ELEMENTARY SCIENCE TEACHER DEVELOPMENT OF PEDAGOGICAL CONTENT KNOWLEDGE THROUGH PROFESSIONAL DEVELOPMENT: AN EXPLORATORY STUDY

A Thesis Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Master of Education

by

Alana Danielle Newell

May, 2012

APPROVAL PAGE

Acknowledgment

This thesis would not have been possible without the unwavering support of the many individuals who extended their guidance and insight when it was needed. I would like to thank my entire thesis committee, Drs. Bernard Robin, Richard Olenchak and Sara McNeil, for their unending patience and encouragement throughout this entire process.

I would also like to extend my utmost gratitude to the faculty and staff at Baylor College of Medicine's Center for Educational Outreach, and especially Dr. Nancy Moreno and James Denk for the unconditional enthusiasm and assistance I have received throughout this process.

I am forever indebted to Drs. Jason King and Richard Olenchak for taking the time out of incredibly busy schedules to share their statistical and methodological expertise with me.

Finally, I would like to thank my family and friends for your unyielding confidence, consistent reassurance, and unconditional love.

ELEMENTARY SCIENCE TEACHER DEVELOPMENT OF PEDAGOGICAL CONTENT KNOWLEDGE THROUGH PROFESSIONAL DEVELOPMENT: AN EXPLORATORY STUDY

An Abstract
of a
Thesis Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Master of Education

by

Alana Danielle Newell

May, 2012

Newell, Alana Danielle. "Elementary Science Teacher Development Of Pedagogical Content Knowledge Through Professional Development: An Exploratory Study." Unpublished Master of Education Thesis, University of Houston, May 2012.

Pedagogical content knowledge (PCK) – a form of knowledge that allows teachers to most effectively pass their content knowledge on to students – is a useful theoretical construct, but has yet to be translated into an efficient tool for improving the practice of teaching. Using existing data from a 2008-10 study from Baylor College of Medicine's Center for Educational Outreach, this research explored teacher professional development content knowledge gains, demographic data, and student content knowledge data in an attempt to isolate some of the attributes of teachers who are more effective at developing PCK in elementary science topics. The ultimate mission of the study is to examine differences in teachers' ability to develop PCK through professional development. A comparison of results from forward stepwise logistic regression and best-subsets logistic regression found that identifying the development of PCK and the attributes of teachers successful in its development are more complex and varied than expected.

TABLE OF CONTENTS

LIST OF TABLES

I. Introduction

Statement of the Problem

A multitude of program designs and best practices exist in the field of elementary school science teacher professional development, all with the ultimate goal of positively affecting student achievement levels by increasing teacher preparedness. Though there are arguably an infinite number of factors that can influence the success of a teacher at improving student outcomes, including learning environment and the students themselves, there are also differences in the ways in which teachers absorb and use their professional development, and specifically develop their pedagogical content knowledge: that is, their ability to transform their knowledge into effective classroom teaching. If an understanding of the way successful teachers build this knowledge can be gleaned, perhaps the practices used by those teachers can be taught to others who are less successful, and in-services can be shaped to better improve those attributes.

Data

The data for this research were collected by the Center for Educational Outreach (CEO) at Baylor College of Medicine during a 2008-2010 study evaluating the effectiveness of various delivery methods of asynchronous online professional development programs. The original research used two cohorts of volunteer fifth grade teachers from Harris County school districts. The 208 participants (cohort one = 124, cohort two = 129, and 45 individuals participated in both cohorts) were randomly assigned to one of four different online professional development module groups: 1) a video of a content talk on the material, 2) an audio podcast of the aforementioned video content talk, 3) a video with very little content discussion, but a demonstration of how the lesson should be taught, and 4) an interactive short course with videos, text and self-assessment questions scattered throughout. In addition to completing a professional development module, participants in both cohorts taught a standardized lesson to their classroom of students.

This thesis research explored de-identified teacher content knowledge scores from the teacher pre-assessment that was taken prior to completing professional development modules, the teacher post-assessment taken immediately following completion of the professional development modules, the associated student content knowledge scores from the student pre-assessment (administered prior to the lesson being taught), and the student post-assessment, (taken following the completion of the elementary science lesson). Data collected concerning teacher participants' demographic characteristics were also investigated.

Approach to Analysis

Two different data analysis techniques were utilized in this research: forward stepwise logistic regression and best-subsets logistic regression. Logistic regression is a method of data analysis appropriate for assessing hypotheses about the relationships present between a categorical dependent variable and any number of categorical or continuous independent variables. Specifically, logistic regression analysis allows researchers to predict the outcome of some dependent variable – in this case group membership – based upon a model created using the independent, or predictor variables.

Stepwise logistic regression is commonly used in exploratory data analyses because unlike traditional regression techniques, it creates a predictive model by including or excluding independent variables based on those variables' ability to improve upon the model as indicated by the model's increased probability of correctly predicting group membership. Forward stepwise logistic regression forms a regression model by beginning with no predictor variables, and adding the available variables one at a time to the model (provided those variables meet the criteria for inclusion), until the model can no longer be improved upon.

Best-subsets logistic regression is similar to stepwise logistic regression in that it evaluates the different individual variables and combinations of variables in search of the best potential model fit, i.e. the model that most accurately predicts group membership. However, instead of adding variables based on probabilities as determined by the statistical software package being used, best-subsets analyses allow the researcher to evaluate all potential combinations of independent variables based on their suitability as a predictor, thus allowing the investigator, and not the computer, to choose the most appropriate predictive model of the data.

Research Questions

- What role, if any, do the attributes teaching experience, highest degree obtained, number of in-service hours and number of science-specific in-service hours play in the development of elementary science teacher's pedagogical content knowledge?
- Are the attributes of teachers who successfully demonstrate the development of pedagogical content knowledge the same for the two different cohorts?
- What differences, if any, will exist between the results of the analyses from the forward stepwise logistic regression and the best-subsets logistic regression?

II. Literature Review

Pedagogical Content Knowledge

Pedagogical Content Knowledge (PCK) was conceptualized by Lee Shulman as the "missing paradigm" in our understanding of how teachers learn to teach (Shulman, 1986). Specifically Shulman postulated that it is a form of knowledge that uses both a teacher's knowledge of the content area and knowledge of pedagogy, and subsequently allows teachers to represent their understanding of a topic in a way in which their students can also understand it. Though, as its name implies, it an amalgam of pedagogical knowledge and content knowledge, PCK is a separate, extra level of knowledge, and one that differentiates an content expert from a novice teacher (Shulman, 1986).

