DISSERTATION

PROFILES OF SUCCESSFUL GRADUATE STUDENTS USING ARTIFICIAL NEURAL NETWORKING TECHNOLOGY AND SIMULTANEOUS MULTIPLE REGRESSION

Submitted by

Joan L. Anderson

School of Education

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2001

UMI Number: 3013821

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

February 23, 2001

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY JOAN L. ANDERSON ENTITLED PROFILES OF SUCCESSFUL GRADUATE STUDENTS USING ARTIFICIAL NEURAL NETWORKING TECHNOLOGY AND SIMULTANEOUS MULTIPLE REGRESSION BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

____

____

____
Advisor

____
Co-Advisor

____
Department Head/Director

ABSTRACT OF DISSERTATION

PROFILES OF SUCCESSFUL GRADUATE STUDENTS USING ARTIFICIAL NEURAL NETWORKING AND SIMULTANEOUS MULTIPLE REGRESSION

The data from graduate student applications at a large Western University were used to develop profiles of successful graduate students, as defined by cumulative graduate grade point average. Two statistical models were employed and compared, artificial neural networking and simultaneous multiple regression. The sample was divided into 3,097 masters and 805 Ph.D. students' university wide who entered between Fall 1990 and Spring 1994. In addition, Non-parametric statistics, including Mann-Whitney U and Spearman Rho were used to determine whether each individual variable was significantly related to graduate grade point average. Non-parametric statistics were used due to the violation of normal distribution by the dependent variable, graduate grade point average.

Results of the artificial neural network and regression models for master's students yielded similar results indicating that the combination of the college the student was applying for, marital status, gender, GRE verbal and analytical scores, and residency region of students could predict 10%-12% of the variance in graduate grade point average. Results of the two models for Ph.D. students were not as similar. The best of five artificial neural networks used the combination of GRE analytical, quantitative and verbal scores; gender, marital status, age, residency region, and citizenship continent to predict 24% of the variance in graduate grade point average. Multiple regression indicated that college, GRE verbal and analytical scores, marital status, and age could predict 9% of the variance in graduate grade point average. Caution should be taken when interpreting the results of the ANN Ph.D. predictive models because two of the five neural networks performed worse than the regression.

Because of having more confidence in the master's data set, the conclusion was that artificial neural networking and simultaneous multiple regression provide similar results in determining which student characteristic variables provide the best profile of students with high graduate grade point averages. A better understanding of the interpretation of artificial neural network outputs in these types of research questions is recommended. In addition, maintaining accurate and complete historical databases for analyses is necessary.

Joan L. Anderson
School of Education
Colorado State University
Fort Collins, Colorado, 80523
Spring 2001

DEDICATED TO:

My Mother and Father

Thank you showing me “I could”

ACKNOWLEDGMENTS

The writer wishes to express her sincere appreciation to the following:

Dr. George Morgan, committee advisor, for being an inspiration to me since the first day I set foot on Colorado State University's campus, for your expertise in statistics, and your gentle way of saying "it's a good start."

Dr. Gene Gloeckner, co-advisor, for your dedication to the task at hand, from the day we first talked about artificial neural networks to the final product. It would have been difficult to remain dedicated without you.

Dr. Antigone Kotsiopulos, committee member, for your ongoing support and belief in me through all these years, your advice, and your friendship.

Dr. Barb Nelson, committee member, for your expertise in editing, your openness, and honesty.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

According to Peter Syverson, Vice President for Research and Information Services for the Council of Graduate Schools in Washington D.C. (personal communication, February 9, 2001), on the average, approximately 50% of Ph.D. students entering graduate school will actually complete the requirements and graduate. This number varies by university, college and department. Students in the hard sciences, such as Biology and Chemistry have higher completion rates than students in disciplines such as the Humanities and Social Sciences. Mr. Syverson attributed this difference to issues of support for students in varying disciplines. In addition, master's students tend to have a higher completion rate (70%) than Ph.D. students (50%).. Although there is little empirical data, Mr. Syverson said that these figures are based on the experience of Deans. Data found on the University of California, San Diego web site indicated that the percent of Ph.D. students finishing their degrees within ten years varied from 48% in the Social Sciences to 75% in the Health Sciences. According to the Graduate School office at Colorado State University, approximately 52% of students entering Colorado State University graduate program will graduate. This is similar to the national average quoted by Mr. Syverson of 50%.

Determining which of the students who apply to graduate school will be most successful is an arduous task at best. An even more difficult task is determining what characteristics of applying students can be used to predict success. Are we asking the right questions, looking at the right statistics, and using the correct procedures?

It is the intention of this study to further explore the identification of characteristics of successful graduate students through the use of statistical methods such as simultaneous multiple regression and artificial neural network technology. Artificial neural networking (ANN) is a predictive statistical model that uses historical data and complicated algorithms in large data bases to find relationships and trends, previously unknown to the researcher, to promote decision support (Foley & Russell, 1998).

Studies have been conducted to identify the characteristics of successful graduate students. Graham (1991) studied the correlation of graduate school of business entrance criteria with success in a Master of Business Administration program. Graduate GPA was used to measure the outcome variable, success. Enright and Gitomer (1989) conducted a study that attempted to identify the characteristics of a successful graduate student. Through interviews with 15 university faculty members, seven competencies where identified. These included the following: a) communication, the ability to share one's ideas, knowledge and insights with others; b) creativity, the ability to produce an unusual number of ideas or generate novel ideas; c) explanation, the giving of a reason or cause for some phenomenon or finding; d) motivation which includes commitment, involvement, and interest in their work; e) planning, the development of a procedure to reach some goal; f) professionalism, skills necessary to successfully accommodate the social conditions of a particular filed; g) and synthesis, the ability to organize information into complex knowledge structures. The Enright and Gitomer (1989) study contrasted with Graham (1991), who studied competencies that should be evident at the time of matriculation, such as grade point average (GPA) and graduate record examination (GRE) scores. Lipschutz (1993) stated that institutions need to try harder to identify the characteristics that are germane to success. Confronting the practices and procedures that define policies at the graduate school level may constitute a manageable series of small wins. One of the most challenging issues is the admissions process.

There are volumes of literature about admissions at the graduate level (Association of American Universities, 1998; Berg, 1993; Council of Graduate Schools, 1994). The results of these studies were inconsistent. Hackman and Price (1995) conducted a survey of all Ph.D. programs nationwide offering degrees in educational administration or educational leadership and found strong similarities between institutions with respect to admissions criteria. They concluded that the average doctoral program requires a minimum combined score of 1000 on the GRE verbal and quantitative exam scores, an undergraduate GPA (UGPA) of 3.0, a graduate GPA (GGPA) of approximately 3.23, a writing sample, a statement of goals, and a personal interview. In contrast, Clark and Palettela (1997), in their book written as a guide to new graduate students, stated that the admissions process and means of evaluation differ across universities, departments, and among faculty. These studies all focused on the admissions process and made recommendations for change based on the strengths and weaknesses within the studied program. There was little attention paid to the issues of success or failure of graduate students such as graduate grade point average or whether a student graduated or did not graduate.

However, there are several studies linking success in graduate school to the admissions process. Gorr, Nagin, and Szcypula (1994) conducted the most notable study. They compared an index used by the admission committee to predict graduate GPA with the predictability of artificial neural networks, linear, and stepwise polynomial regression. Gorr et al. sampled graduate students in a professional school. Results indicated that none of the empirical models were superior in prediction to the currently utilized admissions committee model for selecting graduate students for admission. Fiedler, Foldesy, Matranga, and Peltier (1993) conducted a study of 84 chairpersons of educational administration departments accredited by the National Council for the Accreditation of Teacher Education and found that 48% were not satisfied with the effectiveness of their admissions criteria related to determining success in graduate school. The Council of Graduate Schools (1992) stated that correlating student's admissions credentials with the determination of success or failure is an important factor in determining the success of an institution's admissions process. Other studies have examined the prediction of academic success in graduate school through the use of undergraduate data such as UGPA, type of undergraduate degree earned, and academic caliber of student's undergraduate colleges, concluding that UGPA was a predictor of success in graduate school (Graham, 1991; Hall, 1992). Still other studies utilized admissions data to predict graduate school grades and test scores (Fisher & Resnick, 1990; Gorr et al. 1994; Nelson & Nelson, 1995; Zwick, 1993). These studies determined that GPA and standardized test scores such as the GRE were predictive in determining graduate school grades.

Methodologies for the previously-mentioned studies include qualitative methods such as interviews and focus groups (Enright & Gitomer, 1989) and quantitative methods such as multiple regression techniques (Graham, 1991). Only one study was found which used artificial neural networks. Sadler and Hammerman (1999) conducted a five-year study of graduate admissions and concluded that maintaining or improving the quality of admissions decision making may be achieved by recognizing and using historical patterns. More research is needed incorporating the use of sophisticated, computerized, predictive models on historical graduate admission data to determine which admissions criteria are predictors of graduate school success.

Artificial neural networking (ANN) is a predictive statistical model that uses historical data and complicated algorithms in large data bases to find relationships and trends, previously unknown to the researcher, to promote decision support (Foley & Russell, 1998). This technology uses a computer model that simulates a biological neural network (VanEyden, 1996). Through the use of mathematical algorithms, the software accesses historical data and through a training process, learns to make predictions on new data sets based on the lessons learned from the training set.

Very few studies in the educational field have used artificial neural networks as a methodology. Among the few studies are Everson (1994) who used artificial neural networks as an approach to classifying students as proficient in algebra. Everson concluded that ANN technology could accurately classify these students into the appropriate algebra class based on historical data. Song and Chissom (1993) concluded in their study that the artificial neural network was powerful in predicting university student enrollment.

Studies have been performed to improve the selection process of graduate students. In addition, different methodologies have been utilized to facilitate this process. The results of these studies are inconsistent. Further research is needed in the area of graduate student success and in determining which types of models can accurately predict those students who are most likely to succeed.

Research questions

This study will address the following research questions:

Research question #1. Is there a difference in graduate GPA for master's and for Ph.D. students at Colorado State University between different levels or groups of selected admissions variables?

Research question #2. Is there a combination of admissions variables that predicts graduate grade point average better than any one variable alone and if so, what is that combination?

Research question #3. Do artificial neural networks perform better than traditional statistical techniques in predicting success in graduate school as defined by graduate grade point average?

Variables

The potential predictor variables were:

1. College indicates where the student enrolled. There are eight colleges at Colorado State University: Natural Resources, Veterinary Medicine, Business, Applied Human Sciences, Engineering, Liberal Arts, Natural Sciences, and Agriculture.

2. Department indicates the department or school in which the student enrolled within a college.

3. Ethnicity is the student's race.

4. Gender indicates male or female.

5. Marital status indicates single or married.

6. Entry year is the year the student was enrolled in graduate school.

7. Age at entry is the age of the student when he or she enrolled to graduate school.

8. Citizenship is the country where the student has indicated citizenship.

9. Original residency country is the country of residence reported by the student at the time of first application.

10. Original residency state is the state of residency reported by the student at the time of first application.

11. Undergraduate Colorado State University grade point average is the cumulative undergraduate points for grades received at Colorado State University.

12. Undergraduate credits completed is the total number of credits completed at the undergraduate level.

13. GRE Analytical is the Graduate Record Examination analytical test score.

14. GRE Quantitative is the Graduate Record Examination quantitative test score.

15. GRE Verbal is the Graduate Record Examination verbal test score.

16. GRE Total is the combined Graduate Record Examination test score. Computed by adding the three section scores.

Initially there were two dependent variables, success, defined operationally as the graduate grade point average (GGPA) and graduate or not graduate.

Definition of terms

Artificial neural networks are computerized models that simulates a biological neural network (VanEyden, 1996). Through the use of mathematical algorithms, the software accesses historical data and through a training process, learns to make predictions on new data sets based on the lessons learned from the training set.

Data mining is a computerized technology that uses complicated algorithms in large data bases, to find relationships, and trends, real or perceived, previously unknown to the researcher, to promote decision support. (Foley & Russell, 1998).

Graduate students are students pursuing a master's degree or Ph.D. at Colorado State University.

Matriculation. Students who are formally admitted to Colorado State University graduate school and enroll in the program.

Delimitations

This study focused on graduate admissions and did not investigate undergraduate success and admissions policies and procedures. Research indicated that there is a distinguishable difference between the two. Clark and Palattella (1997) compare and contrast the admissions process at the two levels. They state that undergraduate students seldom meet the admissions committee. They apply to an entire college or university and individual faculty have little say in the admissions process. According to Colbeck (1994), faculty decisions in admissions at the graduate level are a direct reflection of which students they prefer to work with and how the student's proposed research interests are in line with the interests of the unit or department.

The design of the proposed study was restricted to students and the admission data that were previously collected by Colorado State University. In order to obtain a large enough database, data were gathered on those students from Fall, 1990 to Spring,1994. The study also did not include Master of

Business Administration or Doctor of Veterinary Medicine students because the degree completion process differs greatly for them. These programs are typically offered in a cohort model, where students enter and progress through the degree process together. In addition, the requirements for degree completion are different. Final projects may not include a thesis, professional paper, or dissertation, but may require a project or residency at a clinic.

The statistical packages, Statistica, for neural networks and SPSS, were utilized for the data analysis. Several other statistical packages are available.

This study focused on only CSU graduate students and may not be generalized to other institutions. Conducting a nationwide study would be problematic. Different colleges within different universities have varying admissions criteria and standards, which would make comparison difficult.

Limitations and assumptions

An existing archival database was used giving the researcher no control over the data, how it was entered into the University database and how well it was maintained. Surprisingly, certain information, such as undergraduate GPA was not complete, including only CSU undergraduate courses.

Significance

The proposed research was designed to study the potential for the use of sophisticated computerized modeling techniques to the admissions decision-making process. A similar study, conducted by Gorr et al. (1994) used artificial neural network technology to predict graduate student GPA. That study focused primarily on the methodology and not as much on the variables that best predict success in graduate school. As mentioned previously, other studies have focused on the predictor variables but have utilized traditional statistical methods such as regression. Dawes (1971) called for more research concerning the determination of graduate success. The author further stated that decisions made through the use of computers are more systematic, economical, and human than decisions based on intuition.

There is a potential for a decrease in the amount of time spent on admissions decisions by implementing the use of ANN technology. Sadler and Hammerman (1999) stated that finding ways to save time without compromising the process is a reasonable goal.

Several studies have been performed to improve the selection of graduate students. Only one of these studies utilized ANN technology (Gorr, et al. 1994). In addition, different methodologies have been utilized to facilitate this process. The results of these studies are inconsistent (see Chapter 2 for a discussion of these studies). Further research is needed in the area of determining which types of models can accurately predict graduate student success. The purpose of this study is to use simultaneous multiple regression and artificial neural networking to determine which combination of admissions criteria best predict success in graduate school for the master's and Ph.D. students at CSU.

CHAPTER 2

REVIEW OF LITERATURE

In reviewing the literature, a few general areas in the study of admissions and success in graduate school emerged. There are studies that focus on the admissions process without an attempt at predicting success of the students admitted. There are studies that address the issue of success in graduate school without focusing on admissions. There are several studies that focus on the issues of success in graduate school as determined by admissions criteria. The following review will look at each of these areas separately. In addition, a discussion of data mining and artificial neural networks (ANN) will be presented. Finally, a look at those studies in the discipline of education, which used ANN technology, will be included.

Admissions

Studies have been conducted on the admissions processes at the graduate level. Areas that these studies covered include motivation and plans to apply, recommendations and best practices for graduate admissions, the impact on faculty, program evaluations, standardized testing, and affirmative action. The following will address each of these issues individually.

Motivation and plans to apply. There are volumes of literature that study admissions at the graduate level. Studies have been conducted that focus on the motivation and plans of students in applying to graduate school. Delaney

(1999) completed a graduate admissions study that focused on the motivation and enrollment decision factors of a newly redesigned Master of Business Administration (MBA) program. The sample was composed of 228 students who responded to a mailed survey and identified factors that influenced their choices of university and graduate program. Students with a high level of satisfaction with the admissions process indicated that prompt replies to requests for information, individualized attention, correspondence and telephone contact with the admissions office, and the visit to campus were all positive experiences.

Rajecki, Lauer, and Metzner (1998) conducted a study of Purdue psychology undergraduate students to determine if a correlation existed between plans to attend graduate school and knowledge of admissions requirements. They found that 65% of newly admitted freshman students had plans to attend graduate school. This number dropped as the students progressed through the program. Knowledge of admission requirements and GPA accounted for only 1% of the variance in plans to attend graduate school. They recommended that at the undergraduate level, all majors be advised of the routes and barriers to graduate education, and that those career opportunities with a baccalaureate degree be explored.

Recommendations and best practices. There have been several studies concerning recommendations and best practices in regards to the admission process. The American Association of Universities Committee on Graduate Education (1998) presented a report recommending best practices in graduate education. The following admissions recommendations were put forth:

Admissions decisions should be made with the intent of maintaining and improving the quality of graduate programs; departmental admission policies should include the goal of increasing opportunities for underrepresented groups within their graduate programs; and efforts should be made to admit international students while maintaining the domestic talent pool.

