.25	.50	171
Elem.	3.00	0	2.60	.55	2.50	.55	2.40	.55	183
Sec.	1.75	.50	1.75	.50	1.71	.61	2.00	0	163
Sec.	2.40	.55	2.40	.55	2.67	.52	2.40	.55	187
Sec.	2.60	.55	2.00	0	2.10	.57	1.75	.50	176
Elem.	2.25	.50	2.00	.89	2.13	.35	1.67	.82	174
Sec.	2.60	.55	2.60	.55	2.83	.41	2.00	0	176
Elem.	2.25	.50	2.00	0	2.33	.52	2.60	.55	191
Sec.	3.00	0	2.60	.55	2.83	.41	3.00	0	185
Sec.	2.00	0	1.80	.84	2.17	.41	2.25	.50	167
Sec.	2.25	.50	2.40	.55	2.17	.41	2.25	.50	189
Elem.	2.25	.50	2.00	.71	2.50	.55	1.75	.50	176
Elem.	2.25	.50	1.40	.55	2.00	0	2.00	0	176
Sec.	3.00	0	2.75	.50	2.67	.52	2.25	.50	184
Elem.	2.00	0	2.00	0	2.30	.67	1.80	.84	183
Sec.	2.40	.55	2.40	.55	2.50	.55	2.00	0	195

Appendix B: Mean, Mode, Median, and Range of Test Results

Table B.1: Mean, Mode, Median, and Range of ICIS Subtest Scores and *Praxis II* PLT Scores

	Working with Others	**Knowledge of Content**	**Knowledge of Teaching**	**Knowledge of Students**	***Praxis II* PLT**
Mean:					
Elem.	2.22	2.03	2.26	2.15	181
Sec.	2.35	2.31	2.34	2.21	179
Mode:					
Elem.	2.25	2	2.33	2	183/191
Sec.	2	2.75	2.5	2	176
Median:					
Elem.	2.25	2	2.3	2	182
Sec.	2.25	2.4	2.33	2.25	178
Range:					
Elem.	2 (1-3)	1.75 (1.25-3)	0.97 (1.7-2.67)	1.6 (1.4-3)	30 (167-197)
Sec.	1.25 (1.75-3)	1.75 (1.25-3)	1.33 (1.5-2.83)	1.25 (1.75-3)	38 (161-199)

Appendix C: ICIS Subtest Theme Questions

Table C.1: ICIS Knowledge of Content Theme Questions

1.	What are the major concepts students should learn in your class?
2.	How do the content recommendations in the teacher's guide influence your decision about what to teach?
3.	How do you decide what concepts to teach?
4.	What content knowledge should your students possess at the end of the year?
5.	What are some of the content outcomes a student should have mastered before promotion to the next grade/subject level?
6.	What are the major concepts students should learn in your class?
7.	Describe how you would expose students to the most current materials and information? Give some examples.
8.	How would you use outside and community resources to help students learn?
9.	What are some of the important keys you would look for to help you decide what to teach?
10.	How do you provide your students with opportunities to use life experiences in the classroom?
11.	Identify a particular lesson you have taught. What prerequisite knowledge would students need?
12.	How do you incorporate students' prior experiences into your lessons?

Table C.2: Working with Others Theme Questions

1.	What is the purpose of collaboratively working with other teachers?
2.	To what extent have you been involved in school improvement activities? Give some examples.
3.	If you believe a given school policy is not in the best interest of some students in your class what would you do?
4.	Why is it important to include the contributions of people from diverse backgrounds? How can this be done?
5.	What kind of information do you provide to parents?
6.	Describe ways you, as the classroom teacher, might communicate with and involve parents in the education of the child.
7.	Describe how you might change your curriculum of teaching methods as a result of feedback from your students' parents.
8.	Describe what your relationship should be as a teacher with members of the community.
9.	If you were given one community volunteer to work with your class for two hours each week, what would you do?
10.	How should parents be involved in the instructional process?
11.	How should you resolve student peer conflicts?
12.	Describe community/school activities in which you have participated and how they have enhanced student learning.
13.	How might you respond if a parent complained to you about an activity that you were using?
14.	Sometimes the building principal has teachers engage in collaborative work projects such as curriculum committees. [Question not identified.]