Since Shulman's initial conception of PCK, many researchers have explored the concept and added their own personal impressions and conceptualization to the definition. While some view PCK as an "experiential application of other forms of knowledge" which includes content knowledge and pedagogy (Gess-Newsome, 1999; Appleton, 2008), the view used for the purpose of this paper is the transformative view of PCK, in which PCK is seen as the transformation of the knowledge the teacher has of the content, various pedagogical approaches, the class they are teaching etc. Teachers with strong PCK are able to use "the most powerful analogies, illustrations, examples, explanations and demonstrations – in a word, the ways of representing and formulating the subject that make it comprehensible to others" (Shulman, 1986). Accordingly, Shulman asserts that teachers need to be strong in PCK in order to be the best possible teachers (Loughran, Mulhall, & Berry 2004).

The scholarly community has explored the concept of PCK in a multitude of ways, and as a result, there are many variations on the general components of PCK. The two that scholars generally agree upon are 1) a knowledge of students and their misconceptions, and 2) an understanding of the specific subject matter and ways to represent it (Rohaan, Taconis, & Jochems, 2009; van Driel, Verloop, & Vos, 1998; Appleton, 2003).

Despite an effort to create a model of the various formats of knowledge that are components of PCK, most scholars agree that PCK is hard to conceptualize, and because it is a transformative form of knowledge, attempting to view it or separate it out into pieces would be an oversimplification of the concept. (Berry, Loughran, & van Driel, 2008; Rohaan et al., 2009). Thus, although the concept of PCK is a useful one from a research standpoint, there is difficulty in translating the concept into actual improvement upon teaching and teacher education (Berry et al., 2008).

Part of the problem in translating the research knowledge of PCK to actual teaching practice is that aside from being difficult to conceptualize, PCK by nature is person-specific, class/student-specific and subject specific (Rohaan et al., 2009). Consequently, the same teacher teaching the same subject with a different group of students would have to develop his or her own PCK specific to that particular group. This high specificity makes it difficult not only to describe PCK, but also to identify it as it is being developed or occurs.

Additionally, since much of the work teachers do centers around the practice of teaching, and not assessing the reasons surrounding why they teach the way they do, teachers infrequently articulate the reasoning and thought processes behind their PCK

(Rohaan et al., 2009). An additional difficulty is the fact that PCK is not easily observed because it is impossible to know exactly why a teacher does what they do in the classroom, and PCK is not necessarily being used at all given teaching moments nor is it entirely expressed through behavior. Mulhall, Berry, and Loughran (2003) purport that "whether or not a particular action by a teacher is illustrative of that teachers' PCK depends upon the teacher's reasons for that action." PCK is also constructed differently between even experienced teachers, meaning PCK would not necessarily be the same for two teachers, even if they had similar teaching contexts (e.g. student achievement level, age group) and content knowledge (van Driel et al.,1998).

Despite the difficulties in finding ways to define and measure PCK, many researchers have worked on this particular issue, especially in terms of measuring PCK on a large scale (Rohaan et al., 2009). Current methods of measuring PCK are typically quite time consuming, and involve a lot of input from the teachers themselves. Additionally, because the development of PCK is not a linear one, instead being multi-faceted, and is partially an internal construct, it is difficult to identify in an individual person, let alone on any large scale (van Driel et al., 1998; Loughran et al., 2004).

In spite of the various difficulties, researchers continue to attempt to gain insight into the construct because of the potentially great effect that may come from gaining an understanding: ideally an improvement in the quality and efficiency of teacher education and eventually student education (Rohaan et al., 2009). One of the major research studies of PCK investigated whether or not there is some ability to translate the knowledge that successful teachers have built for themselves into some format that can be used by other teachers with less PCK (Mulhall et al., 2003). In order to achieve this, Mulhall, Berry, and Loughran explored the knowledge the successful teachers possessed on particular topics, and through documentation of thought processes etc., the researchers then attempted to create a way to pass on that knowledge to other teachers. While they acknowledge that there are differences between successful teachers' characterizations of their knowledge, they believe that there are enough generalizations to support the belief that such knowledge may apply to and be useful to other teachers.

Generally, it is thought that the development of PCK, and specifically science PCK in this study, is largely experiential; teachers gradually construct their PCK using not only their knowledge of the content and pedagogy that they gain through pre-service and in-service trainings, but also their experience in the classroom as both a teacher and a student, the advice from trusted colleagues, as well as personal beliefs and perceptions of science and science education. (Appleton, 2008). Overall, teacher experience seems to be the greatest determining factor in the creation of PCK, and a firm grasp of the subject matter is generally considered to be a prerequisite to its subsequent development (van Driel et al., 1998).

PCK and Elementary Science Education

Research supports Shulman's postulation that strong PCK is necessary for effective teaching. Logically, this makes sense, as the teachers who are best able to represent topics to their students in the most effective manner by presenting the information in a way that their students will best understand and adequately addressing possible student misconceptions, would reasonably have the best outcomes. Appleton (2008) asserts in terms of science teaching that increased PCK in elementary science teachers will also increase their willingness to teach science content.

Studies have found that elementary teachers are often hesitant to teach science, and frequently those that do will teach using methods and pedagogy that resemble subjects they are more comfortable teaching, such as language arts (Appleton, 2003). Additionally, teachers addressing unfamiliar topics have a difficult time accurately and effectively representing knowledge and understanding potential student problems (van Driel et al., 1998). Van Driel et al. also found that knowledge of pedagogy forms a framework that teachers can use as a crutch when their PCK is limited. Experienced teachers, specifically, were able to maintain the flow of their teaching better, since they had a stronger framework when presented with new knowledge outside of their typical content area.

Ken Appleton (2003) suggests that PCK for elementary science teachers is different than PCK for secondary science teachers. While secondary teachers have the opportunity to develop a high level of PCK in one or two specific areas because of specialization, such as being solely a 10th and 12th grade chemistry teacher and thus building the specific knowledge for those age groups in that subject area, elementary teachers often must teach multiple grades and/or multiple subjects. Even for a science specialist, there are a multitude of areas within science that must be taught, making the process of building adequate PCK more difficult. For self-contained classroom teachers, science is just one subject of many. Through his research, Appleton has concluded that instead of building PCK as secondary teachers do, elementary science teachers generate their PCK through the use of "activities that work."