In a study designed to determine the current status of educational leadership doctoral programs nationwide, 127 universities responded to a survey which provided data on admissions criteria (Hackmann & Price, 1995). The average admissions profile included a minimum score of 1,000 on the graduate records examination (GRE), an undergraduate grade point average (GPA) of 3.00, a graduate GPA of 3.23, and a writing sample including a personal statement and an interview. New trends indicate evidence of portfolios, mentors, internships, field-based programs, and technology skills. It is interesting to note that eight of the responding universities use a point system that gave flexibility to students who did not meet the grade point average (GPA) or Graduate Record Examination (GRE) minimum standards.

The Council of Graduate Schools published a book addressing the essential aspects of good graduate admissions policies, procedures, and practices (Diminnie, 1992). A wide range of admissions functions and the advantages and disadvantages of each were discussed. Each institution must determine the best admissions processes to achieve quality and equity within their institution. In addition, it is imperative that faculty involved in admissions decisions be aware of the issues, laws, and regulations involved in their university or institutional admissions policies, procedures, and guidelines.

The Council of Graduate Schools published another report (Borchert, 1994) addressing commonly-accepted standards of good practice for masters of education programs in the United States. The recommended elements considered in admissions were application materials, including an application form, application fee, a cover letter, letters of recommendation, transcripts from all colleges and universities attended, standardized exam scores, TOFEL scores for international students providing proof of English language competency, and faculty review. The report stressed the need for well-established record-keeping systems. Graduate programs need accurate information to analyze their ratio of applicants to offers, offers to admissions, and admissions to actual enrollments. They need information on student profiles such as race, gender, and citizenship. It should also be possible to track student progress toward degree completion or non-completion. They further recommend that this type of analysis be available on an institution-wide basis.

Impact on faculty. One unique study addressed the issue of faculty and the impact that admission decisions have on their motivation and behaviors. Colbeck (1994) conducted a qualitative study of faculty at two different colleges to gain insight into faculty motivation and behavior. One college was a research-based institution while the other was a metropolitan state university. Admissions practices emerged as having particular importance to faculty at both types of universities. Researchers concluded that the type of student admitted matters to faculty and to their teaching and research work. In addition, students could contribute to research or leave the faculty with little time for research, depending on the caliber of student. Bright students motivate some faculty to teach with thought and depth while other faculty found satisfaction in helping students who were struggling with studies and grades. Overall results showed that highly-selective admissions policies are more supportive to faculty than more open admissions procedures.

Program evaluation. Perhaps the most common form of admission research focused on program evaluation and admissions procedures within certain institutions. Sadler and Hammerman (1999) studied the admissions process for the Harvard University Graduate School of Education. In this five-year study, the researchers explored whether the admission process could be accurately modeled using historical data. This study looked at the degree to which the committee members agree, and what level of bias existed that may compromise fairness in decision making. They concluded that care should be taken to preserve historical records for later analysis.

Multiple logistic regression models can be developed which monitor comparable merit of ratings categories, establish thresholds for consideration, and provide rankings of candidates based on historical probability of admissions. The admissions process at Harvard was a successive step process, beginning with initial ratings by admissions committee members and concluding with a group rating process. The regression models were used to determine the amount of variance explained by each successive stage of the decision process.

The initial ratings process explained 40% of the variance in the final selected cohort. The researchers concluded that the final group rating process was the weakest stage. It was also concluded that rater bias should be calculated and monitored to assure fairness to candidates.

In 1993, Berg conducted a qualitative study of graduate students in an educational leadership doctoral program at Northern Arizona University's Center for Excellence in Education. Six doctoral students in a research methods course participated in a focus group. One of the areas discussed included the strengths, weakness, and improvement suggestions for admissions. The students made the following suggestions for improvement of the admissions process: designation of an adviser at admissions; clear, itemized policies; informal portfolio review prior to admissions process; and interview guidelines for candidates to preview.

A ten-member task force on graduate curricula development at the College of Agriculture, University of Florida (1993) performed another study focusing on program evaluation. University faculty in the College of Agriculture were surveyed using a questionnaire. The task force found that of the 192 respondents, 79% were satisfied with the admission policies, however; faculty at the research and education centers off the main campus were slightly less satisfied (75%), than were those on the main campus (81%).

Standardized testing and affirmative action. When addressing admissions, one cannot ignore the issues of the impact of standardized testing and affirmative action policies on decision making. There have been studies addressing concerns around the use of standardized tests and the admissions decision-making process (Hawkins, 1993; Vaseleck, 1994; Wendel, 1991). These studies challenged the validity of the Graduate Record Examination (GRE) and Graduate Management Admission Test exams. Borchert (1994) reported that "it has long been recognized that some students do not perform as well on standardized tests, and these scores may not be an indicator of potential or talent for advanced study" (p. 32). Borchert further reported that the GRE board warns that scores should never be used as the sole criterion for admissions but rather used in conjunction with other measures of ability or scholarly promise.

Recent studies have attempted to restructure the GRE to increase its validity. Bennett, Morley, Quardt, and Rock (2000) investigated the functioning of computer-delivered responses for use in standardized testing. Enright, Rock, and Bennett (1998) examined the use of alternative item types and configurations to increase the validity of the GRE. Bennett and Sebrechts, (1997) proposed an alternative computer-based problem solving task for admissions assessment. Kezar (2000) in a report on higher education trends from 1997-1999, indicated that universities are moving away from traditional assessment methods and contemplating the use of portfolios, simulation, and case studies in making admissions decisions.

The issue of affirmative action received serious attention in the literature in 1996 (Browne-Miller, 1996; Cross, 1996; Garfield, 1996). There have also been several books published on the issue (Bowen & Bok, 1998; D'Souza, 1991; Sacks, 1978). In a report on public policy issues in 1999 and 2000, the

Association for Governing Boards of Universities and Colleges (2000) reported the following status of the affirmative action policy in higher education. The fact that race may be considered a plus in admissions appears to becoming increasingly vulnerable. In political and legal forums, opponents of affirmative action are active in systematic efforts to challenge admissions efforts and the proponents of affirmative action are gathering evidence to maintain the ability to use race and ethnicity in admissions decision-making. Kezar (2000) stated that there have been research studies on the recruitment and retention of underrepresented groups and concluded that what is needed now was a synthesis of the literature. It is not the purpose of this current research study to focus primarily on these issues but rather to look at all the admissions criteria collected and available for analysis.

Many institutions have formally studied their admissions policies and practices and many boards and councils have laid forth recommendations for best practices; however, studies are inconsistent in reporting the use of graduate admissions criteria across universities. Hackman and Price (1995) conducted a survey of all Ph.D. programs nationwide offering degrees in educational administration or educational leadership and found strong similarities between institutions with respect to admissions criteria. In contrast, Clark and Palettela (1997) in their book written as a guide to new graduate students stated that the admissions process and means of evaluation differed across universities, departments, and faculty. These studies all focused on the admissions process and made recommendations for change based on the strengths and weakness within the studied process. They paid little attention to the issues of success or failure.

Success

Fewer studies have been conducted to attempt to identify the characteristics of successful graduate students. Enright and Gitomer (1989) conducted a study designed to gain a broad understanding of the skills and characteristics attributed to success in graduate school. The authors felt that there was an absence of clearly-defined factors that contributed to success, the relative importance of each, their contribution to success, and how they interact. The researchers interviewed 15 graduate faculty members, predominately psychologists. Results revealed three major findings. First, the graduate education process was characterized as an apprenticeship which yields skills and characteristics that lead to success. Second, a list of these skills was developed and included communication, creativity, explanation, motivation, planning, professionalism, and synthesis. These competencies should be developed during the course of graduate study and may not be completely developed at the time a student enters graduate school. A third finding discussed using simulation testing as a means of demonstrating the competencies previously mentioned. They recommended the use of problem simulations or work samples as a data collection device. The use of simulation testing can provide a work sample that can be analyzed for indicators of successful performance. The authors provided examples of exercises that could be used to assess the development of these skills. These included structured background interviews, completion of an unfinished report, outlining plans for further research to clarify the results of a report or suggest new issues perhaps overlooked in the report, critiquing a paper or report, and peer-group discussions.

In 1987, Powers and Enright published an empirical study of the reasoning or analytical abilities that are required by graduate students for successful performance in graduate school. They postulated that analytical skill was important to successful graduate study, but that little research had been done in this area. Faculty from six different academic fields responded to a survey rating the importance of various skills on a five-point scale scale, one meaning "this skill is not relevant in my field of teaching" to five "there is a critically important difference between marginal and successful students with respect to this skill." Means and standard deviations were calculated for each survey question by field of study. The six fields of study were English, education, psychology, chemistry, computer science, and engineering. Analysis of variance was used to assess differences among the six fields. There were a large number of significant differences within the six fields indicating that each field placed a different level of importance on the different reasoning skills. There were also significant differences on the number of skills that were deemed important within each field. There were, however, general skills that were viewed as important across all disciplines. These include "reasoning or problem solving in situations in which all the information is NOT known" ($\underline{M}$ = 4.24). Reasoning errors like "accepting the central assumptions without questioning them" ($\underline{M}$ = 3.96), "being unable to integrate and synthesize ideas from various sources" (M = 3.96), and "being able to generate hypotheses independently" (M = 3.94) (Powers & Enright, 1987, p. 670). The researchers concluded that graduate admissions decisions were seldom based on the examination of that which is truly expected of graduate students and suggest that the type of information gathered in this research would be a better basis for assessment.

Reilly (1976) conducted a study using critical incidents to collect faculty ratings on graduate student performance. The researcher began by empirically defining sets of criteria which graduate faculty can use to base judgments of graduate student success. These performance factors included independence and initiative, conscientiousness, critical facility, enthusiasm, research and experimentation, communication, teaching skills, and persistence. In addition, data were gathered to determined the difference in importance that faculty members from different disciplines placed on each factor. Data were collected from 227 departments and 1,299 faculty across chemistry, English and psychology. The mean importance rating for each incident within each performance factor was calculated and correlations between disciplines were obtained. Importance ratings were most similar for psychology and chemistry ($r=.87$), and least similar for chemistry and English ($r=.27$) with the largest difference placed in the research and experimentation factor. The mean importance rating of this factor for chemistry was .46, English at .35, and psychology at .52.

Lipschutz (1993) stated that institutions need to try harder to identify the characteristics that are germane to success. Confronting the practices and procedures that define policies at the graduate school level may constitute a manageable series of small wins. One of the most challenging issues is the admissions process. Diminnie (1992) stated that "an important factor in evaluating the success of the admissions process is the determination of student success (or failure) as correlated with his or her admissions criteria" (p. 35).

Admissions and success

Fedler, Foldesy, Matranga, and Peltier (1993) conducted a study of 84 chairpersons of educational administrative departments accredited by the National Council for the Accreditation of Teacher Education and found that 48% were not satisfied with the effectiveness of their admissions criteria for determining success in graduate school. The Council of Graduate Schools (Diminnie,1992) stated that correlating student's admissions credentials with the determination of success or failure is an important factor in determining the success of an institution's admissions process. There are various studies in the literature that link success in graduate school and the admissions process (see Table 1).

Table 1
Summarization of Research Articles Studying Graduate Admissions and Success

Author(s) and Publication Year	Dependent Variable(s)	Significant Independent Variables(s)	Sample & Size	Statistical Analysis
Nelson & Nelson (1995)	Master's Degree Completion	9 hour GPA Quantitative GRE Analytical GRE Admission (probationary or not)	1,533 students mixed disciplines	logistic regression
Gorr, Nagin & Szczypula (1994)	Graduate GPA Accuracy of Decision Making	GPAs from prerequisites Total GPA Residency Transfer status	224 students at North Dakota State Bachelors of Science	Artificial Neural Nets Linear Regression Stepwise-polynomial Regression
Fiedler et al. (1993)	Academic & Career Success (no further def. Provided)	Undergraduate GPA English Proficiency Test Score	84 dept. chairs educational admin.	descriptive statistics Survey Data
King et al. (1993)	Graduate GPA	Quantitative GRE	Potential admittees in a recent class	Regression
Morrow (1993)	Graduate GPA	Analytical GRE Quantitative GRE Verbal GRE	171 students W. Carolina University	Stepwise Multiple Regression
Zwick (1993)	First-year GPA Final GPA	Undergraduate GPA Verbal GMAT Quantitative GMAT	5,000 MBA students	Bayes Regression Models
Hall & Bailey (1992)	First year GPA	MCAT Undergraduate science GPA College selectivity	420 medical students Dartmouth	Pearson Correlation
Graham (1991)	Graduate GPA	GMAT Undergraduate GPA	100 MBA students	Multiple Regression
Fisher & Resnick (1990)	First-year GPA	Total GMAT Undergraduate GPA	530 MBA students	Multiple Regression
Paolollo (1982)	Graduate GPA	Junior/Senior GPA GMAT Full/Part-time Attendance	220 MBA students	Step-wise Linear Regression
Youngblood & Martin (1982)	Graduate GPA	GMAT Undergraduate GPA	406 MBA students	Step-wise Regression
Kirnan & Geisinger (1981)	Master's Comp. Exam	All three GRE scores MAT	114 students A New York Univ.	Step-wise Multiple Regression
Brown & Weaver (1979)	Graduate GPA Degree Completion	Undergraduate GPA Verbal GRE Graduate GPA	129 Students Indiana U./journalism	Regression Discriminate Analysis
Jenkins (1972)	Admission Status (prob/reg)	Undergraduate GPA	107 students CSU Voc. Ed.	Chi-square
Dawes (1971)	Accuracy of decision making	Quantitative models are better than clinical judgment	111 students U of Oregon/Psych.	Multiple Regression Clinical Judgment

Accuracy of admissions committee decision making. Gorr, Nagin, and Szczypula (1994) conducted a notable study. They compared an admissions committee index, artificial neural networks, linear, and stepwise polynomial regression to determine the best predictor of success as measured by GPA. The researchers sampled three years of students from North Dakota State University's College of Pharmacy. Predictor variables included prerequisite college coursework grades in chemistry, zoology, mathematics, and English; total GPA; residency of North Dakota or not, partial transfer student (was some prerequisite coursework done outside of North Dakota State University), and total transfer student (all prerequisite coursework done outside of North Dakota State University). Although the artificial neural network outperformed the regression models by identifying non-linear patterns, it was not a better predictor of the dependent variable. Linear regression was the worst overall, and multiple regression performed the best overall, but results indicated that none of the differences in the empirical models were significant at the .05 level. Results also indicated that none of the empirical models were superior in prediction to the currently utilized admissions committee index. The Gorr et al. (1994) study is notable in that it is closely related to this research in methodology. Further comparative discussion is presented in the results and discussions chapters.

A previous study by Dawes (1971) took a similar look at admissions. The study focused on the decision making process rather than the validity of the admissions variables. Dawes reviewed the literature on admissions criteria as a predictor of success up to 1969 and concluded that the results "revealed disappointing predictive validity" (p. 180). Based on Dawes' conclusion, the focus of the study was on whether psychologists make better clinical judgments on applicants than can be made by a quantitative analysis of the data on which they base their decisions. Faculty rated 111 student files and a rating on a scale of one-five was assigned, five being outstanding. Dawes then took the same students with the variables undergraduate GPA, quality of undergraduate institution (QI), raw GRE scores, and the average rating made by the admission committee (AR) and using multiple regression, the correlation between undergraduate GPA, QI, and GRE scores with the average rating made by the admissions committee was determined. Results showed that quantitative analysis was a better predictor of student's success than the clinical judgment of the admissions committee. This contradicts Gorr, et. al (1994) which determined that admissions committees' judgments on graduate school success were superior to statistical models including artificial neural networking. Dawes further concluded that research is needed concerning the determinants of graduate success which has been answered to by a multitude of studies in the area of admissions and graduate student success.

Probationary vs. regularly admitted students. Jenkins (1972) studied 107 Colorado State University graduate students from Vocational Education. Graduates were classified by type of admission, either regular or probationary. Jenkins determined that there was a significant relationship between undergraduate GPA and the type of admission. A chi-square of 3.84 was needed for statistical significance and the study reported a chi-square of 16.10

(p<.001). In addition, the author studied whether certain biographical information, personal history, undergraduate GPA, teaching experience, and educational background were related to graduate admission type. The results indicated no significant correlation between admission type and any of the biographical factors. Lastly, Jenkins looked at whether undergraduate GPA and biolgraphical factors were predictive of graduate GPA and found no significant correlation.

Nelson and Nelson (1995) conducted a study of graduate students at a medium-sized university to determine which combination of admissions criteria best predicted success in graduate school for those students admitted on a probationary basis. They compared regularly-admitted students to those admitted on a probationary basis in the areas of GPA after the first nine hours of study; GRE verbal, quantitative, and analytical scores; and final graduate GPA. Results indicated that probationary students did not perform as well as regularly-admitted students on any of the measures of achievement such as GRE scores. They also found that a smaller percentage of probationary graduate students actually completed the degree requirements. The most significant variable for predicting completion of the degree was the GPA after the first nine hours of study for both the regularly-admitted students and the probationary students. Jenkins (1972) found that there was a relationship between undergraduate GPA and type of admission. Analytical and quantitative GRE scores were significant predictors for probationary students but not for regularly- admitted students. The logistic regression model accurately predicted 90% of the cases where the student was actually successful in completing the degree.