Table C.3: Knowledge of Teaching Theme Questions

1.	When you ask students questions in class, what do you think about in terms of what to do with their responses?
2.	How would you evaluate the success of your lessons?
3.	How do you get groups of students to work independently and productively?
4.	What are the most important things that you can do to help students succeed?
5.	Describe the teaching strategies you have used to present subject matter. How do you know which one to select?
6.	What kind of assessment feedback do you provide to students?
7.	How can you maximize student achievement in your classroom?
8.	How can you build high expectations for learning?
9.	How do you decide how to structure the lessons for your students?
10.	What opportunities for practice do you provide students on a daily basis in your classroom?
11.	How would you encourage students to respect one another?
12.	What strategies would you employ to foster positive learning environments in your classroom?
13.	How confident are you in your ability to teach all children?
14.	How would you decide what to test or assess in your classroom?
15.	How would you decide what examples or metaphors to use to illustrate points in the lesson?
16.	A variety of approaches to instruction may be utilized. What approaches could you use to deliver instruction?
17.	What steps would you take if a high error or failure rate occurred in your classroom?
18.	How would you decide how fast or slow to move through a lesson?
19.	Describe methods you can use to incorporate critical thinking into your teaching.
20.	What kind of process can you use to score student assessments and tests?
21.	What steps could you take to encourage student – teacher interactions?
22.	What strategies would you use to help a student who is not getting along with others in the class.
23.	What steps could you take to build student interest in your academic field?
24.	How would you assess the effectiveness of your teaching?
25.	How do you go about establishing classroom behavior?
26.	Some districts want all the staff to believe that all children can learn. Research, however, . . . what do you think about this issue?
27.	How do you decide what learning goals are appropriate and inappropriate for your students?
28.	How will you react if a student asks a question that you cannot answer?
29.	How would you decide when and for what reasons to form student-learning groups?
30.	How would you know when to alter your lesson plans?
31.	Select a lesson you have recently taught. How could that lesson be improved?
32.	What resources do you plan to use and/or make available to students in your classrooms?
33.	How could a lesson that doesn't go well be improved?
34.	What could you do to ensure that all students' comments are considered during a discussion?
35.	What steps can you take to reduce wasted time in the classroom?

36.	What does the term classroom discipline mean to you?
37.	What teaching methods could be most effective for you in encouraging students to develop and defend judgments?
38.	What kinds of classroom environment would you hope to establish?
39.	What do you do when a student responds to your question with an obviously incorrect answer?
40.	What is mean by the phrase, “All children are capable of learning”?
41.	If I were to walk into your classroom, what would I observe in terms of the classroom climate?
42.	When a student responds in an unusual way that may illustrate a related point, how would you respond?
43.	What are some of the formal and informal diagnostic techniques you might use in assessing your students?

Table C.4: Knowledge of Students Theme Questions

1.	What interests would typical students in your class have?
2.	Describe some different learning styles you have observed. How can you adjust your lesson?
3.	How could you adapt your lessons to meet the needs of special education students?
4.	Describe how you could adjust your lessons or activities to respond to the composition of students in your classroom.
5.	What do you feel is important for you to know about your students?
6.	What specific student activities and teacher strategies are needed to help special education students?
7.	What kinds of skills would you expect typical students in your class to possess?
8.	Describe a situation in which you might have to differentiate instruction for students with varying development and ability levels.
9.	How would you recognize and respond to individual differences?
10.	What are the most important developmental characteristics of students you might teach?
11.	Describe how you could incorporate students' current knowledge into the design of a lesson.
12.	What kinds of teaching models or methods would work best for typical student in your classes?