Specifically, these teachers who have little PCK and perhaps feel uncomfortable teaching science, cope by finding activities that are easily managed, get students involved, have predictable outcomes and allow the students to gain knowledge without much teacher intervention. Appleton (2003), found that the use of activities that work has a close relationship with the teachers' science PCK. This method of teaching not only allows the teacher to cope with little knowledge of the science content, but it also serves as a source of or supplement to their science PCK. As teachers use activities that work, they become more confident in their science teaching, and may even gain their own content knowledge through the process (Appleton, 2003). Thus, teachers may be able to begin to form their own science PCK through the process of using these activities. Appleton found, however, that activities that work might not be sufficient for the development of science PCK on their own. The problems that arise with exclusively using this form of science teaching include the fact that teachers who only use activities that work may develop a view of science in which they believe all problems can be solved by experimentation only, ignoring the open-ended side of science.

There have been attempts to better align science content with an associated pedagogy in order to guide teachers in the process of developing science PCK (Loughran, Mulhall, & Berry, 2008). Loughran et al. noted that by sensitizing pre-service teachers to the topic of PCK, they were less likely to use PCK as simply a theory and more likely to look at it as a way in which they might be able to manage their professional knowledge.

PCK and Professional Development in Elementary Science

There is a general sense by some researchers that elementary teachers are ill prepared to teach, particularly when it comes to science (Berry et al., 2008). The consequence of this lack of preparedness is a lack of confidence in the ability to teach it, and in many cases, a complete avoidance of the subject, wherein the teachers either completely forego science education, or teach very little, and often in a format better suited for another subject that they are more comfortable with (Appleton, 2003). Heller et al. (2012) found that "teachers with strong content and content-specific pedagogical knowledge have been shown to provide higher quality instruction." Thus it is incredibly important to provide teachers with such knowledge in order to improve not only their ability to understand the content, but also to improve teaching.

There is evidence that professional development programs are capable of improving teacher content knowledge, as well as a teachers' understanding of student thinking and concept development, classroom practice and student outcomes (Borko, 2004; Hill, 2009), so it is reasonable to suggest that elementary teachers participate in effective, content-focused science professional development. In addition to strengthening teachers' science content knowledge, there is also a need for professional development that can enhance a teachers' PCK (Appleton, 2008; Goodnough & Hung, 2009), and in fact, quality pre-services or in-services have been found to increase PCK (Clermont, et al., 1993; van Driel et al., 2002).

Unfortunately there is evidence that that there is a dearth of effective professional development programs, specifically those that are cognizant of teachers' weaknesses and link those with learning opportunities to remedy them (Hill, 2009; Borko, 2004). With an increased need for quality professional development for a large number of teachers, as well as the time constraints associated with the profession, online teacher professional development has become an increasingly popular option (US Department of Education, 2009). Studies of the effectiveness of online professional development programs have discovered that online learning can be equally, and in some cases more effective than traditional face-to-face learning environments (US Department of Education, 2009; Dede et al., 2008). Because of the relative newness of the medium, however, more research is needed to identify the best practices and most effective methods to be used within online professional development (Dede et al., 2008). Dede et al. (2008) go on to suggest that future research about online professional development address topics such as the best program designs, the pedagogical strategies that are most effective in an online environment, and the overall effectiveness of the programs. There is also a push towards more empirical research that builds upon existing research to create "usable knowledge" that can actually impact the practice of teaching and teacher professional development.

Comparison of Methods

Stepwise techniques are commonly used in exploratory data analyses for many reasons. Proponents of the methods appreciate the ability to obtain an overview of the relationships present among the variables in a study (Grayson, 1997; McCann, Short, & Stewin 1986; Thomas & Galambos 2004; Peng 2002), especially in scenarios in which there are no existing hypotheses regarding which variables may be the most accurate predictors in a model. In those situations, proponents of stepwise methods believe that the procedure allows the researcher to select the most parsimonious predictive model by excluding extraneous variables, while simultaneously allowing the researcher to better understand the interactions between the dependent and independent variables (Astin & Denon, 2009).

Despite the outward appearance of having a number of benefits, stepwise regression techniques have some very vocal detractors. Critics of the techniques have three major complaints. The first such criticism is that stepwise methods used in most statistical software packages do not use the correct degrees of freedom (*df*) in the evaluation of explained variance, causing any such calculation to be incorrect (Thompson, 1989). Thompson (1989) asserts that this occurs because the calculations consider only variables which are accepted into the model, ignoring the fact that a greater number of variables had been considered for entry into the model. Proponents argue that the issue incorrect *df* can be overcome by simply lowering the p-value used to gauge statistical significance (Astin & Denson, 2009).

The second major criticism of stepwise regression techniques is that they do not, in fact, select the best possible predictors for any given model, though many researchers use them to do precisely that. By selecting variables one at a time for entry into the model based on their individual predictive powers, the procedure ignores the possibility that several variables that individually may not have much power might create a superior predictive model when joined together (Thomson, 1989; Whitaker, 1997; Thompson, 2001). Astin and Denson (2009) agree with this assertion, however, they do not believe that stepwise methodology should be abandoned altogether, but that it should be used in conjunction with other analysis techniques if the main purpose of the research is to determine the best possible subset of predictors.

Finally, the third major criticism of the stepwise technique is that researchers can incorrectly assume that a variable's order of entry into the model is indicative of its relative importance to the model, which may not be the case at all. Sampling error can distort variables' order of entry, thus giving an invalid "importance" to some variables, particularly because the entry of a variable into the model is largely dependent on the predictive powers of the variables entering the model before it (Thompson, 1989;

Whitaker, 1997; Thompson , 2001). Stepwise proponents argue that sampling error exists whether stepwise methods are used or not; sampling error at any given step would result in an inflated regression coefficient for that variable in non-stepwise procedures (Astin & Denson 2009). Astin and Denson (2009) go further, and state that in small sample sizes stepwise approaches actually can minimize the risk of sampling error because there are a smaller number of independent variables involved.