Graduate Record Examination scores and undergraduate grade point average. Brown and Weaver (1979) conducted a study of 129 graduate students from Indiana University's graduate journalism program. They studied whether undergraduate GPA, undergraduate major GPA, verbal GRE score, and quantitative GRE score were predictors of graduate GPA, and completion of the degree program. Of the four predictors studied, it was determined that although weak, undergraduate GPA was the best predictor of graduate GPA ($r=.27$). With all four variables combined, they could account for 26.5% of the variability in graduate GPA. Other major findings of the study concluded that a high overall undergraduate GPA and a high verbal GRE score increased the students' chances for high grades in graduate school. In addition, high graduate grades increased a student's chance for graduate degree completion.

Kirnan and Geisinger (1981) conducted a study to determine if certain admissions data - - the three GRE scores, Miller Analogies Test (MAT) score, and undergraduate GPA -could predict performance on a master's comprehensive exam. A student group of 114 from a large New York University were included in the study. Statistically significant correlations were found for all three GRE scores, GRE verbal ($r=.43$, $p<.001$), GRE quantitative ($r=.27$, $p<.01$), and GRE advanced ($r=.32$, $p<.01$). The MAT score was also a significant predictor ($r=.34$, $p<.001$). Undergraduate GPA was not a significant predictor at the .05 level of performance on the master's comprehensive exam.

Career success. A study by Fielder, Foldesy, Matrange, and Peltier (1993) sought to determine what admission criteria used by the National Council for Accreditation of Teacher Education accredited doctoral programs of educational administration were related to academic and career success. Department chairs from 84 programs nation-wide were surveyed, and they indicated that previous grades (30%) and English writing proficiency test scores (26%) were the most effective predictors of academic success. In addition, the study determined that almost half of the department chairs where not satisfied with the effectiveness of their admissions criteria in predicting success in graduate school (48%). The GRE test scores and letters of recommendation from personal sources where determined to be the least effective predictors of academic success. This study was based on the opinion of department chair persons who used personal experience as their motive for determining which measures are predictive of graduate school success. It is important to note that not all students applying to graduate school have English proficiency test scores. International students use this type of test score. In addition, the results of this study are not consistent with the literature that indicates GRE test scores to be accurate predictor's of success in graduate school.

For example, a study of the Department of Government at Harvard University (King, Bruce, & Gilligan, 1993) quantitative GRE scores were determined to be the best predictor of grades in graduate school, but grades were not good predictors of success in graduate school or professional success. The authors further concluded that admissions committees were able to make accurate predictions about which students would succeed if admitted. This study is contradictory to a previous study by Dawes (1971) who concluded that empirical models are better predictors of success than clinical judgments made by admissions committees. However, Gorr et al. (1994) later confirmed the King et al. results concluding that statistical models and artificial neural networking did not produce superior results to the admissions committee decisions.

A study at Western Carolina University (Morrow, 1993) sought to validate the admissions standards of their master's degree program in counseling. Four members of the faculty rated 171 graduate students as exemplary or marginal with respect to their demonstrated ability and/or potential to represent the program and profession. These ratings where subjective and faculty did not have access to the admissions records in making these determinations. The criteria for making these recommendations were not discussed in the article. The author then correlated the faculty ratings with the following admissions criteria: undergraduate GPA, GRE verbal scores, GRE quantitative scores, and GRE analytical scores. Undergraduate GPA correlated highest with the faculty ratings ($r=.42$, $p<.05$). The two GRE scores did not yield significant correltations. Additional tests were run using graduate GPA and GRE analytical scores (these score were not currently being used in the admissions decision-making process). Graduate GPA correlated significantly with faculty rating ($r=.79$, $p<.01$) and analytical GRE scores also correlated significantly with faculty ratings ($r=.60$, $p<.05$). Additional tests were run using graduate GPA as the outcome variable and it was determined through stepwise multiple regression that GRE analytical scores were the most powerful predictor of graduate GPA ($r=.57$, $p<.01$) accounting for 33% of the variance. The addition of undergraduate GPA increased the correlation to .64, accounting for an additional 8% of the variance. It was not surprising to learn that the admissions standards were changed as a result of the study. Minimum GPA and GRE scores were increased and analytical GRE scores were added to the admissions criteria.

Graduate Management Aptitude Test (GMAT) scores and undergraduate grade point average (GPA). A study conducted in 1982 by Youngblood and Martin concluded that GMAT scores and undergraduate GPA ($r=.51$) were the best predictors of success in graduate school as defined by graduate GPA. Later studies also attempted to determine the predictability of GMAT test scores on graduate student success. A 1990 study by Fisher and Resnick studied 530 incoming MBA students at Baruch College in New York. The study sought to relate factors used in the MBA admissions with first year GPA. Factors included sex, age, years since completion of undergraduate degree, date of acceptance, undergraduate GPA, undergraduate GPA of junior/senior years, total GMAT, verbal GMAT, and quantitative GMAT. Using stepwise regression, results indicated that the best predictor of first-year GPA is the combination of undergraduate GPA plus total GMAT ($r=.27$). No significant difference was found between using the quantitative GMAT score or the verbal GMAT score. Also, there was no predictive value in the junior/senior GPA.

Paolollo (1992) studied 220 graduates from a Master of Business Administration (MBA) program at a medium sized university and found junior/senior GPA to have predictive significance of graduate GPA at the .01 level. These results were inconsistent with Fisher and Resnick's (1990) results that indicated that junior/senior GPA had no predictive value. Paolollo's predictor variables included GMAT scores, junior/senior undergraduate GPA, number of hours required by the MBA program, age, undergraduate major, attendance (full-time verses part-time), veteran status, sex, undergraduate university, marital status, and other graduate degrees held. Paolollo (1992) found four variables to be significantly related to GGPA. Undergraduate GPA ($r=.35$, $p<.001$), GMAT score ($r=.26$, $p<.001$), attendance (full-time verses part-time) ($r= -.18$, $p<.01$), and gender ($r=.14$, $p<.05$) indicating that females had higher graduate GPAs. None of the other variables were significant at the .05 level.

Graham (1991) studied the correlation of graduate school of business entrance criteria with success in a Master of Business Administration program as defined by graduate GPA. A total of 100 students were studied, 50 who had taken the GMAT and 50 who had taken the Miller Analogies Test (MAT). The predictor variables included number of semesters in the program, undergraduate GPA, age, sex, ethnicity, marital status, GMAT scores, MAT scores, number of years since undergraduate degree, and the type of undergraduate degree. In the GMAT group, the GMAT score had a higher correlation with graduate GPA ($r=.41$) than any of the other variables. In a stepwise regression, the GMAT score was the first to enter the equation and accounted for half of the variance. The GMAT score was the only significant variable at the .05 level in the stepwise regression equation ($R=.41$). In the MAT group, MAT score, ethnicity, and

Undergraduate GPA had significant correlation with Graduate GPA. The first variable to enter the stepwise regression model was ethnicity, followed by Undergraduate GPA, and then the MAT score. The author noted that the strong relationship of ethnicity (R=.43) was interesting in that it was not consistent with the GMAT group. Graham mentioned that this study was consistent with previously cited research in which standardized test scores and undergraduate GPA are the most important predictor variables.

Zwick (1993) examined the degree to which Graduate Management Admissions Test (GMAT) scores and undergraduate GPA (UGPA) predicted first-year and final GPAs of students in a doctoral program in Business and Management. Over 5,000 students were studied using an empircal Bayes regression approach. The results indicated that UGPA was a better predictor alone than the GMAT verbal and quantitative scores. However, using all three predictors was more effective than any one alone. Final GPA was more accurately predicted than the first-year GPA. In comparing these predictors to a group of master's of business students, these predictors were more accurate at the master's level than the doctoral level. They attributed this difference to the increased selectivity of the doctoral program.

Hall and Bailey (1992) conducted a study of the first-year academic performance of graduate students at Dartmouth Medical School. The study looked at 420 students entering Dartmouth between 1982 and 1986. The purpose was to determine whether the admission criteria used for selection were predictive of the first-year academic performance. The criteria included the

Medical College Admissions Test (MCAT), undergraduate science GPA, and college selectivity (the academic caliber of the student's undergraduate college). Results indicated that college selectivity scores were predictive of undergraduate science GPA. College selectivity groups could account for 69.9% of variance in science GPAs; however, there was no significant difference between the groups for MCAT scores and first-year GPAs. The MCAT scores correlated significantly with first-year GPA, ranging from .334 to .469 ($p<.001$). The MCAT and undergraduate science GPA predictability on first-year GPA improved when college selectivity scores were included. Overall, the authors determined that the combination of the three variables were useful in identifying successful graduate students.

Conclusions

It is clear from the literature that there are numbers of studies on the predictability of admissions criteria related to graduate school success. The studies are surprisingly similar in design. Samples consisted of students from the particular school or department being studied. The predominant dependent or outcome variables utilized in these studies were first-year graduate GPA and overall graduate GPA. Only two studies used degree completion as a definition of success (Brown & Weaver; 1979; Nelson & Nelson, 1995). The differences lie in the significant independent or predictor variables, which were highly inconsistent. Some studies found that undergraduate GPA was a significant predictor of graduate success (Brown & Weaver, 1979; Youngblood & Martin, 1982), while others claimed the inverse (King et al. 1993; Kirnan & Geisinger, 1981). Although standardized test scores were a common outcome, there is disagreement on which standardized scores were the better determinant. These results could be due in part to the fact that different schools and departments utilize different GRE scores in making admissions decisions. Quantitative GRE scores were significant in Nelson and Nelson (1995) and King et al. (1993). Verbal GRE score was significant in Brown and Weaver (1979). Analytical GRE scores were significant in Nelson and Nelson and Morrow (1993).

Similar disagreement was seen with the use of quantitative, verbal, and total GMAT scores. Sample sizes ranged from 84 (Fiedler et al. 1993) to 5,000 (Zwick, 1993) with the majority of the studies ranging from 100-500 graduate students. Methodologies covered quantitative methods such as correlation (Hall & Bailey, 1992), and multiple regression techniques (Graham, 1991; Kirnan & Geisinger, 1981; Youngblood & Martin, 1982). Logistic regression was used by Nelson and Nelson (1995) and discriminate analysis in the study by Brown and Weaver (1979). The use of logistic regression was most likely due to the dichotomous dependent variable of complete or non-complete. Only one study used artificial neural networks (Gorr et al. 1994). More research is needed incorporating the use of sophisticated computerized predictive models on historical graduate admission data to determine which admissions criteria are predictors of graduate school success. Sadler and Hammerman (1999) concluded that maintaining or improving the quality of admissions decision making might be achieved through recognizing and using historical patterns.

Data mining and artificial neural networks

Artificial neural networks are a form of data mining. Data mining is a relatively new computerized technology that uses complicated algorithms to find real or perceived relationships and trends in large databases, previously unknown to the researcher, to promote decision support. Data mining is currently in a state of growth. With new products and consumer demand on the rise, the market for software utilizing data mining is expected to expand from its current $3.3 billion to $8.4 billion by the year 2000 (Foley & Russell, 1998).

Data mining has been referred to as a statistical process of analyzing data stored in a data warehouse (Decker, 1998). A data warehouse is an extensive data repository consisting of information from all facilities of an organization's operations, including external sources, that is maintained to support decision-making. Data within the warehouse is manipulated to create easy access by data mining tools. Data mining is possible, but more difficult without a data warehouse. Smaller, downsized versions of data warehouses can be created. These are known as data marts and focus on one particular area of a database such as credit card users (Decker, 1998). Data marts are less expensive and can operate in a much smaller environment.

There are many types of data mining techniques. These include market basket analysis, memory-based reasoning, cluster detection, link analysis, decision trees and rule induction, and artificial neural networks (Berry & Linoff, 1997). Artificial neural networking is a predictive form of data mining that through the use of historical data uses complicated algorithms to determine relationships and trends in large data bases (Foley & Russell, 1998).

An ANN is a model that simulates a biological neural network (VanEyden, 1996). They are models of the interconnections in the human brain adapted for use on digital computers. In general, ANNs learn from a training set of data, generalizing patterns within the data for the purpose of prediction (Barry & Linoff, 1997). There are three basic steps in developing an ANN. First, the network is trained with existing data, then tested with data composed of known outputs. Testing is a process of confirming the predictability and accuracy of the previously trained network. Finally, the network is used to perform predictions on data with unknown outputs.

Examples of areas where ANNs have been applied successfully include the medical field where ANNs have been used to determine predictive patterns so that the appropriate treatment can be prescribed. Lapedes, Steeg, and Farber (1994) used ANNs to predict new protein structures and studied DNA sequencing (Lapedes, Barnes, Burks, Farber, & Sirotkin, 1988). ANNs have been studied in the assessment of patients' response to drug treatments (Valafar & Valafar, 1999), drug dose prediction (Lada, Brier, & Zurada, 1999) and heart rate variability (Bezerianos, Papadimitriou, & Alexopoulos, 1999). ANNs are used in stock market prediction by technical analysts, credit assessment to identify characteristics of applicants for classification as a good or bad credit risk, and foreign currency exchange rate forecasting (Pandya, Kondo, Talati, & Jayadevappa, 2000). Additional uses include engine management, where ANNs can monitor engine functions and aid in achieving goals such as minimized fuel consumption (VanEyden, 1996), and flood occurrence predictions (Bodri & Cermak, 2000). A more detailed description of the internal functions of the ANN will be outlined later. First, it is important to briefly discuss the working of the biological neuron to aid in understanding the function of the ANN.

The human brain, specifically the cerebral cortex, is composed of many layers of interconnected neuron cells. Within the cerebral cortex, characteristics such as intelligence, interest, awareness, and the abilities to adapt and learn reside. There are over one hundred billion neurons in the human brain (see Figure 1). These neurons are specialized nerve cells which transmit information to other neurons or other parts of the body and are made up of three parts. The soma contains the cell nucleus and other activity support systems. In the soma, incoming signals are added up over time until the soma decides when and how to respond to these inputs. When the input summation process reaches a certain threshold, the neuron will fire, if that threshold is not reached, the neuron remains inactive. The dendrites, surrounding the soma are the receptors for signals generated by other neurons creating a sort of network. The axon is the outgoing connection. When the neuron fires an electrical impulse, it travels down the axon to the boutons, which are connected to other dendrites in other neurons (VanEyden, 1996).

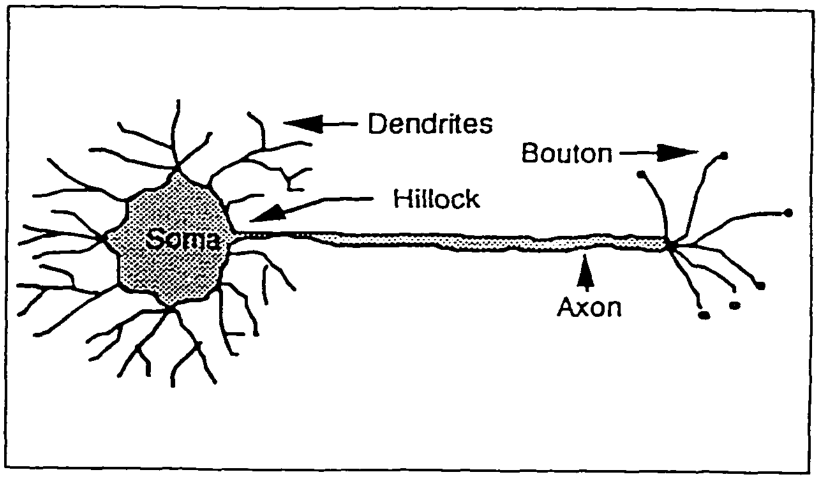

Figure 1. A biological neuron (VanEyden, 1996, Pg. 25).

How does the human brain learn? VanEyden (1996) discusses four schools of thought that are worth noting.

1) The traditionalist: the human brain learns because of mental states, not physical states of the mind.
2) The theory of reduction: The integration of psychological and neurobiological theories. Reductions are explanations of phenomena in one theory described in terms of a more basic theory.
3) Connectionism theory: Learning is a matter of making connections between stimuli and responses, where a response is any behavior and stimuli is any input that affects this behavior.
4) The theory of cognition: Views the brain as a connectionist model characterized by connections with different strengths between processing units. Neural networks can be used to test this cognitive

theory or more simply, to test current ideas based on biological research and mathematical theory.

How does an ANN function? In its simplest form, an ANN is composed of three layers (see Figure 2). An input layer, where the network receives data; a hidden layer, which performs functions analogous to the biological neuron; and an output layer, which collects features discovered and produces responses. ANNs can be simple models with few input units, one hidden layer, and a single output unit, or they can be composed of numerous hidden layers and numerous output units. A further explanation of the steps that an ANN completes in the training process will be explained in Chapter Three.