Appendix D: *Praxis II* PLT Descriptions of Topics Covered

Table D.1: *Praxis II* PLT Representative Descriptions of Topics Covered in Each Category

I.	**Students as Learners:** Student Development and the Learning Process • Understands the theoretical foundations of how students learn • Knows the major contributions of foundational theorists to education • Understands the concepts and terms related to a variety of learning theories • Knows the distinguishing characteristics of the stages in each domain of human development (i.e., cognitive, physical, social and moral) • Understands how learning theory and human development impact the instructional process Students as Diverse Learners • Understands that a number of variables affect how individual students learn and perform • Recognizes areas of exceptionality and their potential impact on student learning • Understands the implications and application of legislation relating to students with exceptionalities on classroom practice • Recognizes the traits, behaviors, and needs of intellectually gifted students • Recognizes that the process of English language acquisition affects the educational experience of English language learners (ELLs) • Knows a variety of approaches for accommodating students with exceptionalities in each phase of the education process Student Motivation and Learning Environment • Knows the major contributions of foundational behavioral theorists to education • Understands the implications of foundational motivation theories for instruction, learning, and classroom management • Knows principles and strategies for classroom management • Knows a variety of strategies for helping students develop self-motivation
II.	**Instructional Process:** Planning Instruction • Understands the role of district, state, and national standards and frameworks in instructional planning • Knows how to apply the basic concepts of predominant educational theories • Understand how scope and sequence affect instructional planning • Knows how to select content to achieve lesson and unit objectives • Knows how to develop observable and measurable instructional

	objectives in the cognitive, affective, and psychomotor domains • Is aware of the need for and is able to identify various resources for planning enrichment and remediation • Understands the role of resources and materials in support student learning • Knows how to develop lessons as part of thematic and/or interdisciplinary units • Recognizes their role in collaborating with instructional partners in instructional planning Instructional Strategies • Understands the cognitive processes associated with learning • Understands the distinguishing features of different instructional models • Knows a variety of instructional strategies associated with each instructional model • Knows a variety of strategies for encouraging complex cognitive processes • Knows a variety of strategies for supporting student learning • Knows basic strategies for promoting students' development of self-regulatory skills • Understands the design of different group configurations for learning • Understands the use and implications of different grouping techniques and strategies • Knows how to select an appropriate strategy for achieving an instructional objective • Understands the concept of monitoring and adjusting instruction in response to student feedback • Recognizes the purpose of reflecting upon, analyzing and evaluating the effectiveness of instructional strategies • Knows the characteristics of different types of memory and their implications for instructional planning and student learning • Recognizes the role of teachable moments in instruction Questioning Techniques • Knows the components of effective questioning • Understands the uses of questioning • Knows strategies for supporting students in articulating their ideas • Knows methods for encouraging higher levels of thinking • Knows strategies for promoting a safe and open forum for discussion Communication Techniques • Understands various verbal and nonverbal communication modes • Is aware of how culture and gender can affect communication • Knows how to use various communication tools to enrich the learning environment • Understands effective listening strategies

III.	**Assessment:** Assessment and Evaluation Strategies • Understands the role of formal and informal assessment in informing the instructional process • Understands the distinctions among the different types of assessment • Knows how to create and select and appropriate assessment format to meet instructional objectives • Knows how to select from a variety of assessment tools to evaluate students' performance • Understands the rationale behind and the uses of students' self and peer assessment • Knows how to use a variety of assessment formats Assessment Tools • Understands the types and purposes of standardized tests • Understands the distinction between norm-referenced and criterion-referenced scoring • Understands terminology related to testing and scoring • Understands the distinction between holistic and analytical scoring • Knows how to interpret assessment results and communicate the meaning of those results to students, parents/caregivers, and school personnel
IV.	**Professional Development, Leadership and Community:** • Is aware of a variety of professional development practices and resources • Understands the implications of research, views, ideas and debates on teaching practices • Recognizes the role of reflective practice for professional growth • Is aware of school support personnel who assist students, teachers, and families • Understands the role of teachers and schools as educational leaders in the greater community • Knows basic strategies for developing collaborative relationships with colleagues, administrators, other school personnel, parents/caregivers, and the community to support the educational process • Understands the implications of major legislation and court decisions relating to students and teachers