Together, all of these weaknesses of stepwise regression lead its critics to believe that it is prone to producing inaccurate models (Antonakis & Dietz, 2011; Kieffer, Reese, & Thompson, 2001; Thompson, 2001; Whitaker 1997).

Though best-subsets logistic regression is a newer, less frequently used method of finding the best variables for a given predictor, there are compelling arguments that it may be a preferred alternative to using a traditional stepwise regression (King, 2003). Best-subsets allows for a comparison of all potential regression models via summary statistics, which permits the researcher to be active in the model selection process. King (2003) points out that this benefit can be particularly helpful when there are models that perform similarly, and thus the deciding factor between the two models may be non-statistical, as in the ease of attaining certain scores or information, or the associated costs of a given variable. In short, the researcher possesses the knowledge of not only the accuracy of the model(s), but also the context in which it will be used, permitting a more thorough assessment of the utility of each model.

A major drawback of best-subsets logistic regressions is the laborious nature of the analysis. Creating, comparing and selecting the appropriate predictors are all quite work intensive processes (King, 2003), and unlike stepwise regression procedures, many statistical software packages, including the Statistical Package for the Social Sciences (SPSS) used in this research, will not calculate best-subsets regressions. Thus the procedure must be carried out with additional techniques and calculations by the investigators, which can be significantly more time consuming than the stepwise methods readily calculated with most major statistical software.

Another consideration with best-subsets logistic regression, is that although the procedure avoids the two other pitfalls of stepwise logistic regression, it has the same issues with *df* that stepwise procedures do: namely, that probability estimates computed for the models do not take into account that some *df* should have been lost because more parameters were evaluated than were included in the models. Thus, observed probability values for best-subsets regressions are necessarily invalid.

III. Methodology

Purpose of the Study

The intention of this research was to build upon the original study by exploring the differences between teachers that demonstrated the ability to successfully pass on their content knowledge gains to their students, and their peers who were able to gain content knowledge themselves, but were not as successful in passing that knowledge on to their students. The original study found no relationship between high teacher knowledge gains and student knowledge gains, and it was my belief that the difference in the teachers' levels of success was that the teachers able to pass their knowledge onto their students were more adept at developing pedagogical content knowledge. By identifying successful teachers and comparing their demographic data to less successful teachers, the hope was to find patterns that might provide some insight into the characteristics that aid in the development of elementary science teacher PCK.

Research Methodology and Design

I examined the original study data from the teachers in both cohorts who received professional development, and classified those participants according to their knowledge gains. I first identified teachers with above average personal content knowledge gains within each cohort, and then subdivided the above average group into teachers whose students also had above average gains, and those whose students had below average gains. Using demographic information collected about each participant in the original study as independent variables, I then completed both a forward stepwise logistic regression and a best-subsets logistic regression on each cohorts' dataset in an effort to find a model that would accurately predict a teachers' ability to have passed on his or her content knowledge gains to students. The objectives of the research were to : 1) Investigate the role of teaching experience, level of degree obtained, in-service participation and science-specific in-service participation in the development of PCK as demonstrated by above average student content knowledge gains from pre- to post-assessment. 2) Compare the differences between the results of the models for each cohort. 3) Compare the differences in the results from the two regression procedures.

Procedures

I began my research by calculating the descriptive statistics of the content knowledge score gains for all participants in the original study who received professional development (i.e. the entirety of cohort one, and four of the five groups in cohort two). Specifically, I was interested in the mean content knowledge score gain for each cohort as the designated cut-off between the teachers with above average content knowledge gains, and those with below average content knowledge gains. All data from teachers with below average content knowledge gains were removed from further analysis.

After separating the teachers, I turned to the student data. I calculated the descriptive statistics for all student content knowledge score gains for each cohort. I then found the mean content knowledge score gain for the classroom of each of the teachers with above average content knowledge gains, and used those values to separate the teachers into high success and low success groups. If the mean gain for a teacher's class was greater than the student mean gain for the overall cohort, the teacher was classified as highly successful at developing pedagogical content knowledge as demonstrated by passing content knowledge onto his or her students, and was thus designated to the "High Group". If a teacher with above average content knowledge gains had a classroom mean gain of less than the student mean gain of the overall cohort, they were classified as having low levels of success at developing PCK, and were designated to the "Low Group". This designation was codified within the dataset as 0=Low Group Member, 1=High Group Member. A Pearson chi-square was then calculated in order to identify any interactions between the different professional development module groups and the High/Low group designation.

After the classifications were made, the demographic factors to be used as independent variables in the analysis were identified and added to the dataset. The independent variables were all chosen because background research suggested that they might play a role in the development of PCK, and because there were very few missing cases in the dataset. The independent variables were as follows:

- Teaching Experience in years (TEXP)
- Highest Degree Obtained (HDO) – (coded for entry as 1= Bachelor's Degree, 2 = Bachelor's Degree plus Graduate Work, 3= Master's Degree, 4=Doctoral Degree)
- Total number of in-service hours in the year prior to participation in the original study (INSERV)
- Total number of science-specific in-service hours in the year prior to participation in the original study (SCIINS)

Once the independent variables were isolated and coded for entry (in the case of HDO), a forward stepwise logistic regression was run for each cohort using the likelihood ratio method provided in the statistical software.

Finally, best-subsets logistic regression models were computed using the procedure outlined in King (2003) (Appendix A).

All statistical analyses were completed with the Statistical Package for the Social Sciences (SPSS) software, version 19 for Mac OSX.

IV. Results

Cohort One

There were 124 teachers in the first cohort, and the mean gain for the overall group was 3.11 points (SD = 2.51) from the pre-assessment to the first post-assessment (Table 1). Each individual module's mean gain score was also calculated.

Table 1
Cohort One Mean Gains and Standard Deviations

	Teachers			Students		
	N	Mean Gain	SD	N	Mean Gain	SD
Video Content	29	2.76	2.85	608	2.81	3.37
Audio Only	33	3.42	2.61	737	3.09	3.36
Video Demo	31	3.19	2.33	639	3.77	3.76
Short Course	31	3.03	2.18	644	3.70	3.52
All Cohort One	124	3.11	2.51	2628	3.34	3.53

After the division of the participants into high and low groups, 51 teachers had a gain from the pre-assessment to the first post-assessment that was greater than the mean. In this case, that meant any gain greater than three since the gain scores were all whole numbers. Of those 51 teachers, only 17 of them had students with mean content knowledge gains greater than the overall group's gains (Table 2).