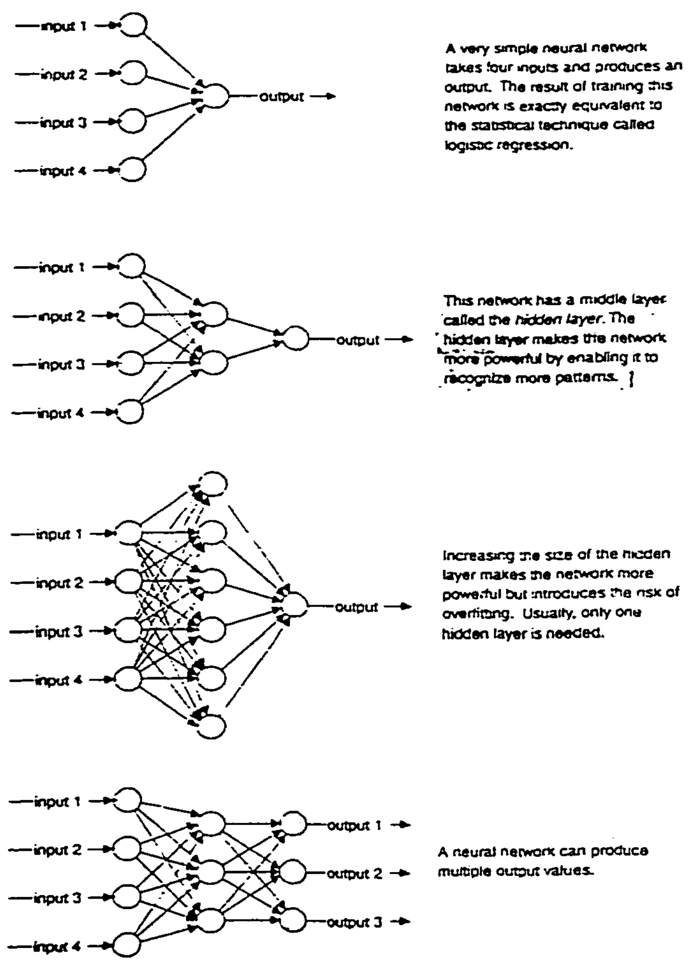

Figure 2. Four examples of artificial neural networks (Berry & Linoff, 1997, Pg. 296).

The advantages of artificial neural networks

Using artificial neural networks on research has its advantages. Discussions of the advantages of power and non-linearity, learning ability, ease of use, control of complexity, and accurate results are all present in the literature. Each of these is discussed in the following section.

Power and non-linearity. Easily, the most important advantage of ANN's is their non-linear power. ANNs are sophisticated modeling techniques that can perform extremely difficult and complex functions. Traditionally, linear models have been the most commonly-used techniques to handle predictive functions. However, often this assumption of linearity was not valid. ANNs are non-linear modeling tools, more in tune with real world situations (Statsoft, 1998).

The ability to learn. Through the use of complicated algorithms, ANNs have the ability to learn from experience, much like the biological neuron. By using data with known outputs, the network can run through what is termed a training mode and actually learn to make accurate predictions based on new data introduced (Zang, Patuwo, & Hu, 1997).

Ease of use and availability. According to Statsoft (1998), the level of knowledge required to successfully use and apply neural networks is much lower than that needed to apply traditional statistical techniques. ANN software packages are becoming increasingly more available and affordable in off-the-shelf applications. Users who simply want to load the software, run through a simple tutorial, and trust the output to be valid, will find the system fairly easy to use compared to traditional methods.

Control of complexity and versatility. ANNs can perform fairly simple modeling functions such as the basic three layer network described earlier, up to extremely complex networks with a large number of input units and multiple hidden layers. The complexity of the network can be easily adjusted and several examples can be run on the same data set to optimize the results. ANNs can easily be adjusted to perform prediction, classification, or even clustering functions. This control is one of the reasons that ANNs are increasing in popularity (Berry & Linoff, 1997).

Accurate results. ANNs have proven themselves in industry to produce accurate results (Berry & Linoff, 1997). They can perform well in very complicated domains such as time series designs and fraud detection.

Disadvantages of artificial neural networks

Although the literature supports many advantages to using artificial neural networks, there are also disadvantages. These include the nature of the input data, the inability for ANNs to explain the results, the possibility of ANNs converging on inferior results, and the size of the database needed. Each of these disadvantages is discussed in greater detail in the next section.

Nature of input data. ANNs will only accept numerical input data. All categorical data must be recoded. Input data must also be massaged or transformed to usually values between 0 and 1. This requires additional time, CPU power, and disk space. The choice of transformation may also effect the outcome of the ANN. Fortunately, most software applications provide data transformation modules to help users through this process (Cabena, et al. 1998; Berry & Linoff, 1997).

Inability to explain results. The inability to explain results has been claimed to be the biggest disadvantage of ANNs (Berry & Linoff, 1997). If the output must be explained, such as denying a loan or admissions to graduate school, ANNs may not be the tools of choice. It is recommended that ANNs be used when results need to be acted on, not explained. The upside to this disadvantage, is that ANNs can provide sensitivity analysis, explaining which input variables where more important then others. Sensitivity analysis is one way of over-coming the lack of ability to explain the rules.

ANNs may converge on an inferior solution. Converging on an inferior solution is also referred to as finding the local minimum instead of the global minimum. As the network trains, its goal is to minimize the difference between the current error and the desired error. Even though ANNs can be adjusted to lower the error, it is never certain whether the error could be lower. To further understand local verses global error, the concept of error surface provides insight. Each of the weights and thresholds of each of the inputs are a point in space. These errors can be plotted to form an error surface. The objective in training an ANN is to find the lowest possible point on that surface. As the ANN trains, it is basically exploring this surface. When it comes to an area where an additional run of the network either increases the error or no change is detected, it assumes it has reached the minimal error. This point on the surface could simply be a depression in the surface and not the lowest possible point (Statsoft, 1998).

Size of database. ANNs require very large databases of hundreds, sometimes thousands of cases. Several hundred are needed to train the network and hundreds more to test the network before actual predictions can take place. This could be prohibitive to certain studies where this amount of data is not available (Statsoft, 1998).

As was indicated by the previous section of the literature review on admissions and success in graduate school, the most common methodology utilized by researchers is multiple regression. Snyder (1996) stated that one of the easiest places to integrate neural networks is to follow on the study of regression. Snyder stated that the goal of regression is to determine a functional relationship between a dependent variable and one or more independent variables. The researcher hypothesized a relationship and then used a computer-based software package, such as SPSS, to determine the best fit for that regression model. In contrast, a neural network model is presented with independent data as input and dependent data as output and is trained until it can accurately predict the output data given the input data. The network is then given new input data that it has not been exposed to and it proceeds to predict an output value.

Reasons why artificial neural networks are superior to traditional statistical methods

Despite the disadvantages of artificial neural networks, there are researchers who maintain that these networks are superior to traditional statistical methods. The reasons for superiority fall into four major categories, data-driven; self-adaptive models; generalizability; general and flexible functional forms; and non-linearity. Each of these is discussed in the following section.

ANNs are data-driven, self-adaptive models. As a result of the data-driven, self-adaptive ability of ANNs, there are few assumptions about the models for the problems that are being studied. They learn from a training set composed of examples and have the ability to detect subtle relationships among the data regardless of whether these relationships are unknown or hard to describe (Zhang, et al. 1997). ANNs are well suited to finding solutions to problems that require knowledge that is difficult to specify, providing enough data is available. The ability to find solutions coupled with the ability to learn makes these very useful tools in determining what variables indicate success in graduate school. Zhang et al. stated that it is better to have data than good theoretical guesses.

ANNs can generalize. After they have learned through the training process, ANNs can often infer the unseen part of a population even if there was missing or "noisy" data in the training sample. Prediction of future behavior (the unseen part) is performed from examples of past behavior. This aspect of generalizability is ideal in answering any predictive question.

ANNs have more general and flexible functional forms. ANNs have more general and flexible functional forms than the traditional statistical methods can effectively deal with (Zhang et al. 1997). All predictive models assume there is a relationship between the input (independent variables) and outputs (dependent variables). Traditional statistical methods are limited in their abilities to identify these underlying relationships. With the flexibility of ANNs ability to determine weights of inputs, number of hidden layers, and learning algorithms, the ability to uncover relationships is greatly enhanced. Relationships between variables that can predict success in graduate school may be discovered through the use of ANNs that would be difficult, at best, to determine through the use of traditional methods.

ANNs are non-linear. Traditional statistical predictive models have by-and-large been linear in nature. However, more often than not, the underlying relationships are non-linear in nature. Real-world systems are often non-linear (Zhang et al. 1997). There are non-linear models, which have been developed to address this issue. However, these methods are still limited in that some relationship must be hypothesized with little knowledge of the actual underlying law. Oftentimes there are numerous non-linear relationships which traditional non-linear models are not capable of detecting. The data driven aspect of an ANN makes it capable of performing non-linear modeling without any prior speculation into the relationships that exist. Non-linear modeling make ANNs a more flexible and general tool for prediction.

Zhang et al. (1997) conducted a review of the literature from 1989 to 1995 which compared ANNs to other traditional statistical methods and the results were mixed. Of the 24 studies examined, 11 compared ANNs to regression techniques. Three studies found ANNs to be inferior to regression techniques in predictability; three studies found ANNs to be no better; and five studies reported that ANNs were superior to traditional regression techniques in predictive functions. Similar results were found in a series of studies on predicting scores for an AIDS risk test (Chance, MacLin & Lykins, 1993; Lykins & Chance, 1992b; MacLin, Chance, & Lykins, 1993). All three studies confirmed that ANNs outperformed regression in AIDS risk predictability.

Research in the educational field and artificial neural networks

Very few studies in the educational field have applied artificial neural network technology in their research. Baker and Martin (1998) took an interesting approach to using neural network technology to further understand the underlying brain functions that take place in learning. They examined the way that neural networks are formed during the training phase and use that information to postulate how neural connections are formed in the human brain during the learning process. The researchers took an interesting approach, in that previous literature uses the brain to understand how a neural network functions, not the inverse.

In a study of the Italian schooling system (Carbone & Piras, 1998), ANN technology was implemented into the public school system. Its purpose was to predict which students were at risk of failure, to optimize an intervention strategy for those "at-risk" students, to verify the effectiveness of the interventions, and to produce scenarios that determine optimum strategies of intervention. The project is called the "Palomar Project."

Clariana (1998) wrote a paper describing ANN technology applications in curriculum design and classroom data management. The study suggested several applications for use: automatic indexing of textbooks allowing teachers to locate specific pages in tests that support lesson plans; cross-referencing any two sets of standards, like curriculum and state mandated tests to verify that curriculum meets the state test requirements; identifying videos, books, Internet sites and instructional software to create a resource-based lesson directly supporting a teachers thematic unit.

An empirical study conducted by Everson (1994) used ANN technology to make classification or placement decisions based on proficiency in algebra. A look at prior academic achievement, math test scores, and test anxiety was used to determine proficiency in basic algebra. In addition, the study went on to assess whether ANNs outperformed traditional statistical techniques. It was concluded that ANNs outperformed these traditional methods and may lead to higher rates of classification accuracy.

Another empirical study conducted by Song and Chissom (1993) used ANN technology to predict enrollments at the University of Alabama. They concluded that ANNs were effective when applied to forecasting enrollments.

Gorr et al. (1994) used incorporated ANN technology in a study to determine whether linear regression, stepwise polynomial regression, ANNs or traditional admissions committee indexes were the best predictors of graduate GPA. The researchers sampled three years of students from North Dakota State University's College of Pharmacy. Predictor variables included prerequisite college coursework grades in chemistry, zoology, mathematics, and English; total GPA; residency of North Dakota or not, partial transfer student (was some prerequisite coursework done outside of North Dakota State University), and total transfer student (all prerequisite coursework done outside of North Dakota State University). Although the artificial neural network outperformed the regression models by identifying non-linear patterns, it was not a better predictor of the dependent variable. Linear regression was the worst overall, and multiple regression performed the best overall, but results indicated that none of the differences in the empirical models were significant at the 95% level. It was concluded that none of the empirical methods had better predictive capabilities than the traditional practitioner index.

Dawes (1971) concluded that decisions made systematically with the aid of a computer do not mean that the decisions have been "dehumanized." Dawes states that the conclusions drawn in this research "indicate that such decisions may be less capricious and more valid than those made by the decision maker relying on their own intuitions. Such decisions are more human" (p. 187).

It is the intent of this research to further the Gorr et al. (1994) study of the use of ANNs for predicting success in graduate school through the examination of admission criteria.

CHAPTER 3

METHODOLOGY

An ex post facto study of Colorado State University student admissions data was conducted to determine the specific variables that were predictive of success in graduate school. Success was defined by graduate grade point average (GGPA). An artificial neural network (ANN) computer program, Statistica, was utilized to analyze the data. ANN technology is one form of data mining that utilizes archival data to extract a decision rule from a sample of data to apply to new data. Sadler and Hammerman (1999) stated that the attempt to recognize and use historical patterns in data can improve the quality of decisions and reduce the time required by the decision-making process. A more detailed discussion of data mining and ANN technology is found in Chapter Two. Studies have been conducted using ANN technology for predictive purposes (Everson, 1994; Gorr et al. 1994; Song et al. 1993). The most notable is a study conducted by Gorr et al. (1994) using ANN technology to determine graduate success as measured by graduate GPA. SPSS was also utilized to run the traditional statistics, such as descriptive statistics, Kruskal-Wallis, Spearman rho, and multiple regression.

Participants and site

The original sample was composed of 5,206 students who entered Colorado State University Graduate School between Fall 1990 and Spring 1994. Due to the nature of artificial neural networks (ANN), a large sample size is preferable to maximize the predictive capability of the model. Wiess and Kulikowski (1991) stated that a sufficiently large training set will provide accurate performance measures. The sample was not a random selection of students, but rather all graduate students who entered the university graduate school between Fall 1990 and Spring 1994.

Data acquisition and procedure

There were several databases for graduate students available at Colorado State University. It was the original intent to utilize the database from the Graduate School. However, after meeting with a representative from the Graduate School, it was determined that undergraduate GPA was not available. A second source was available from the Office of Budget and Institutional Analysis. After two meetings with the database manager, it was also determined that the undergraduate GPA variable was not available. The next step was to contact the main university database. The university database manager told the researcher that the undergraduate GPA variable would be available in these data. Based on that information, the main university database was chosen. In addition to the undergraduate information, several other potential predictor variables found in the existing literature were said to be available. Much to the disappointment of the researcher, after human subjects approval, proposal approval, acquisition, and examination of the database, undergraduate GPA and undergraduate credits completed referred to only those students who had taken undergraduate courses at Colorado State University and over three-fourths of the undergraduate GPA data were missing. Based on the review of literature and an examination of the variables available in the university database, Table 2 indicates dependent and independent variables that were initially chosen.

Table 2
Variables, Levels, and Percent of Complete Data

Variable	Level of Measurement	# of Levels	Mean	Percent Complete
Independent				
College	Nominal	8		100%
Department	Nominal	14		100%
Ethnicity	Nominal	6		84%
Gender	Nominal	2		100%
Marital Status	Nominal	2		98%
Age @ Entry	Interval		31	99%
Citizenship Country	Nominal	106		99%
Residency State	Nominal	52		100%
UG CSU GPA	Interval		3.3	28%
UG CSU Credits	Interval		54.5	36%
GRE Analytical	Interval		566	86%
GRE Verbal	Interval		510	86%
GRE Quantitative	Interval		589	86%
GRE Total	Interval		1664	86%
Dependent				
Graduate/Not Graduate	Nominal	2		100%
GRAD GPA	Interval		3.7	99%

From this list, the following decisions were made about which variables to use. College and Department reported similar information. To reduce the number of variable levels, the variable college was used. There also appeared to be a high level of multicollinearity between citizenship, college, and ethnicity. In addition, the variable ethnicity was only 84% complete. It was decided to eliminate ethnicity as a variable. Undergraduate GPA and credits were also eliminated due the amount of missing data. GRE total was reporting the same information as the individual GRE scores. It was decided to look at the GRE scores on an individual basis rather than as a total to enrich the level of analysis on GRE scores. Finally, in looking at the dependent variables, the percentage of students who graduated verses those that did not was far removed from the national average. In the chosen database, 84% of the students graduated, while the national average was approximately 50%. There was no explanation as to why these two percentages were so different, so it was decided to eliminate the graduate/not graduate variable as a potential outcome and use the gradaute grade point average (GGPA).

Two variables were recoded to reduce the number of levels. Citizenship country was recoded to citizenship continent resulting in 6 levels and residency state was recoded to residency region resulting in 6 regions. Table 3 reports the final list of variables used in the study.

Table 3
Variables, Levels, and Percent of Complete Data

Variable	Level of Measurement	# of Levels	Mean	Percent Complete
Independent				
College	Nominal	8		100%
Gender	Nominal	2		100%
Marital Status	Nominal	2		98%
Age @ Entry	Interval		31	99%
Citizenship Continent	Nominal	6		99%
Residency Region	Nominal	6		100%
GRE Analytical	Interval		566	86%
GRE Verbal	Interval		510	86%
GRE Quantitative	Interval		589	86%
Dependent				
GRAD GPA	Interval		3.7	99%

Human Subject's approval and permission from the data warehouse manager were obtained prior to meeting with the programmer to obtain a copy of the data. Once obtained, the data were loaded into Statistica, a computer-based statistical program and recoded to meet the parameters of the software and merge groups of students which represented very small numbers within certain levels of certain variables.

The original sample size of 5,206 was reduced to 3,902 cases (or students). Eliminated were those cases that were non-degree seeking (postdoctoral students). Also, all Master of Business Administration (MBA) and Doctor of Veterinary Medicine (DVM) cases were eliminated due to the unique nature of these programs. The data were then divided into master's students and Ph.D. students for comparative analyses. This data set was then transferred to SPSS, another statistical computer program, to run frequencies and other descriptive statistics. A detailed discussion of descriptive statistics is available in Chapter Four. SPSS was used instead of Statistica because of the familiarity the researcher had with its functionality.

Data analysis using artificial neural networks

The software programs Statistica and SPSS were used to analyze the data. Initially, descriptive statistics and frequencies were run to further define the sample. The next step was to use the ANN technology available in Statistica to develop a predictive model. The Statistica program automatically randomly divides the sample into three subsets. A larger group to train the network, one to test the network, and one to verify the test. Weiss and Kulikowski (1991), to insure accurate testing, recommend using this verification technique on large to medium sample sizes. The larger group was used to train the network. ANN technology has the ability to "learn" from a training sample. Learning is defined as "choosing or adapting parameters within the model structure that work best on the sample at hand and other samples like them" (Weiss & Kulikowski, 1991, p. 4). The ANN model chose the variables that were the best predictors of success. The smaller groups were used to test and verify the resulting model.

There are seven steps outlined by VanEyden (1996) which explain the process that an ANN completes to perform the training process (see Figure 3 for an indication as to where each step is initiated within the ANN). Each of these will be discussed in detail.

Step 1) Computation of the net weighted output:

Each input unit or independent variable is assigned a weighted value. Different weights can be assigned to the same input unit based on the number of hidden layers within a network. The weights of all the input variables are summed and a net weight is calculated. The formula is as follows:

$$Ui = \sum_{j=1}^{n} (Xi \,.\, Wij)$$

Ui = net weighted input
i = current processing unit
j = processing unit connected to i
Wij = the weight between j and i
Xi = originating processing unit

Step 2) Convert net input to an activation level:

This is equivalent to the excitement level of a biological neuron. Each neuron has an activation level that tells the neuron what to do with the signal after the weights have had their effect. When the activation level reaches the threshold value, the neuron will fire and pass the net weighted value through a transfer function. Small changes in inputs, when the combined inputs are near the threshold value, can have large effects on the output. On the other hand, large changes in the inputs of the unit can have little effect on the output when the combined units are far from the threshold. This property gives ANNs their non-linear behavior (Berry & Linoff, 1997).

Step 3) Transfer of the weighted sum:

The product of the weighted sum is transformed into a working or transfer output. The weighted sum of the effective inputs is computed and transformed by a transfer function. The transfer function defines how the activation value is converted to an output value. There is disagreement in the literature with the actual number of transfer functions and their definitions. However, there is agreement that the choice of transfer function depends on the nature of the data. If the dependent variable is nominal, a transfer function resulting in values of 0 and 1 is preferred. If the output is continuous, then a transfer function scaling outputs between 1 and -1 is preferred (VanEyden, 1997; Statsoft 1998). Statistica, the neural network software utilized in this research, supports two types of transfer functions, scaled and nominal.

Step 4) Computation of the competition function of the output function:

This function allows the output units to compete to determine which will be active. There can only be one or two winners. The one or two winners will be chosen on the basis of the highest value of the product after applying the transfer function to the weighted sum value. This is where the learning process takes place. Only the winners will learn to adapt.

Step 5) Computation of current error and error function:

The difference between the current output and the desired output is calculated. The resulting value is the current error.

Step 6) Estimation of the back propagation error value:

There are numerous algorithms that can be applied to ANNs that allow the network to manipulate the error and learn. The back propagation algorithm was chosen because it is the most common and easiest to understand (Cabena, Hadjinian, Stadle, Verhees, & Zanasi, 1998). In essence, the information obtained is "back propagated" to the previous layer. The current error equals the back-propagated value. The back-propagated "error" value is determined by multiplying each incoming connection weight by the back-propagated value and adding this to the error in the source processing element. The hidden units estimate their error as a weighted sum of the error signals received from the connected output units. After this process, all the connections are updated in proportion to their error signal.

Step 7) Learning:

Using the back propagated error value, the weights of the connections are modified. When using competition, only the "winning" weights are modified. The network reruns the steps until the error achieved is acceptable. An error of 0, or no error, is never acceptable. This means that the network has memorized the data and the prediction process when running the network on new data is eliminated.

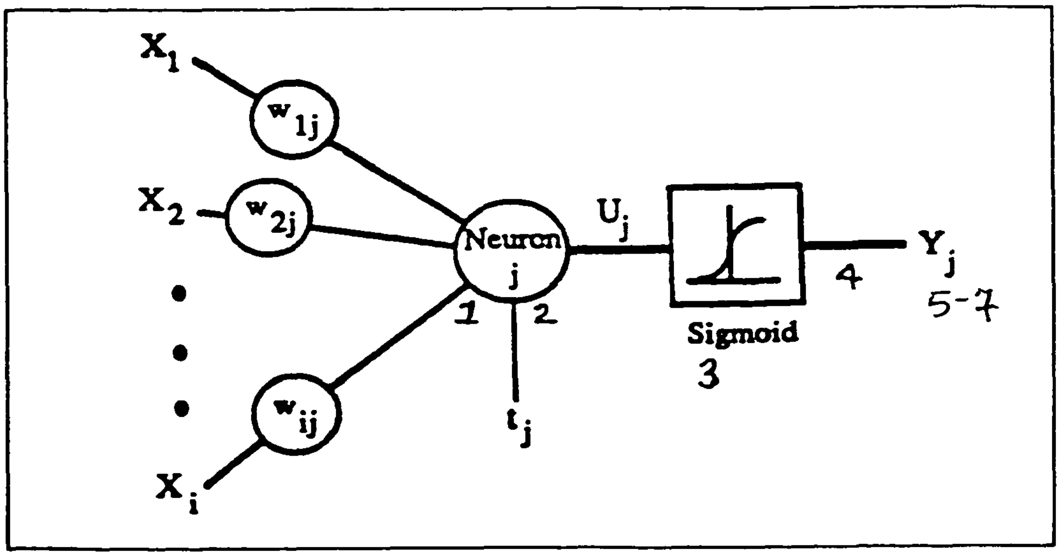

Figure 3. An artificial neuron procession element (VanEyden, 1996, Pg. 40)

Several ANN models are available in Statistica. The Multilayer Perceptron (MLP) model was chosen for analysis for several reasons. First, according to the Statistica documentation, a MLP is capable of modeling most real world problems, sufficiently and adequately (Statsoft, 1998). Second, in performing multiple preliminary pilot tests on the data sets, the MLP consistently outperformed other network types such as the Radial Basis Function and linear models for predictability. A decrease in prediction error and a higher percent in the variability of GGPA's could be accounted for. Third, in an interview with Rajiv Menta, CEO of Synergentics, Inc and neural networking expert (Personal communication, 2000), MLP networks are a good choice for this type of research problem. A back propagation algorithm was chosen because it is the easiest to understand which is advantageous to this entry level ANN research.

Data analysis using SPSS

Further analysis on the identified predictor variables was performed using Kruskal-Wallis, Spearman rho, Mann-Whitney, correlation, and multiple regression. To answer research question one, Mann-Whitney U, Kruskal-Wallis with Mann-Whitney U post hoc, and Spearman rho were used. The use of non-parametric statistics was necessary to due to the violation of normal distribution by the dependent variable GGPA. Mann-Whitney U was used on the nominal two- level variables, Kruskal-Wallis was used for the multi-level nominal variables, and Spearman rho was used on interval variables. To answer research question two, correlation and simultaneous multiple regression was used.

Model comparison

The same variables used by the ANN were related to the dependent variable of graduate GPA to validate their predictive capabilities and to further determine the accuracy of the ANN. A similar process was documented in Zhang et al. (1997) and Gorr et al. (1994) in which neural network predictive capabilities were compared to regression results to determine which model best predicted a particular outcome variable such a grades in graduate school. As discussed in Chapter Two, these results were mixed. Zhang et al. (1997) found that ANNs were superior to multiple regression techniques. Gorr et al. (1994) determined that ANNs have no better predictive power than other traditional statistical techniques. This study was performed to shed some additional light on this controversy. The following Chapter Four reports and interprets the results of the ANNs, descriptive analysis, correlations, as well as simultaneous multiple regression, and non-parametric statistics including Mann Whitney U, Kruskal-Wallis, and Spearman rho.

CHAPTER 4

RESULTS

Data analysis included descriptive statistics to better define the sample. Non-parametric statistics, (Kruskal-Wallis, Mann-Whitney U, and Spearman rho) were used to answer research question one because the dependent variable of graduate grade point average (GGPA) was not normally distributed. Simultaneous multiple regression and artificial neural networks were performed to assess the best combination of predictors of GGPA. Comparisons were discussed. All analyses were performed on both the master's students group and the Ph.D. student group.

Descriptive statistics

SPSS was used to gather descriptive statistics about the sample. Results of the nominal variables are reported in Table 4 including the n and percentage. Table 5 reports the n, mean, standard deviation, and skewness of the interval level variables.

Degree and college. Master's degree seeking students composed 79% of the sample, and 21% were Ph.D. seeking. The College of Applied Human Sciences held the highest number of students with 25%, followed by Engineering with 18%, Liberal Arts with 15%, and Natural Sciences with 12%.

Demographics. The gender split was fairly even with 56% male and 44% female. Marital status was also fairly even with 47% married and 51% single. The data did not contain numbers for divorced students.

Table 4

Descriptive Statistics of Nominal Variables for Graduate Students

Variable		N	% of Total	% Masters	% Ph.D.
Degree					
	Masters	3097	79		
	Ph.D.	805	21		
Gender					
	Female	1735	45	48	32
	Male	2167	55	52	68
Marital Status					
	Single	1688	55	55	39
	Married	1347	45	44	60
Citizenship Continent					
	North America	3368	86	89	74
	South American	53	1	1	3
	Europe	74	2	1	5
	Asia	293	8	6	12
	Africa	23	1	0.5	1
	Mideast	77	2	1	4
Residency Region					
	Colorado	1978	51	54	38
	West	434	11	11	10
	Midwest	382	10	9	12
	Northeast	243	6	6	7
	South	209	5	5	5
	International	656	17	14	27
College					
	Applied Human Sciences	978	25	28	15
	Natural Sciences	481	12	7	32
	Natural Resources	406	10	11	10
	Vet Med	247	6	5	10
	Business	237	6	8	0
	Engineering	704	18	17	20
	Liberal Arts	617	16	19	5
	Agriculture	232	6	5	8

Table 5

Descriptive Statistics of Interval Variables for Graduate Students

	Total				Masters				Ph.D.			
Variable	N	M	SD	Skew-Ness	n	M	SD	Skew-ness	n	M	SD	Skew-ness
Independent Variables												
Age@Entry	3901	30.92	7.82	0.87	3096	30.8	7.70	0.91	805	31.6	8.2	0.73
GREVerbal	3364	538	112.4	-0.12	2650	511	110.4	-0.11	714	505	119.5	-0.14
GREQuantitative	3364	538	125.3	-0.32	2650	579	125.3	-0.25	714	629	117.3	-0.61
GREAnalytical	3340	562	119.4	-0.3	2635	563	118.0	-0.31	705	575	124.3	-0.29
Dependent Variables												
GGPA	3881	3.71	0.32	-3.34	3078	3.69	0.33	-3.13	803	3.76	0.26	-4.61

Citizenship and residency. The majority (51%) of the students were Colorado residents at the time they applied to CSU graduate school. The database was not updated if a student's residency status changed while attending graduate school. Students who did not report a residency state were coded as International, comprising 17% of the sample. The Western region of the United States consisted of 11%, and 10% were from the Midwest. The majority of the students (86%) were residents of the North American continent that is composed of the United States and Canada. In addition, 7% of the students were residents of Asian countries and the other 4 regions had 1% or 2% each. The mean age of the students in the database was 31. Means for graduate GPA was 3.71, but this measure was highly skewed with more than half having a 3.80 or above. The GRE scores were above national averages, with a 538 for verbal and quantitative, and 562 for analytical.

Research question #1

Is there a difference in graduate GPA for master's and for Ph.D. students at Colorado State University between different levels or groups of each of nine admissions variables (citizenship continent, residency region, age of student when entering graduate school, gender, marital status, college, GRE verbal score, GRE analytical score, and GRE quantitative score).

Citizenship continent. A Kruskal-Wallis test was performed in SPSS. Kruskal-Wallis was the chosen statistic because the dependent variable, GGPA was not normally distributed. Small but statistically significant differences were found ($p<.001$) between graduate GPA's based on a students' citizenship continent, at both the master's and Ph.D. level. Table 6 shows the n and mean GGPA for each continent for both the master's and Ph.D. separately, and the table provides the results of Kruskal-Wallis as a chi-square.

Table 6

Kruskal-Wallis Analysis of Citizenship Continent and Graduate Grade Point Average

Citizenship Continent	Masters		Ph.D.	
	n	Mean GGPA	n	Mean GGPA
1. Europe	36	3.74	38	3.74
2. North America	2752	3.70	597	3.78
3. Asia	193	3.56	99	3.71
4. South America	31	3.66	22	3.70
5. Mideast	46	3.58	31	3.67
6. Africa	15	3.47	8	3.56
Test statistic	χ^2 (5)=87.71, $p<.001$		χ^2 (5)=30.88, $p<.001$	

At the master's level, Europeans had the highest GGPA, with North Americans second. These were reversed at the Ph.D. level. Gliner and Morgan (2000) suggest that a Mann-Whitney U test is an appropriate, but liberal, post hoc test for the Kruskal-Wallis to determine the significance of differences between pairs of groups (continents in this case). Table 7 and 8 reports the differences between the citizenship continents for master's and Ph.D. students.

Table 7

Mean Differences in Graduate Grade Point Average by Continent for Masters Students

Continents	1	2	3	4	5	6	Post hoc
1. Europe		.035	.174**	.078	.160**	.270**	1>3,5,6
2. North America			.144**	.043	.125**	.235**	2>3,5,6
3. Asia				-.102	-.020	.091	
4. South America					.082	.192*	4>6
5. Mideast						.111	
6. Africa							

** p <.01 comparing each pair of means using the Mann-Whitney test.
* p <.05 comparing each pair of means using the Mann-Whitney test

Table 8

Mean Differences in Graduate Grade Point Average by Continent for Ph.D. Students

Continents	1	2	3	4	5	6	Post hoc
1. Europe		-.036	.032	.045	.069	.185**	1>6
2. North America			.068**	.081*	.105**	.221**	2>3,4,5,6
3. Asia				.013	.037	.153	
4. South America					.024	.140	
5. Mideast						.116	
6. Africa							

** p <.01 comparing each pair of means using the Mann-Whitney test.
* p <.05 comparing each pair of means using the Mann-Whitney test

In the master's group, Europeans had significantly higher GGPA's than Asians (p<. 001), Mid Easterners (p=.007), and Africans (p=.005). North Americans had significantly higher GGPA's than Asians ($p<.001$), Mid Easterners ($p<.001$), and Africans ($p=.002$). Also South Americans were significantly higher on GGPA ($p=.033$) than Africans. This indicates that among master's students, European and North American citizens have higher GGPA's than Asians, Mideasterners, and Africans. Effect sizes ranged from medium ($d=.4$) to large ($d=.8$). This indicates a moderate to strong relationship between citizenship continent and GGPA at the master's level.

Europeans did not perform as well at the Ph.D. level. Only one significant difference was found. Europeans had higher GGPA's than Africans ($p=.036$). North Americans, however, had significantly higher GGPA's than Asian ($p<.001$), South American ($p=.033$), Mideastern ($p=.002$), and African students ($p=.006$). Effect sizes ranged from small/medium ($d=.3$) to large ($d=.8$) indicating a slightly moderate to large relationship between citizenship continent and GGPA at the Ph.D. level.

Residency region. A significant difference was found between the GGPA's of the six residency regions using a Kruskal-Wallis test for both master's and Ph.D. students at $p<.001$ (see Table 9).