Table 2
Cohort One Group Designations by Mean Gain Scores

	All Above Mean Teacher Gains	Teachers with Above Mean Student Gains (High Group)	Teachers with Below Mean Student Gains (Low Group)
Video Content	9	2	7
Audio Only	16	2	14
Video Demo	13	8	5
Short Course	13	5	8
All Cohort One	51	17	34

Module designation had a statistically significant correlation with group designation (Pearson Chi-square = 8.433, df = 3, p = 0.038), however the sample sizes were too small to accurately investigate the effects of the predictor variables on each module group individually.

A forward stepwise logistic regression (cut value = 0.500) completed on the overall group found that none of the independent variables were useful predictors in designating teachers to the above or below average student gains groups.

Table 3
Cohort One Observed and Predicted Frequencies for High or Low Group Designation by Forward Stepwise Logistic Regression

Observed	Predicted		% Correct
	Below Average Student Gains	Above Average Student Gains	
Below Average Student Gains	32	0	100%
Above Average Student Gains	16	0	0%
Overall % Correct			66.70%

The best-subsets logistic regression for this cohort confirmed the findings of the forward stepwise logistic regression. In logistic regression techniques, R^2 assesses effect size, though it cannot be considered an accurate variance-accounted-for measure since it is based upon log likelihood functions which are not really a sum of squares as they would be in linear regressions (King, 2003). Mallow's C_p is a measure of predictive squared error, and is expected to be equal to or less than the number of predictors plus one (p+1, p=predictors), when the model fit is correct (King, 2003). Thus for any model to be considered as a predictor, the C_p value would need to be less than 5.0. The extremely low R^2 and high Mallow's C_p values verify that none of the models in Table 3 are sufficient predictors of a teacher's ability to successfully develop PCK from the provided professional development. Furthermore, the consistency seen among both the C_p and R^2 values for all models suggests that all of the chosen independent variables performed similarly poorly, with none showing potential for further exploration in this particular setting.

Table 4
Cohort One Best Subsets Models

Predictor Variable(s) in Model	R^2	$R^2_{\text{L-Adj}}$	C_p
TEXP	0.004	-0.018	44.25
HDO	0.009	-0.013	44.04
INSERV	0.004	-0.018	44.25
SCIINS	0.003	-0.019	44.30
TEXP HDO	0.010	-0.034	43.98
TEXP INSERV	0.006	-0.038	44.15
TEXP SCIINS	0.005	-0.039	44.19
HDO INSERV	0.014	-0.030	43.80
HDO SCIINS	0.012	-0.032	43.89
INSERV SCIINS	0.005	-0.040	44.22
TEXP HDO INSERV	0.014	-0.053	43.79
TEXP HDO SCIINS	0.013	-0.055	43.87
TEXP INSERV SCIINS	0.007	-0.061	44.12
HDO INSERV SCIINS	0.016	-0.051	43.73
TEXP HDO INSERV SCIINS	0.016	-0.075	43.72

Cohort Two

The second cohort had 129 participants, however, 25 of those teachers received no professional development, and thus they could not be included in this particular analysis. The 104 participants who received professional development had a mean gain of 3.38 points (SD=3.75) from the content pre- to the first post-assessment. The students of these teachers had a mean gain of 1.06 points (SD=2.83) from pre- to post-assessment.

Table 5
Cohort Two Mean Gains and Standard Deviations

	Teachers			Students		
	N	Mean Gain	SD	N	Mean Gain	SD
Video Content	25	4.24	3.89	483	1.00	2.60
Audio Only	27	4.19	4.35	551	1.06	2.77
Video Demo	27	2.96	3.38	517	1.05	3.07
Short Course	25	2.12	3.02	507	1.12	2.86
Cohort Two*	104	3.38	3.75	2058	1.06	2.83

*Lesson only group removed

The 2010 cohort had 46 participants with above average content knowledge gains, 24 of whom also had students with above average content knowledge gains.

Table 6
Cohort Two Group Designations by Mean Gain Scores

	Above Mean Teacher Gains	Teachers with Above Mean Student Gains (High Group)	Teachers with Below Mean Student Gains (Low Group)
Video Content	14	8	6
Audio Only	17	8	9
Video Demo	9	6	3
Short Course	9	3	6
All Cohort Two*	49	25	24

*Lesson only group removed

Professional development module designation did not have a statistically significant correlation with group designation in cohort two.

Table 7
Cohort Two Observed and Predicted Frequencies for High or Low Group Designation by Forward Stepwise Logistic Regression

	Predicted		
Observed[a]	Below Average Student Gains	Above Average Student Gains	% Correct
Below Average Student Gains	0	23	0
Above Average Student Gains	0	24	100
Overall % Correct			51.10%

[a]Two cases were omitted due to incomplete data.

A forward stepwise logistic regressions completed on cohort two (cut value = 0.500) showed that none of the independent variables satisfied the statistical criteria for entry into the model, meaning that none of the variables had any statistically significant relationship with the dependent variable – the designation group of the teachers.

As with cohort one, the best-subsets logistic regression for cohort two confirmed a lack of statistically significant correlation between the predictors and the intercept. Additionally, the models for cohort two performed similarly to the models in cohort one in that is there are consistently low and R^2 and high C_p values for all of the tested models.

Table 8
Cohort Two Best Subsets Tables

Predictor Variable(s) in Model	R^2	$R^2_{\text{L-Adj}}$	C_p
TEXP	0.001	-0.021	42.36
HDO	0.003	-0.019	42.27
INSERV	0.004	-0.019	42.25
SCIINS	0.001	-0.021	42.37
TEXP HDO	0.006	-0.039	42.13
TEXP INSERV	0.005	-0.040	42.19
TEXP SCIINS	0.040	0.002	42.33
HDO INSERV	0.007	-0.038	42.10
HDO SCIINS	0.004	-0.041	42.22
INSERV SCIINS	0.006	-0.039	42.15
TEXP HDO INSERV	0.011	-0.058	41.92
TEXP HDO SCIINS	0.007	-0.062	42.09
TEXP INSERV SCIINS	0.008	-0.061	42.04
HDO INSERV SCIINS	0.009	-0.060	42.03
TEXP HDO INSERV SCIINS	0.015	-0.079	41.77

Comparisons

There are many factors that may have contributed to the non-significance of the results. Given the number of variables available and the many categories within each one, the sample size, particularly within cohort two may have been a limiting factor in the analysis. Including the four different levels of the "highest degree obtained" (HDO) variable, there were seven different independent variables involved in each regression. While the 124 and 104 cases in cohort one and cohort two respectively are sufficient, the small sample size of each module did not allow for investigation of possible relations between the module designation groups and the level of content knowledge gains by the students. It is entirely possible that different models would have had better results within the professional development module groups in each cohort, but unfortunately there was no way to accurately explore this potential interaction given the sample sizes.