Table 9

Kruskal-Wallis Analysis of Residency Region and Graduate Grade Point Average

Residency Region	Masters		Ph.D.	
	n	Mean GGPA	n	Mean GGPA
1. West	351	3.71	83	3.79
2. Colorado	1654	3.70	307	3.78
3. Midwest	281	3.73	99	3.76
4. Northeast	187	3.73	56	3.76
5. South	165	3.72	44	3.79
6. International	440	3.61	214	3.72
Test statistic	$\chi^2(5)=59.27$, $p<.001$		$\chi^2(5)=21.78$, $p=.001$	

The Mann Whitney U test found no significant differences among the North American regions, but did find significant differences between the international students and several regions. Table 10 and 11 reports these group differences.

Table 10

Mean Differences in Graduate Grade Point Average by Region for Masters Students

Regions	1	2	3	4	5	6	Post hoc
1. West		.010	-.023	-.024	-.015	.093**	1>6
2. Colorado			-.033	-.034	-.025	.083**	2>6
3. Midwest				-.012	.083	.116**	3>6
4. Northeast					-.095	.117**	4>6
5. South						.108**	5>6
6. International							

** $p < .01$ comparing each pair of means using the Mann-Whitney test.

Table 11

Mean Differences in Graduate Grade Point Average by Region for PhD Students

Regions	1	2	3	4	5	6	Post hoc
1. West		.050	.037	.032	.045	.077**	1>6
2. Colorado			.022	.017	-.010	.062**	2>6
3. Midwest				-.045	-.032	.040	
4. Northeast					-.028	.045	
5. South						.073*	5>6
6. International							

** p <.01 comparing each pair of means using the Mann-Whitney test.
* p <.05 comparing each pair of means using the Mann-Whitney test

At the master's level, differences were found between international students and students from the western United States (p<.001), from Colorado (p<.001), the Midwest (p<.001), the Northeast (p<.001), and the South (p<.001). Students residing in the United States have higher GGPA's than International students. The effect sizes were small (d=.2) to small/medium (d=.3) indicating that the relationship between residency region and GGPA was weak to slightly moderate for master's students. Among Ph.D. students, significant differences were found between International students and students from the Western United States (p=.004), Colorado (p<.001), and the South (p=.042). Effect sizes did not fluctuate and were small (d=.2) indicating a weak relationship between residency region and GGPA for Ph.D. students.

These findings are consistent with the differences found between citizenship continents. Students from the United States have higher GGPA's than international students, however, there is no significant difference among students within the United States.

Gender. Due to the violation of the assumption of normal distribution, Mann-Whitney U tests were performed to determine the relationship between gender and GGPA and marital status and GGPA. Table 12 reports that the relationship between gender and GGPA is statistically significant for both master's students and Ph.D. students. At the master's level, females have higher GGPA's than males ($p < .001$). The effect size is small to medium ($d = .33$) indicating a slightly moderate relationship between gender and GGPA for master's students. Female Ph.D. students also have higher GGPA's than males ($p = .011$) however there is a smaller difference between the mean GGPA's at the Ph.D. level. It is important to realize that statistical significance and practical significance are not the same thing (Morgan, Griego & Gloeckner, 2000). With large sample sizes, it is fairly easy to find significant statistical differences when the actual difference is quite small. This is the case with Ph.D. GGPA. Although there is a statistically significant difference in GGPA between genders, the difference is only .03 of a grade point and a very small effect size ($d = .1$) indicating a weak relationship between gender and GGPA for Ph.D. students.

Table 12

Mann-Whitney Analysis of Gender and Graduate Grade Point Average

Gender	Masters		Ph.D.	
	n	Mean GGPA	n	Mean GGPA
1.Female	1465	3.75	259	3.78
2. Male	1613	3.64	544	3.75
Test statistic	Z=-11.09, p<.001		Z=-2.56, p=.011	

Marital status. Marital status is significant at the master's level but, as Table 13 reports, is not significant at the Ph.D. level. Married master's students have higher GGPA's than single master's students ($p<.001$). The effect size is small ($d=.2$) indicating a weak relationship between marital status and GGPA for masters students. There is no statistical significant difference between married and single Ph.D. students.

Table 13

Mann-Whitney Analysis of Marital Status and Graduate Grade Point Average

Marital Status	Masters		Ph.D.	
	n	Mean GGPA	n	Mean GGPA
1. Married	1341	3.73	479	3.77
2. Single	1675	3.67	316	3.74
Test statistic	Z=-4.99, p<.001		Z=-1.43. p=.153	

College. A Kruskal-Wallis test determined that there was a significant difference between the average GGPA's and the eight colleges attended by students. The chi-square test statistic showed significant differences at the master's level ($p<.001$) and Ph.D. level ($p<.001$) as reported in Table 14.

Table 14

Kruskal-Wallis Analysis of College and Graduate Grade Point Average

College	Masters		Ph.D.	
	n	Mean GGPA	n	Mean GGPA
1. Natural Resources	323	3.66	80	3.71
2. Vet Med.	168	3.72	79	3.81
3. Applied Human Sciences	855	3.80	123	3.84
4. Engineering	540	3.61	163	3.74
5. Liberal Arts	569	3.69	38	3.80
6. Natural Sciences	225	3.72	255	3.73
7. Agriculture	165	3.60	65	3.75
8. Business	233	3.55	NA	NA
Test statistic	$\chi^2(7)=267.03$, $p<.001$		$\chi^2(6)=35.85$, $p<.001$	

At the master's level, the College of Applied Human Sciences, Liberal Arts and Veterinary Medicine have the highest GGPA while Agriculture, Engineering, and Business have the lower GGPA. Applied Human Sciences, Veterinary Medicine, and Liberal Arts also had the highest GGPA's at the Ph.D. level, however Natural Sciences, Engineering, and Natural Resources have the lower GGPA's.

Mann-Whitney U tests were performed to analyze the individual paired group differences between colleges on GGPA (see Table 15 and 16). At the master's level, the College of Applied Human Sciences students GGPA was significantly higher than all other seven colleges: Natural Resources ($p<.001$), Veterinary Medicine ($p=.001$), Engineering ($p<.001$), Liberal Arts ($p<.001$), Natural Sciences (p<. 001), Agriculture (p<. 001), and Business (p<. 001). Liberal Arts students had significantly higher GGPA's than Natural Resources (p<. 001), Engineering (p<. 001), Business (p<. 001), and Agriculture (p<. 001). Veterinary Medicine GGPA's were significantly higher than Business (p<. 001), Engineering (p<. 001), Agriculture (p<. 001), and Natural Resources (p=. 007). In addition, Natural Sciences students GGPA's were significantly higher than Agriculture (p<.001), Engineering (p<.001), Natural Resources (p=.009), and Business (p<.001). Students in the College of Business had significantly lower GGPA's than students in Engineering (p=. 021) and Natural Resources (p<. 001), and Engineering students GGPA's were significantly lower than students in Natural Resources (p=.028). Effect sizes range from small (d=.1) to large (d=.8) indicating varying strengths of relationship between college and GGPA.

Table 15

Mean Differences in Graduate Grade Point Average by College for Masters Students

Colleges	1	2	3	4	5	6	7	8	Post hoc
1. Natural Resources		-.0568	.114**	-.137**	.049*	-.034**	-.055**	.057**	1<2,4,6,7 1>3,5
2. Vet Med.			.171**	-.081**	.106**	.023	.019	.113**	2>3,5,8 2<4
3. Business				-.251**	-.064*	-.148**	-.169**	-.058	3<4,5,6,7
4. Applied Human Sciences					.187**	.103**	.082**	.193**	4>5,6,7,8
5. Engineering						-.084**	-.104**	.069**	5<6,7,8
6. Liberal Arts							-.021	.09	
7. Natural Sciences								.111**	7>8
8. Agriculture									

** p <.01 comparing each pair of means using the Mann-Whitney test.
* p <.05 comparing each pair of means using the Mann-Whitney test

At the Ph.D. level, the College of Applied Human Sciences had significantly higher GGPA's than Natural Resources (p<.001), Engineering

(p<.001), Natural Sciences (p<.001), and Agriculture (p<.001). Veterinary medicine had significantly higher GGPA's than Engineering (p=.003), and Natural Sciences (p=.003). Natural Resources had a significantly lower GGPA than Veterinary medicine (p=.004). Effect sizes are small (d=.1) to small/medium (d=.4) indicating a weak to moderate relationship between college and GGPA.

Table 16

Mean Differences in Graduate Grade Point Average by College for Ph.D. Students

Colleges	1	2	3	4	5	6	7	Post hoc
1. Natural Resources		-.106**	-.033**	-.033	-.088	-.025	-.038	1<2,3
2. Vet Med.			-.027	.073**	.018	.080**	.068	2>4,6
3. Applied Human Sciences				.100**	.045	.108**	.095**	3>4,6,7
4. Engineering					-.055	.078	-.053	
5. Liberal Arts						.063	.049	
6. Natural Sciences							-.013	
7. Agriculture								

** p <.01 comparing each pair of means using the Mann-Whitney test.

Age of student when entering graduate school. A Spearman Rho correlation was used to determine the relationship between student's age and GGPA (see Table 17). Spearman rho was the chosen statistic because of the skewness of the GGPA variable. The assumption of normally distributed GGPA's was markedly violated. In the Ph.D. group, age was significant at the p=.006 level (rho= -.097). The rho^2 was .009 (.9%). This means that .9% of the variation in GGPA can be predicted by the students age at the time they enter graduate school. According to Cohen (1988) a rho=-.097 indicates a small effect size or that the relationship between GGPA and age at the time of entry to graduate school is weak. The correlation is negative indicating that older students entering the Ph.D. program receive lower GGPA's. Age was not significant at the master's level.

Table 17

Spearman Rho Analyses of Interval Level Variables and Graduate Grade Point Average

	Masters		Ph.D.	
	rho	p	Rho	p
Age at Entry	0.000	0.986	-0.097	0.006
GRE Verbal	0.195	<.001	0.263	<.001
GRE Quantitative	-0.01	0.620	0.033	0.374
GRE Analytical	0.164	<.001	0.207	<.001

GRE Verbal scores. A Spearman Rho indicated that there was a statistically significant relationship between GRE verbal scores and GGPA for both master's and Ph.D. students. The results can be seen in Table 17. At the master's level, GRE verbal scores were significant at p<. 001 (rho=. 195). The rho^2= .04 indicating that 4% of the variance in master's students GGPA can be predicted by GRE verbal scores. The effect size is small to medium, meaning that GRE verbal scores have a modest relationship to master's students GGPA's. The correlation is positive denoting that the higher the GRE verbal scores the higher the GGPA for master's students.

GRE verbal scores for Ph.D. students had similar results. GRE verbal scores were significant at the p<. 001 level (rho=. 263). The rho^2=. 069 which denotes that 7% of the variance in Ph.D. GGPA's can be predicted by GRE verbal scores. The effect size is medium, and the correlation is positive meaning that the higher the GRE verbal scores the higher the GGPA for Ph.D. students.

GRE Quantitative scores. As Table 17 reports, GRE quantitative scores are not statistically significant at either the master's or the Ph.D. level. This indicates that for all colleges combined, GRE quantitative scores have no predictive value in determining GGPA.

GRE Analytical scores. GRE analytical scores were significant at both the master's and Ph.D. levels (see Table 17). At the master's level, GRE analytical scores were significant at $p<.001$ (rho=.164). The rho²=. 018 indicating that only .2% of the variance in GGPA for master's students can be predicted by GRE analytical scores. The effect size is small to medium and the correlation is positive, indicating that the higher the GRE analytical score, the higher the GGPA for master's student; however this relationship is weak.

At the Ph.D. level, GRE analytical scores were significant at $p<. 001$ (rho=. 202). The rho²=. 040 increasing the percentage of variance of GGPA that can be predicted by GRE analytical scores to 4% at the Ph.D. level. The effect size is also small to medium and the correlation is positive.

Research question #2

Is there a combination of admissions variables that predicts GGPA better than any one variable alone and if so, what is that combination?

Multiple Regression results

Simultaneous multiple regression was performed on the combination of variables outlined in Chapter Three. Table 18 and 19 reports the correlations for master's and Ph.D. students and Table 20 and 21 report the results of the multiple regression for master's and Ph.D. students. It is important to note the direction of the correlation for nominal variables in these four tables. A negative correlation indicates that the level within the nominal variable with the lowest code number has higher GGPA. The nominal variables were coded as follows:

Gender: female = 1 male = 2

Marital status: married = 1 single = 2

Citizenship continent: North America = 1 other continents = 0

Residency region: Colorado = 1 other regions = 0

College: The College of Applied Human Sciences = 1 other colleges = 0

Correlation results. As reported earlier Table 18 indicates that at the master's level, there is a significant positive relationship between GGPA and North American continent, college of Applied Human Sciences, GRE verbal, and GRE analytical at the $p < .01$ level. There are significant negative relationships with gender and marital status, indicating that females and married students have higher GGPAs. These variables may be significant predictors in the regression model. It is also interesting to note that continent, region, and college are all significantly related at the $p < .01$ level. Thus, these variables are reporting somewhat similar or overlapping information. In the regression model it is possible that as one of these variables loads, it may cause another not to due to the fact that they could be reporting similar information, called multicollinearity.

At the Ph.D. level, the relationship of the variables to GGPA is almost identical, with the exception of age and gender. If you will recall, age was not significantly related to GGPA at the master's level, however, it is statistically significant at the Ph.D. level. In addition, gender is significant at the r=.063 level for Ph.D. students and r=.165 at the master's level.

Table 18

Correlation's for Admissions Variables for Masters Students

Variable	Graduate GPA	N. America or Other Continent	Colorado or Other Region	Age At Entry	Gender	Marital Status	College of AHS or Other College	GRE Verbal	GRE Quantitative	GRE Analytical
Graduate GPA	1.000	.124**	.026	-.010	-.165**	-.089**	.196**	.138**	.006	.140**
North America or Other Continent		1.000	.324**	-.047**	-.125**	.042	.106**	.420**	-.166**	.241**
Colorado or Other Region			1.000	.056**	-.114**	-.125**	.165**	.086**	-.190**	-.042*
Age at Entry				1.000	-.005	-.044*	-.005	-.006	-.007	-.053**
Gender (males higher)					1.000	-.082**	-.287**	.010	.332**	.032
Marital Status (singles higher)						1.000	-.043*	.024	.030	.072**
College of AHS or Other College							1.000	.194**	-.348**	-.205**
GRE Verbal Score								1.000	.330**	.578**
GRE Quantitative Score									1.000	.596**
GRE Analytical Score										1.000

*$p < .05$. **$p < .01$.

Table 19

Correlation's for Admissions Variables for Ph.D. Students

Variable	Graduate GPA	N. America or Other Continent	Colorado or Other Region	Age At Entry	Gender	Marital Status	College of AHS or Other College	GRE Verbal	GRE Quantitative	GRE Analytical
Graduate GPA	1.000	.125**	.076	-.091**	-.063*	-.089**	.177**	.218**	.043	.193**
North America or Other Continent		1.000	.370**	-.083*	-.119**	.136**	.187**	.580**	-.203**	.322**
Colorado or Other Region			1.000	.028	-.180**	-.095**	.446**	.224**	-.313**	-.010
Age at Entry				1.000	-.054	-.050	-.011	-.045	.023	-.037
Gender (males higher)					1.000	-.131**	-.150**	-.115**	.184**	-.091**
Marital Status (singles higher)						1.000	-.122**	.116**	.056	.131**
College of AHS or Other College							1.000	.021	-.460**	-.220**
GRE Verbal Score								1.000	.176**	.646**
GRE Quantitative Score									1.000	.541**
GRE Analytical Score										1.000

*$p < .05$. ** $p < .01$.

Simultaneous Regression results. According to the Table 20, the F (33.58) is statistically significant, indicating that one or more of the independent variables is a significant predictor of GGPA at the master's level. The combination of region, gender, marital status, college, GRE verbal scores, and GRE analytical scores can predict GGPA. College is the best predictor (β=.208, p<.001). Gender and marital status are both significant predictors, however, female gender (β=.-127, p<.001) is a slightly better predictor than married marital status (β=.-110, p<.001). Region other than Colorado is also a significant predictor (β=.-049, p=.015), as are GRE verbal (β=.096, p<.001) and GRE analytical (β=.113, p<.001) scores. Although the regression model indicates that there is predictability in this combination of variables the percentage of variance that can be predicted by these variables is relatively small (Adjusted R^2=.102). This combination of variables can predict 10% of the variance in GGPA. The effect size is medium (R=.324) meaning there is moderate relationship between the combination of college attended, gender, marital status, residency region, the GRE verbal, and the GRE quantitative scores with GGPA for students at the master's level.