Another potential limiting factor of this investigation is the fact that the student gains from pre- to post-assessments were relatively small for both groups (cohort one – mean gain=3.34 points, SD=3.53; cohort two – mean gain= 1.06, SD=2.83). Such small content knowledge gains may have made it difficult to get an accurate picture of the differences between the groups.

An additional possibility that should be taken into consideration is that the subject of the analysis may not have been an accurate portrayal of the development of PCK. The original study was highly standardized, and the instructions the teachers were provided with directed them to avoid additional research on the topics covered. It is therefore entirely possible that examples and illustrations a teacher may have wanted to use during the course of instruction were avoided in order to comply with the rigidity of the original study. Thus, while there may be differences in the student gains for these teachers, they may not have been demonstrative of that teacher's development of PCK.

Although the results from both cohorts were unremarkable in their lack of statistical significance, there were several worthwhile outcomes. First, since teaching experience was the variable which prior research had indicated was the greatest determining factor in the development of PCK, there was an expectation of correlation with the dependent variable. It was therefore entirely unanticipated that there was not a statistically significant relationship between teaching experience and the ability for teachers demonstrating high content knowledge gains to pass their knowledge on to their students. Perhaps even more surprising was the fact that the data for the models in the best-subsets regression that included teaching experience were incredibly similar to the other models not including teaching experience. Thus, it would seem that in this scenario, teaching experience is no more a predictor of success than any other demographic variable tested.

The second outcome of note is the utility of the best-subsets logistic regression as compared to the forward stepwise logistic regression, even when the data is not statistically significant. Though on the surface both the regressions provided the same information, namely that a statistically significant correlation does not exist between the dependent variable and any of the independent variables, the forward stepwise logistic regression provided substantially less information. The best-subsets model allows for deeper inspection into the data, permitting a closer examination of the relationships between the different variables. The best-subsets approach revealed that the high Mallow's C_p and low R^2 values are consistent for all variables involved in the study, whereas the stepwise technique provided no additional insight into the variable interactions. If, in fact, there had been any relationships that were closer to significance, or that were not consistent with the other values in the model, it would have been useful to observe, and could perhaps lead to further investigations, especially given that the nature of the research is exploratory. Researchers who rely solely on stepwise methods would perhaps miss out on some of the nuances of this rich data set.

V. Conclusions and Summary

This chapter will summarize the entirety of the thesis research, address the posed research questions, evaluate the meanings of this study in the context of the existing body of knowledge in the field, and examine the implications and potential future directions of this investigation and related research.

Summary

Pedagogical content knowledge (PCK), a transformative amalgam of a teacher's knowledge of content and pedagogy, is exceedingly useful as a theoretical construct – allowing researchers to identify and define one of the aspects that separate effective teachers from less successful ones – but has yet to find any practical application inside the classroom. This research began with the intent of isolating four of the demographic attributes belonging to those teachers who were more successful at developing elementary science PCK through professional development, with the hope that the information gleaned could be used, to help less effective teachers develop PCK either on their own or through PCK-specific professional development programs.

This thesis research explored existing student and teacher assessment data from a Baylor College of Medicine (BCM) study in which two cohorts of teachers underwent one of four online professional development modules, and then were required to teach a standardized lesson to their students. Teacher content knowledge assessment score gains were used to identify teachers with greater than average gains, and those teachers were separated into two groups: those who were successful at learning elementary science content and passing it on to their students (as determined by above average student content knowledge gains), and those who were successful at learning the content, but were unable to pass on that knowledge to their students. The underlying belief was that teachers who had above average content knowledge gains themselves, and were able to pass that knowledge on to their students were demonstrating the effective development of PCK in elementary science topics, whereas their peers with below average student gains did not develop elementary science PCK.

Forward stepwise logistic regression and best-subsets logistic regression were employed as data analysis methods to evaluate four different demographic variables – 1) number of years of teaching experience, 2) highest degree obtained, 3) number of in-service hours over the previous year, and 4) number of science-specific in-service hours over the previous year – in an effort to discern if any of these attributes were significantly correlated with the successful development of PCK in the participants.

Neither form of logistic regression found statistically significant associations between any of the independent variables and the dependent variable (i.e. membership in the successful group of teachers).

This research hoped to address the following questions:

- What role, if any, do the attributes teaching experience, highest degree obtained, number of in-service hours and number of science-specific in-service hours play in the development of elementary science teacher's pedagogical content knowledge?
- Are the attributes of teachers who successfully demonstrate the development of pedagogical content knowledge the same for the two different cohorts?

- What differences, if any, will exist between the results of the analyses from the forward stepwise logistic regression and the best-subsets logistic regression?

PCK and Research Results

While previous studies suggest that the development of pedagogical content knowledge is largely experiential, that did not appear to be the case in this research. In fact, the best-subsets logistic regression models that incorporated teaching experience as an independent variable performed no differently than the models that did not include teaching experience.

One of the possible explanations for this finding is that the underlying assumption of this thesis research may be false – that is, the cause of the differences in the participants' success levels was not due to disparities in the development of elementary science PCK. The scripted nature of the lessons involved in the original study, as well as the fact that it is generally impossible to know when a teacher is applying his or her PCK during the course of instruction without asking for specific reasoning behind teaching methods, make it difficult to determine exactly how much PCK, if any, the participants in the study actually developed and utilized. This ambiguity is consistent with current research into PCK, which asserts that actions indicative of a particular teachers' PCK are incredibly hard to pinpoint without a great deal more observation and teacher input than occurred with this particular research. In the future, efforts to ascertain which data are demonstrative of an individuals' development of PCK should ideally include questions throughout the course of the study that explicitly request examples of reasoning behind methods, etc. to assist in the determination of a teacher's levels of PCK before further assumptions can be made. Because the data explored in this research was originally collected without PCK in mind, it is infinitely more difficult to draw conclusions about the causes behind the participants' varying levels of teaching effectiveness.