Table 20

Multiple Regression Results for Masters Students

Variable	B	Std. Error	β	p
North America or Other Continent	.0433	.024	.042	.075
Colorado or Other Region	-.0303	.012	-.049	.015
Age at Entry	.0002	.001	-.005	.775
Gender	-.0783	.013	-.127	<.001
Marital Status	-.0683	.012	-.110	<.001
Applied Human Sciences or Other College	.1400	.014	.208	<.001
GRE Verbal Score	.0003	<.001	.096	<.001
GRE Quantitative Score	.0001	<.001	.022	.429
GRE Analytical Score	.0003	<.001	.113	<.001
Constant	3.85	<.001		

Note. R=.324, R^2=.102, F=33.58, p<.001

Table 21 reports a significant F of 7.713. This means that one or more of the independent variables are significant predictors of GGPA. The combination of age, marital status, college, GRE verbal, and GRE analytical scores predict GGPA. As with the master's students, college is the best predictor of GGPA (β=.143, p=.001). GRE verbal (β=.142, p=.012) and GRE analytical (β=.136, p=.024) were the next strongest predictors. Single marital status (β=.-117, p=.002), and young age (β=.-087, p=.019) also added predictive value.

Although this combination of variables had predictive power, the R^2=.080, indicating that 8% of the variance in GGPA can be accounted for by the combination of these variables. The effect size (R=.304) is medium meaning that there is a moderate relationship between the combination of college a student attends, their GRE verbal and analytical scores, their marital status and age with GGPA at the Ph.D. level.

Unlike the master's students, region and gender were not considered to have predictive value at the Ph.D. level. Age was a significant predictor at the Ph.D. level but not at the master's. This is consistent with the results from research question one, in which age was not considered significantly related to master's level GGPA.

Table 21

Multiple Regression Results for Ph.D. Students

Variable	B	Std. Error	β	p
North America or Other Continent	-.0071	.032	-.011	.083
Colorado or Other Region	-.0142	.024	-.026	.561
Age at Entry	-.0030	.001	-.087	.019
Gender	-.0233	.022	-.041	.289
Marital Status	-.0633	.020	-.117	.002
Applied Human Sciences or Other College	.1070	.033	.143	.001
GRE Verbal Score	.0003	<.001	.142	.012
GRE Quantitative Score	<.0001	<.001	.016	.765
GRE Analytical Score	.0003	<.001	.136	.024
Constant	3.80	.104		

Note. R=.304, R^2=.080, F=7.713, p<.001

In summary, the multiple regression model for master's students indicates that students from the College of Applied Human Sciences, who are not residents of the state of Colorado, are married females, and have higher GRE verbal and analytical score have higher GGPA's. The regression model for Ph.D. students indicates that students from the College of Applied Human Sciences who are younger and married with higher GRE verbal and analytical scores have higher GGPA's.

Artificial neural network results

All available independent (input) variables were used to run five multi layer perceptron (MLP) neural networks at the master's and Ph.D. levels to determine predictive models for GGPA. Tables 22 and 23 report these results. Indicators of network performance were reported as a multiple correlation (R) and by mathematically squaring $\underline{R}$, an $\underline{R}^2$ was obtained. Correlation is defined by Statsoft as a multiple $\underline{R}$. These numbers are most closely related to the $\underline{R}$ and $\underline{R}^2$ statistic reported in simultaneous regression. An adjusted $\underline{R}^2$ was available in the regression output; however, no adjusted $\underline{R}^2$ was provided by the ANN output.

ANN results at the Master's level. As seen in Table 22, the second ANN test produced the best multiple R results of the five test runs. The significant predictor variables reported by this network included college, gender, marital status, GRE analytical, and GRE verbal scores. College was the first to load and determined to be the best predictor of GGPA, or it can be said that the performance of the network would deteriorate the most if college were not included as a predictor variable. The next variable to load was marital status and then gender. GRE analytical loaded before GRE verbal, followed by region, age, and citizenship continent. The multiple correlation was $\underline{R}=.334$ and the $\underline{R}^2$ was .11. This indicates that 11% of the variance in GGPA at the master's level can be predicted by a combination of college, gender, marital status, GRE analytical, and GRE verbal scores.

Table 22

Artificial Neural Network Results on Masters Students

Input Variables		
GRE Verbal Score	College	Residency Region
GRE Quantitative Score	Gender	Citizenship Continent
GRE Analytical Score	Marital Status	Age at Entry

	Test 1	Test 2	Test 3	Test 4	Test 5
R	.296	.334	.249	.310	.327
R^2	.088	.112	.062	.096	.107

ANN results at the Ph.D. level. The best network for predicting GGPA at the Ph.D. level outperformed the best network at the master's level. The fifth test run had superior performance as can be seen by Table 23. The predictor variables indicated by this network included GRE analytical scores, gender, marital status, GRE quantitative scores, GRE verbal scores, age, residency state, and citizenship country. GRE analytical score was the first variable to load into the network, followed by gender and marital status. GRE quantitative and verbal scores were next, in that order, followed by age, residency region, and citizenship country. The multiple correlation was R=.486 with an R^2 of .236. This means that 24% of the variance in GGPA for Ph.D. students can be predicted by a combination of GRE analytical scores, gender, marital status, GRE quantitative scores, GRE verbal scores, age, residency state, and citizenship country.

Table 23

Artificial Neural Network Results on Ph.D. Students

Input Variables		
GRE Verbal Score	College	Residency Region
GRE Quantitative Score	Gender	Citizenship Continent
GRE Analytical Score	Marital Status	Age at Entry

	Test 1	Test 2	Test 3	Test 4	Test 5
R	.282	.232	.390	.414	.486
R^2	.080	.054	.152	.171	.236

Research question #3

Do artificial neural networks perform better than traditional statistical techniques in predicting success in graduate school as defined by graduate grade point average?

Comparing Master's models

As Table 24 reports, the ANN and simultaneous regression models performed remarkably similarly at the master's level. They both loaded the same variable first, college. The next three variables to load for each model were the same, but they loaded in different order. The ANN model loaded marital status, gender, and GRE analytical scores in that order. The regression model loaded gender, GRE analytical scores, and marital status, in that order. Both models loaded GRE verbal scores and region as the fifth and sixth variables to have predictive value. The ANN model added age and citizenship continent as seventh and eighth predictive variables.

Table 24

Comparing Artificial Neural Networking to Simultaneous Multiple Regression

	Student Level and Type of Analysis			
	Best Master's ANN	Master's Regression	Best Ph.D. ANN	Ph.D. Regression
Predictor Variables In order	College Marital Status Gender GRE Analytical GRE Verbal Region Age Continent	College Gender GRE Analytical Marital Status GRE Verbal Region	GRE Analytical Gender Marital Status GRE Quantitative GRE Verbal Age Region Continent	College GRE Verbal GRE Analytical Marital Status Age
R	.334	.324	.486	.304
R^2	.122	.105	.236	.092

In comparing the R and R^2 statistics, there are also similarities. The adjusted R^2 was not used for comparison purposes because there was no adjusted R^2 statistic available in the ANN output. There was only .01 difference between the multiple Rs of the models with the best ANN model being slightly higher (R=. 334). Similarly, the R^2 statistics only yielded a difference of .017 with the best ANN model being slightly higher at R^2=. 112. There is a 1.7% difference in the predictive ability of these models and the predictive variables are very similar. It is also important to note that of the five ANN models run, two of the models indicated better predictability of GGPA, while three of the models indicated lower predictability.

Comparing Ph.D. models

In contrast to the master's models, the two Ph.D. models performed very differently (Table 24). The regression model loaded College, GRE verbal, and GRE analytical scores as the first three variables in that order. In contrast, the ANN model did not indicate college to have predictive value. It loaded GRE analytical scores first, followed by gender, and marital status. The regression model did not find gender to have predictive value. Regression did however load marital status fourth. Regression loaded age as the fifth and final variable to have predictive value. The ANN model also included age, but placed GRE quantitative and GRE verbal scores before age. The ANN model also included region and citizenship continent as having predictive value, loading them seventh and eighth.

There was also a greater difference in the multiple correlation ($\underline{R}$) and $\underline{R}^2$ statistics between the two Ph.D. models. The best ANN indicated an $\underline{R}$ of .486 where as the regression model reported an $\underline{R}$ of .304. The $\underline{R}^2$ for the best ANN model was .236 and the regression model had a $\underline{R}^2$ of .092. There was a 15% difference in the percentage of variance in GGPA predicted between the two models. Of the five ANN models run, three indicated better predictability of GGPA than the regression model, however two of the models had lower predictability.

It is clear that the comparison of the two models is mixed. Regression outperformed three of the five ANNs at the master's level, but only slightly. In contrast three of the ANNs outperformed regression at the Ph.D. level, the best one by a large margin, showing a higher multiple correlation and a 15% greater percentage of variance than could be predicted by the regression. A more detailed discussion of this comparison follows in Chapter Five.

CHAPTER FIVE

DISCUSSION

This chapter discusses each significant independent variable as it relates to the dependent variable of graduate grade point average (GGPA) at both the master's and the Ph.D. levels, including a comparison of the simultaneous regression models to the ANN models at both the master's and the Ph.D. levels. A general discussion of the use of both models and limitations related to the database follow. The chapter ends with a discussion of recommendations for future research and practice and a conclusion.

Variable and model discussion

Citizenship continent. Citizenship continent was significantly related to GGPA at both the master's and Ph.D. levels ($p < .001$). North American and European students have higher GGPA's. This may be due to the fact that language and culture differences are greater for students coming from Asia, Africa, South America, and the Mideast to attend graduate school. It is also important to note that both the master's group and the Ph.D. group were heavily skewed towards North American students. Master's students were 90% North American and Ph.D. students were 87% North American.

Citizenship continent was not a significant predictor for either group of students in the regression models, probably due to the overlap with the residency region variable. However, continent did load as the last predictor variable in both of the ANN models, indicating that according to the ANNs, citizenship continent has predictive value on GGPA. North American citizens have higher GGPAs than other continents, but this relationship is weak.

Citizenship continent was not included as a predictor variable in the existing literature, however, Graham (1991) did find ethnicity to have statistical significance (R=. 426) in predicting GGPA for Master of Business Administration (MBA) students. White students received the higher GGPAs. Even though these variables may not be measured or categorized the same, ethnicity and citizenship continent could be reporting somewhat similar information if Graham had a lot of international Masters of Business Administration students.

Residency region. The region that a student reported for residency was significantly related to GGPA for both the master's students (p<. 001) and Ph.D. students (p=. 001), however, it is important to note that the Mann Whitney U post hoc determined that the only statistically significant differences were between international students and students within the North American regions. In re-coding the original data, all students who did not report a region of the United States for residency were considered International students. These findings are consistent with the relationships with citizenship continent. International students have lower GGPA's than students who are residents of the United States. No significant differences were found among students within the United Sates. Gorr et al. (1994) included residency as a predictor variable in determining GPA for North Dakota State Bachelors of Science students and found no statistical significance. It is possible that if international students were not included as a level in the residency region variable, similar results to the Gorr et al. study would have been found.

Residency region had predictive value in both the regression and ANN models for master's students, however, it was one of the last variables to load in each, indicating that the predictive value was weak. Residency region was a significant predictor in the Ph.D. ANN model, but was not considered to have predictive value in the Ph.D. regression model. It is interesting to note that both residency region and citizenship continent loaded in the ANN models for both groups, but only residency region loaded in the master's regression model.

College. College was significantly related to GGPA for both master's students ($p < .001$) and Ph.D. students ($p < .001$). Students within the College of Applied Human Sciences received higher GGPA's than any other college at Colorado State University. This does not necessarily mean that students within the College of Applied Human Sciences are smarter than other students on campus. Perhaps the professors within this college are more lenient graders, or perhaps there was more support for these students, giving them a greater chance to succeed. It should be noted that the difference in mean GGPA's between the College of Applied Human Sciences and the college with the lowest GGPA was only .25 of a grade point for master's students and .13 for the Ph.D. students, indicating that although college is a significant predictor of GGPA, the difference in these mean GGPA's in small. Once again, the difference between statistical significance and practical significance applies. The difference in GGPA between colleges was statistically significant; however these differences were of little practical significance.

College loaded as a significant predictor of GGPA for both the regression model and the ANN model at the master's level. It loaded as the first variable in both models. The regression model for Ph.D. students also determined college to be a significant predictor of GGPA; however, the ANN model did not determine college to have any predictive value. No other studies within the literature used college as a possible predictor of GGPA. Other studies focused on admissions within a particular college or program rather than university wide as this study did.

Gender. Gender was significantly related to GGPA for both master's students ($p < .001$) and Ph.D. students ($p = .011$). Females earned higher GGPA's than males in both groups. A possible reason for this relationship is that there were a higher percentage of international males (master's, 7%; Ph.D., 10%) than international females (master's, 3%; Ph.D., 2%), who as discussed previously received lower GGPA's than North American students. There are also a higher percentage of males in colleges other than Applied Human Sciences (master's, 44%; Ph.D., 60%) than females (master's, 29%; Ph.D., 24%). Recall that students in the College of Applied Human Sciences received the highest GGPA's of all the other colleges at Colorado State University.

Gender was also considered a significant predictor of GGPA at the master's level in both the regression model and the ANN model, loading second in the regression model and third in the ANN model. For the Ph.D. students, gender was a significant predictor in the ANN model, but not in the regression model.

Two studies were found in the literature that examined the relationship of gender to GGPA. Paolillo (1982) found gender to be significantly related to GGPA when the two variables were correlated (r=. 14, p=. 05), but was not a significant predictor in a stepwise multiple regression model. This finding is consistent with the results of this study with the Ph.D. student group where gender was significantly correlated to GGPA (r= -. 06), but was not a significant predictor in the regression model. Graham (1991) included gender as a possible predictor variable of GGPA for master's students in a stepwise regression model and found no statistical significance. The current study contradicts Graham and Paolillo's findings and determined that gender was a significant predictor of GGPA at the master's level.

Marital status. Martial status was significantly correlated to GGPA at the master's level (p<. 001) but was not significantly related to GGPA at the Ph.D. level using the Mann Whitney U. Married master's students receive higher GGPA's than single master's students. In an attempt to determine a reason for this relationship at the master's level, multiple cross tabs were run to determine if there was a greater number of married females, married students in the College of Applied Human Sciences, or married students that were North American citizens. These variables were chosen because of the significant relationship they had with GGPA. Percentages between married and single students within these areas differed by only a couple of percentages in each case. To further add to the perplexity of the relationship of marital status to GGPA, marital status loaded as a significant predictor of GGPA in both regression models (master's and Ph.D.) and both ANN models (master's and Ph.D.) indicating that when grouped with other predictor variables, married marital status had predictive value. Two studies in the literature (Paolillo, 1982; Graham, 1991) found no statistically significant predictability of marital status on GGPA for master's students.

Age. Age was determined to be significantly related to GGPA at the Ph.D. level (rho=-. 097, p=. 006), but was not significant at the master's level. Younger Ph.D. students receive higher GGPA's. However, it is important to note that this relationship is weak. Age can only account for .9% of the variability in GGPA for Ph.D. students (rho^2=. 009). A possible explanation for this relationship could be that more Ph.D. students are returning to the University after taking time for family, careers, and other life issues. Study skills could be rusty and the rigor of a Ph.D. program is high. In general, the earlier in life a student enters a Ph.D. program, the more likely they are to receive a higher GGPA.

Age was considered a significant predictor variable in both the master's ANN model and the Ph.D. ANN model. In addition, the Ph.D. regression model indicated that age has predictive value. All three models load age as one of the last variables indicating the predictive value is low. Fisher and Resnick (1990), Graham, (1991), and Paolillo (1982) included age as a potential predictor of GGPA for master's students and found no predictive value. This study is somewhat consistent with those results. The ANN model indicated that age had weak predictive value at the master's level, but it was the second to last variable to load. It is possible that the ANN model for master's students identified an non-linear relationship between age and GGPA that the linear regression model was not able to detect.

GRE Verbal score. GRE verbal scores were significantly correlated to GGPA for at both the master's level (rho=.135, p<. 001) and the Ph.D. level (rho=.225, p<. 001). The higher the GRE verbal scores the higher the GGPA at both levels. GRE verbal scores also loaded as significant predictors in both the regression and ANN models at both student levels indicating that GRE verbal scores are a definite, if relatively weak predictor of GGPA at Colorado State University. This finding is consistent with the existing literature. Brown and Weaver (1979) determined that the GRE verbal score was a clear predictor of graduate grades for 129 journalism students at Indiana University (no statistical results were provided). Morrow (1993) found a significant correlation (r=. 42, p=. 01) between GRE verbal scores and GGPA. Morrow found a much stronger relationship between GRE verbal scores and GGPA than was found by this study. Kirnan and Geisinger (1981) determined that there was a significant relationship (r=. 43, p<. 001) between GRE verbal scores and grades on a master's comprehensive examination. However, Nelson and Nelson (1995) performed a logistic regression model with master's degree completion as the outcome variable and determined that GRE verbal scores were not significant predictors of degree completion for probationary students. They did however find statistical significance between GRE verbal scores and degree completion for regularly admitted students ($p < .001$).