Alternately, the study findings may be correct; it is entirely possible that teaching experience is not influential in the development of all types of PCK. Since it is a highly specific form of knowledge, it is entirely possible that for this particular level of teachers, subject matter and form of professional development, other attributes that were not explored are more significant in PCK development.

PCK and Elementary Science Education and Research Results

Further obscuring the determination of the development of PCK in this particular study is the fact that the standardized lessons given to the students may very well fall under the category of "activities that work." Since the inquiry-based lessons allowed the students to learn with very little teacher guidance or input, and much of the guidance provided was scripted, it is possible that student content knowledge gains had very little to do with the quality of the teaching. The original study data included some cases in which there were teachers who had very small content knowledge gains themselves, but whose students managed large content knowledge gains. This suggests the possibility that what I presumed to be evidence of the development of PCK could be confused with or confounded by the impact of an effective activity.

Since "activities that work" play a role in the development of a teachers' PCK, but do not develop it fully, it is also conceivable that the varying levels of teacher effectiveness were caused by individual teachers' comfort level with and knowledge of how to use such activities instead of a true ability to pass on elementary science content to students. Thus, it is feasible that teachers who did not gain knowledge themselves found success because they had an existing framework in place allowing them to guide their students through an inquiry-based lesson without having to completely develop science PCK on their own, while teachers who were adept at learning themselves, but could not pass on their knowledge, may have simply had less familiarity with the use of such activities.

PCK and Professional Development in Elementary Science Education and Research Results

The original BCM study confirms existing research assertions that content-focused professional development can lead to both teacher content knowledge gains and student content knowledge gains, as indicated by overall score gains for all groups in both cohorts. It would seem, though, that the development of a teachers' elementary science PCK cannot be taken for granted. While it is possible that quality in-service education for teachers can aid in the development of PCK, the outcomes of this thesis research suggest that professional development that simply increases content knowledge or even a teacher's content-specific pedagogy (as was used in the third module – the video lesson demonstration), cannot, on its own, lead to the development of elementary science PCK. Professional development should specifically target the development of PCK in order to ensure that teacher's gains are effectively and consistently passed on to students.

Comparison of Methods within This Research

While proponents of stepwise methods might feel that such an analysis has a use in data explorations like this one, unless there are extreme time constraints, I would agree with King (2003) that the benefits of the technique do not outweigh its criticisms, and that best-subsets is a much more useful alternative. Best-subsets logistic regression, while time-consuming, provides the investigator with additional information even when no statistically significant results are discovered. If, for instance, a certain model had performed differently – either better or worse – than the competing models, it might be worth examining what potentially caused such differences. Forward stepwise logistic regression, on the other hand, provided no knowledge beyond the fact that none of the variables were accepted into the model, thus preventing the researcher from making further decisions about the direction of the exploration.

In the case of this research, best-subsets logistic regression was exceedingly useful; it allowed me to observe that all four independent variables performed similarly in the construction of an appropriate PCK model, which then prompted questions about the role of teaching experience in the development of PCK, a finding which may guide future research. The ability to compare models when results are not statistically significant is an attribute of best-subsets logistic regression that cannot be understated, and it is surprising that such a useful characteristic was not mentioned in the existing body of research.

Implications for Future Research

Despite the fact that none of the independent variables in this research demonstrated any correlation to what I believe was the development of PCK in the participants, the data set from the original study has many other facets that have yet to be explored, including a wealth of qualitative data. Future research may wish to investigate possible correlations between successful teachers and other demographic attributes, as well as attitudinal factors that may be gleaned from the rest of the data.

Beyond the scope of this research and dataset, it would be worth investigating the effect of including PCK-specific training within elementary teacher professional development, especially in the field of online professional development. There is much that can be gained from discovering the differences in the development of PCK through various online professional development methods. Additionally, the potential differences in teaching efficacy when teachers are cognizant of developing PCK, as compared to their peers who are not, should be examined in the context of online professional development as well.

Finally, further methodological studies exploring the wealth of information arising from best-subsets logistic regression, including the utility of non-significant data, would also be of interest.

Conclusion

Overall, though this study may not have satisfactorily answered the initial research questions, it has raised some intriguing new ones. For instance, how does one go about deciding whether or not PCK is actually being developed in any research? How would one identify the demographic factors that may play a role in elementary school teachers' development of PCK? How can these attributes and this knowledge contribute to, and eventually develop, effective professional development that can help teachers build and enhance their PCK?

Hopefully future research will address these and other questions regarding PCK, and improvement will continue to be made within the field.

References

Antonakis, J., & Dietz, J. (2011). Looking for validity or testing it? The perils of stepwise regression, extreme-scores analysis, heteroscedasticity, and measurement error. *Personality and Individual Differences, 50,* 409-415.

Appleton, K. (2003). How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice. *Research in Science Education, 33,* 1-25.

Appleton, K. (2006). Science pedagogical content knowledge and elementary school teachers. In K. Appleton (Ed.) *Elementary Science Teacher Education: International Perspectives on Contemporary Issues and Practice* (pp 31-54). Mahwah, NJ: Lawrence Erlbaum Associates.

Appleton, K. (2008). Developing science pedagogical content knowledge through mentoring elementary teachers. *Journal of Science Teacher Education*, *19*, 523-545.

Astin, A., & Denson, N. (2009). Multi-campus studies of college impact: Which statistical method is appropriate? *Research in Higher Education, 50,* 354-367.

Berry, A., Loughran, J., & van Driel, J. H. (2008). Revisiting the roots of pedagogical content knowledge. *International Journal of Science Education*, *30*, 1271-1279.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*, 3-15.

Clermont, C.P., Krajcik, J.S., & Borko, H. (1993). The influence of an intensive in-service workshop on pedagogical content knowledge growth among novice chemical demonstrators. *Journal of Research in Science Teaching*, *30*, 21–43.

Dede, C., Ketelhut, D. J., Whitehouse, P., Breit, L., & McCloskey, E. M. (2008). A research agenda for online teacher professional development. *Journal of Teacher Education*, *60,* 8-19.

DesJardins, S., & Lindsay, N. (2008). Adding a statistical wrench to the "toolbox." *Research in Higher Education, 49,* 172-179.

Gess-Newsome, J., & Lederman, N. (Eds.). (1999). *Examining pedagogical content knowledge: The construct and its implications for science education*. Hingham, MA, USA: Kluwer Academic Publishers.

Goodnough, K., & Hung, W. (2009). Enhancing pedagogical content knowledge in elementary science. *Teaching Education*, *20*, 229-242.

Grayson, J. P. (1997). Academic achievement of first-generation students in a Canadian university. *Research in Higher Education, 38,* 659-676.

Heller, J., Daehler, K., Wong, N., Shinohara, M., & Miratrix, L. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching*, *49*, 33-362.

Hill, H. (2009). Fixing teacher professional development. *The Phi Delta Kappan*, *90*, 470-476. Retrieved from http://www.jstor.org/stable/20446155

Kieffer, K., Reese, R. & Thompson, B. (2001). Statistical techniques employed in "AERJ" and "JCP" articles from 1988 to 1997: A methodological review. *The Journal of Experimental Education, 69,* 280-309. Retrieved from http://www.jstor.org/stable/20179991

King, J. E. (2003). Running a best-subsets logistic regression: An alternative to stepwise methods. *Educational and Psychological Measurement, 63,* 392-403. Retrieved from http://epm.sagepub.com/content/63/3/392

Loughran, J., Mulhall, P. & Berry, A. (2004). In search of pedagogical content knowledge in science: Developing ways of articulating and documenting professional practice. *Journal of Research in Science Teaching*, *41*, 370-391.

Loughran, J., Mulhall, P. & Berry, A. (2008). Exploring pedagogical content knowledge in science teacher education. *International Journal of Science Education*, *30*, 1301-1320.

Lumpe, A., Czerniak, C., Haney, J. & Beltyukova, S. (2012). Beliefs about teaching science: The relationship between elementary teachers' participation in professional development and student achievement, *International Journal of Science Education*, *34*, 153-166.

Mallows C_p equals number of parameters for all my Regression runs. (2010). In *IBM.com* Retrieved from http://www-01.ibm.com/support/docview.wss?uid=swg21481274

McCann, S., Short, R., & Stewin, L. (1986). Perceived teacher directiveness, student variables, grades and satisfaction: Aptitude x treatment interactions? *Instructional Science, 15,* 131-164.

Mulhall, P., Berry, A., & Loughran, J. (2003). Frameworks for representing science teachers' pedagogical content knowledge. *Asia-Pacific Forum on Science Learning and Teaching*, *4*, Article 2. Retrieved from http://www.ied.edu.hk/apfslt/v4_issue2/mulhall/index.htm#contents

Peng, C., Lee, K., & Ingersoll, G. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96,* 3-14.

Peng, C., & So., T. (2002). Logistic regression analysis and reporting: A primer. *Understanding Statistics, 1,* 31-70.

Peng, C., So, T., Stage, F., & St. John, E. (2002). The use and interpretation of logistic regression in higher education journals: 1988-1999. *Research in Higher Education, 43,* 259-293.

Rohaan, E. J., Taconis, R., & Jochems, W. (2009). Measuring teachers' pedagogical content knowledge in primary technology education. *Research in Science and Technological Education*, *27*, 327–338.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*, 4-14.

Thomas, E., & Galambos, N. (2004). What satisfies students? Mining student opinion data with regression and decision tree analysis. *Research in Higher Education, 45,* 251-269.

Thompson, B. (1989). Why won't stepwise methods die? *Measurement and Evaluation in Counseling and Development, 21,* 146-148.

Thompson, B. (2001). Significance, effect sizes, stepwise methods and other issues: Strong arguments move the field. *The Journal of Experimental Education, 70,* 80-93. Retrieved from http://www.jstor.org/stable/20152666

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies, Washington, D.C.

van Driel, J. H., De Jong, O., & Verloop, N. (2001). The development of preservice chemistry teachers' pedagogical content knowledge. *Science Teacher Education*, *86*, 572-589.

van Driel, J. H., Verloop, N., & Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, *35*, 673-695.

Wuensch, K.L.(2011). Binary Logistic Regression with PASW/SPSS©. Dr. Karl L. Wuensch's SPSS-Data Page, East Carolina University Department of Psychology. Retrieved from http://core.ecu.edu/psyc/wuenschk/MV/Multreg/Logistic-SPSS.pdf

Whitaker, J. S. (1997). *Use of stepwise methods in discriminant analysis.* Paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX. Retrieved from http://ericae.net/ft/tamu/STEPWIS.htm

Yoon, K. S., Duncan, T., Lee, S., & Shapley, K. (2008, March). *The effects of teachers' professional development on student achievement: Findings from a systematic review of evidence.* Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY. Retrieved from http://www.pdal.net/inc/docs/AERA%202008%20Paper_final_PD%20research%20review.pdf

Appendix A: Best-Subsets Logistic Regression Procedure

Adapted from King, 2003 and Mallow's C_p, 2010

with the help of Jason E. King

- Run a standard logistic regression of the full model –the model with all four variables included – saving the estimated predicted probabilities (pred).
- Create a new variable *z* by transforming the dependent variable (dvar) via the equation:

  z = log(pred /(1 – pred)) + ((dvar – pred)/(pred * (1 – pred)))

- Create the new variable *u* which is a case weight variable as follows:

  u = pred * (1 – pred)

- Run a linear regression predicting the new dependent variable, *z*, using the full set of independent variables. Include the case weight variable, u, so that weighted least squares estimates are obtained. (As a result of these modifications, the linear regression command will estimate logistic regression models.)
- Run a linear regression for each potential variable combination, recording the R^2 and adjusted R^2 values and enter in to the best-subsets regression table.
- Calculate Mallow's C_p values for each model with the following equation:

  Cp = (SSe/sigma_sq) + 2*p_star – C

  SSe = sum of squares error for the model

  Sigma_sq = mean square error of the full regression model

  p_star = the number of variables in the model (predictors plus intercept)

  C = the sum of the case weights for the model