GRE Quantitative scores. GRE quantitative scores were not significantly correlated to GGPA for either the master's students or the Ph.D. students. Also, the regression and ANN models found no predictive value for master's students, neither did the regression model for Ph.D. students. However, GRE quantitative scores were determined to have predictive value by the Ph.D. ANN model. GRE quantitative score was the fourth variable to load. It is interesting that the correlation between the GRE quantitative score and GGPA for Ph.D. students was not significant, yet the ANN model determined predictive value. Once again, it is possible that the ANN found a non-linear relationship between GRE quantitative scores and GGPA that the linear regression model was not able to detect. Predictive value in the GRE quantitative score was also found in several other studies. King et al. (1993) reported that the GRE quantitative score had better predictive value of GGPA than either of the other GRE scores (verbal and analytical) for students in the Harvard Government Department. Morrow (1993) found GRE quantitative scores to be significantly correlated to GGPA ($r = .40$, $p = .01$) however, it was the weakest correlation of all three GRE scores. Kirnan and Geisinger (1981) found a correlation ($r = .27$, $p < .001$) between GRE quantitative scores and a master's comprehensive examination. Nelson and Nelson's (1995) logistic regression model found statistical significance between GRE quantitative scores and master's degree completion for probationary students ($p = .01$) but no significance for those student admitted regularly. Brown and Weaver (1979) determined that GRE quantitative scores had predictive value for GGPA only for journalism students entering with no previous professional experience. The predictive value of GRE quantitative scores is inconclusive and the predictive value is weak.

GRE Analytical scores. GRE analytical scores were significantly correlated to GGPA at both the master's level (rho=.136, $p<.001$) and the Ph.D. level (rho=.202. $p<.001$). The higher the GRE analytical score, the higher the GGPA. In addition, it was determined to be a significant predictor of GGPA by both regression and ANN models for both student levels. Although GRE analytical scores were only used as potential predictor variables by two of the studies utilizing GRE scores, both found statistical significance. Morrow (1993) found GRE analytical scores to have the highest correlation to GGPA of all three GRE scores ($r=.57$, $p=.01$). In addition Nelson and Nelson's (1995) study found GRE analytical scores to be statistically related to master's degree completion for probationary students ($p=.03$) but not related to degree completion for regularly admitted students.

Model comparison

Referring back to Table 24 which compares the simultaneous regression models for both master's and Ph.D. student groups and the ANN models for both master's and Ph.D. student groups, there are similarities and differences. Within the master's student group, the two models are remarkably similar. The squared multiple correlation's differ by only a percentage. The variables that loaded are also very similar. The ANN model added two variables to the list of those with predictive value, age and citizenship continent. It can be concluded that the regression model and the ANN model perform similarly for the master's student group, with the ANN perhaps having a higher sensitivity to variables with weaker predictive power than the regression model.

Within the Ph.D. group, however, there are markedly different results. The best ANN model appears to have outperformed the regression model when comparing the multiple correlations. In addition, the variables that were determined to have predictive value are dissimilar. Caution needs to be taken in interpreting these results. The N for the Ph.D. ANN model was only 520 cases. This is considered a very small sample size for ANNs and in performing the five test models, results were very different for each test. Correlation's ranged from R=.282 to R=.486. In addition, the variables determined to be significant predictors were not consistent across the five test networks. This study reports the results of the top-performing network based on correlation numbers, but once again, caution needs to be taking in interpreting the results. Additional tests need to be run with a larger sample in order to be confident of the accuracy of the output.

More confidence can be taken in the results of the master's group. Each of the five test models performed by the ANN were consistent in the correlation and which variables were determined to be predictors. It is determined that traditional statistical techniques such as simultaneous regression, and artificial neural networking perform basically the same for these types of research questions. Regression could account for 11% of the variance in GGPA and ANNs could account for 12%. These results are consistent with the only other study found in the field of educational research where ANN technology was compared to traditional statistical methods (Gorr et al. 1994). It was determined that ANNs performed no better than traditional methods in determining GGPA. Studies in other fields, however, did determine ANNs to outperform traditional statistical methods (Chance, MacLin & Lykins, 1993; Everson, 1994; Lykins & Chance, 1992b; MacLin, Chance & Lykins, 1993) In a review of literature by Zhang et al. (1997) which compared ANN technology to traditional statistical methods across multiple types of research projects, results were mixed. Of the 11 reviewed studies, three studies found ANNs to be inferior to regression; three studies found similar results to this study and concluded them to be no different; and five studies reported ANNs to be superior to traditional statistical methods. Perhaps this can be attributed to the high technological level of ANNs. Even though the researcher had help from experts in the field, including the engineers at Statsoft, the developers of the software, and did what was reasonable, it was very difficult to understand the results of the ANN software.

Limitations of the database

Creating and maintaining an accurate historical database is paramount for analysis of the admissions process. It was found during the course of this research that Colorado State University is not doing that very well. There were several databases available on campus for use in this study; however all were incomplete. The amount of missing data was a large detriment to the accuracy of the results in this study. Of the original list of 15 independent variables selected for the study, only eight were at least 98% complete. Ethnicity was 84% complete and was later eliminated from the list of independent variables due to issues of multicolliniarity and missing data. The GRE scores were 86% complete but were retained as independent variables based on the reviewed literature. Several past studies found GRE scores to be good predictors of graduate student success (Brown & Weaver, 1979; King et al. 1993; Kirnan & Geisinger, 1981; Morrow, 1993; Nelson & Nelson, 1995). Another important variable in predicting graduate success, according to the past literature, was undergraduate grade point average (UGPA). This variable was only 28% complete. With this amount of missing data, it was determined that this variable would not be useful in this study. In addition, the variable of undergraduate credits completed was eliminated based on the fact that it was only 38% complete. Several people told the researcher that undergraduate information was available, but on further investigation it was determined that only information for students who completed their undergraduate studies at Colorado State University (CSU) had undergraduate information in the database. Another discrepancy in the data was the percentage of students who graduated. The CSU database indicated that 85% of the entering graduate students actually graduated. It was originally thought that graduate or not graduate could be an interesting dependent variable, however, due to the difference between the national figure of approximately 50% and the database figure of 85%, it was decided that this would be eliminated as an accurate dependent variable.

Impact on the study

The unavailability of these variables actually caused the study to change. Of the nine variables used in the study, only three could actually be considered admission criteria, the three GRE scores. The remaining variables are more associated with the description or profile of a successful graduate student; college, gender, marital status, age, citizenship continent, and resident region. Although this is valuable information for a profile of successful graduate students, admissions decisions can not be based on these factors. It would be illegal to use several of the variables and impractical to use others. Assessing the predictive value of undergraduate information on graduate student success is extremely difficult, if not impossible, if this information is not available to researchers in a complete, accurate, and electronic format. Borchert (1994) in a report by the Council of Graduate Schools states that institutions should establish record keeping systems that allow them to analyze their admissions process. Diminnie (1992) stated the following:

> An important factor in evaluating the success of the admissions process is the determination of student success (or failure) as correlated with his or her admissions credentials...To effectively evaluate the admissions process, a centralized database that tracks the applicant pool is a necessity. (p. 35)

Colorado State University would be well served to better maintain student record databases to facilitate future analysis for admissions criteria and processes.

Recommendations for future practice and research

Practice. Caution needs to be taken in applying ANN technology in admissions decision making. Everson (1994) concluded that further research in the way neural networking computing can be used in conjunction with traditional statistical methods might improve our ability to accurately assess appropriate educational experiences for students. However, this researcher feels that further analysis and understanding of ANN technology is warranted before accurate practical application can be validated. Zhang et al. (1997) concluded that even though ANNs provide a great deal of promise, they also involve a great deal of uncertainty. Zhang et al. stated that the limited number of empirical studies result in the too common use of words such as "seem" and "appear." In drawing this conclusion they state several specific points which were also experienced in this research study. First, ANNs have no specific way to explain the relationship between inputs and outputs. ANNs have been termed "black-box methods". This makes interpreting the results very difficult and making comparison to other models arduous at best. Second, there are no confirmed methods to determine which network structure best fits the function. As a result, a process of trial and error and experimentation is used to determine network complexity and fit. Lastly, ANNs require larger data sets than traditional methods. When these large data sets are used in both ANNs and traditional methods, statistical significance is found but there may be little or no practical significance.

In addition to better understanding ANN technology, the use of different types of traditional statistical methods for assessing which admissions variables are accurate predictors of graduate student success may yield different results. This study utilized simultaneous multiple regression. Perhaps the use of other types of regression, such as stepwise multiple regression, would yield different results. Identifying different outcome variables, such as whether or not a student completes the degree, the use of logistic regression or discriminate analysis may also lead to different conclusions.

Monitoring and assessing admissions policies and criteria need to be ongoing processes. Diminnie (1992) concluded that monitoring the outcomes of admission policies and academic success should be a continuing process. Accurate and complete historical databases should be developed and maintained in such a manner to facilitate analysis. Researchers should have availability and access to these databases to regularly perform studies on admissions criteria, acceptance standards, and graduate student success.

Finally, there are ethical issues that need to be addressed. The use of variables such as ethnicity and age may or may not provide helpful information in regards to what makes a successful graduate student. Regardless of their informative value, basing admissions decisions on either of these factors is illegal. As researchers, is it ethical to include these variables in studies when no practical action can be taken based on the outcomes?

Future research. Other outcome variables to define graduate student success should be studied. Brown and Weaver (1979) and Nelson and Nelson (1995) used the outcome variable of degree completion as a determinant of graduate student success. Both studies found UGPA and GRE scores to be predictors of graduate degree completion. Another study by Fiedler et al. (1993) looked at career success as a determinant of graduate student success. They concluded that UGPA and an English proficiency test score were accurate predictors of graduate student career success. Tracking degree completion and employment of students after graduation may provide valuable insight into what type of admissions criteria are the best predictors of long term success.

The use of college as a predictor variable could be problematic. In fact, it is not an admission criterion. Different colleges have different admissions criteria and standards. It is difficult and perhaps not accurate to compare students across colleges. There are two proposed solutions to this. First, standardization of GGPA across colleges might reduce the effect that the variable college places on the other predictors. Another solution, and perhaps a better one, is to look at colleges on an individual basis. Nelson and Nelson (1995) studied admissions criteria across eight different colleges and concluded that predictors of graduate student success depended on the major area of study.

A final consideration for admissions committees and future researchers in the area of graduate admissions is that the studied admissions variables taken together were able to accurately predict 12% of the variance in GGPA. This leaves abundant room for additional assessments of a student's potential success or failure. Such assessments may include more qualitative requirements such as letters of recommendation, writing samples, interviews, portfolios, and prior professional experience. Clark and Palattella (1997) reported on a study by David Bordwell of the University of Wisconsin, Madison who concluded that the applicants writing sample was a more accurate predictor of applicant promise than the GRE scores. Hall and Baily (1992) concluded that although standardized test scores and UGPA had predictive value on first year GPA, interview scores outperformed both in predicting scores on a professional medical exam. Hagedorn (1996) concluded that portfolio assessment and group interviews were found to be more effective models for predicting the success of women, minorities and older students. What is needed is the development of quantitative models or procedures to objectively measure or evaluate these types of admissions requirements.

Conclusion

Admissions are a complex issue. Even though we study admission variables in an attempt to predict the success of graduate students, the final decision as to whether a student is admitted or not will always involve human judgment. This study was able to determine certain variables that could predict 10-12% of the variance in GGPA, and perhaps there are other variables which could improve that number, but ultimately humans will make the final decision. Gorr et al. (1994) and King et al. (1993) both concluded that human judgment was equal to or better in selecting successful graduate students than statistical models. When humans judge humans, some clues can be obtained from the numbers, but determining which students are admitted to graduate school is not solely a numbers game. This research supports these findings, in concluding that quantitative data can offer some level of help in determining which graduate students will be successful. This help, however, is small and offers little practical significance in our ability to predict individual student success.

REFERENCES

Association of American Universities. (1998, October). Association of American universities committee on graduate education: Report and recommendations. Washington, DC.

Baker, J.C., & Martin, F.G. (1998). A neural network guide to teaching. Bloomington, IN.: Phi Delta Kappa Educational Foundation.

Bennett, R.E., Morely, M., Quardt, D., & Rock, D.A. (2000). Graphical modeling: A new response type for measuring the qualitative component of mathematical reasoning. Applied Measurement in Education, 13.

Bennett, R.E., & Sebrechts, M.M. (1997). A computer-based task for measuring the representational component of quantitative proficiency. Journal of Educational Measurement, 34,64-77.

Berg, C. (1993, November). Evaluating an educational leadership doctoral program: Results of a focus group simulation for educational leadership doctoral students. A paper presented at the Annual Meeting of the Arizona Educational Research Organization in Tucson, AZ.

Berry, M.J.A., & Linoff, G. S. (1997). Data mining techniques for marketing, sales, and customer support. New York: Wiley.

Bezerianos, A., Papadimitrious, S., & Alexopolulos, D. (1999). Radial basis function neural networks for the characterization of heart rate variability dynamics. Artificial Intelligence in Medicine,15,215-234.

Bodri, L., & Cermack, V. (2000). Prediction of extreme participation using a neural network: Application to summer flood occurrence in Moravia. Advances in Engineering Software, 31,311-321.

Borchert, M.E. (1994). Master's education: A guide for faculty and administrators. A policy statement. Washington, DC: Council of Graduate Schools

Bowen, W.G. , & Bok, D. (1998). The shape of the river: Long-term consequences of considering race in college and university admissions. Princeton, N.J.: Princton University Press.

Brown, T.R., & Weaver, D.H. (1979, July). More than scores needed to predict graduate success. Journalism Educator. pp.13-15.

Brown-Miller, A. (1996). Shameful admissions: The losing battle to serve everyone in our universities. San Francisco, CA: Jossey-Bass.

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. (1997). Discovering data mining: From concept to implementation. Upper Saddle River, NJ: Prentice Hall.

Carbone, V., & Giuseppino, P. (1998). Palomar project: Predicting school renouncing dropouts, using artificial neural networks as a support for educational policy decisions. Substance Use and Misuse, 33,717-750.

Chance, D.C., MacLin, O.M., & Lykins, S. (1993). Multiple regression and artificial neural networks compared to multivariate nearest network technique. Proceedings of the sixth Annual Conference on Applied Mathematics in Edmond OK.

Clariana, R.B. (1998, March). Smarter tools, better teachers: Applying neural network technology to curriculum alignment. A paper presented at the ninth Annual Conference of the Society for Information Technology and Teacher Education. Washington, DC.

Clark, R.E., & Palattella, J. (1997). The real guide to graduate school: What you better know before you choose humanities or social sciences. US: Lingua France Books.

Cohen, J. (1988). Statistical power analysis for the behavioral scientist (2nd. Ed.) Hillsdale, NJ: Erlbaum.

Colbeck, C. (1994, November). The contexts of academic work: What matters most to faculty. A paper presented at the Annual Meeting of the Association for the Study of Higher Education Tucson, AZ.

Cross, T. (1996). Why the Hopwood ruling would remove most African Americans from the nations most selective universities. The Journal of Blacks in Higher Education, 11,66-70.

Dawes, R. M. (1971). A case study of graduate admissions: Application of three principles of human decision making. American Psychologist, 180-188.

Decker, P. (1998, March/April). Data mining's hidden dangers. Banking Strategies,74, 6-14.

Delaney, A.M. (1999). A collaborative approach to designing graduate admissions studies: A model for influencing program planning and policy. A paper presented at the Annual Forum of the Association for Institutional Research in Seattle WA.

Diminnie, C. B. (1992). An essential guide to graduate admissions. A policy statement. Washington, DC: Council of Graduate Schools.

D'Souza, D. (1991). Illiberal education: the politics of race a nd sex on campus. New York: Free Press

Enright, M.K., & Gitomer, D. (1989). Toward a description of successful graduate students (GRE Board Research Rep. No. 85-17R). Princton, NJ: Educational Testing Service.

Enright, M.K., Rock, D.A., & Bennett, R.E. (1998). Improving measurement for graduate admissions. Journal of Educational Measurement,35,250-267.

Everson, H.T. (1994, April). Using artificial neural networks in educational research: Some comparisons with linear statistical models. A paper presented at the Annual Meeting of the National Council on Measurement in Education in New Orleans, LA.

Fiedler, M.L., Foldesy, G., Matranga, M, & Peltier, G. (1993-94). Admission criteria used by NCATE accredited doctoral programs in educational administration. National Forum of Educational Administration and Supervision Journal, 10, 45-50.

Fisher, J.B., & Resniock, D.A. (1990). Standardized testing and graduate business school admission: A review of issues and an analysis of a baruch college MBA cohort. College and University, 65, 137-148.

Foley, J. & Russell, J.D. (1998, March 16). Mining your own business. Informationweek, 673, 18-20.

Garfield, L.Y. (1996). Squaring affirmative action admissions policies with federal guidelines: A model for the twenty-first century. Journal of College and University Law,22, 895-934.

Gorr, W.L., Nagin D., & Szczypula, J. (1994). Comparative study of artificial neural network and statistical models for predicting student grade point averages. International Journal of Forcasting,10, 17-34.

Graduate Curricula Development Task Force (1993, January). Evaluation of graduate education in the institute of food and agricultural sciences. Gainesville, FL: University of Florida.

Graham, L.D. (1991). Predicting academic success of students in a master of business administration program. Educational and Psychological Measurement, 51,721-727.

Hackmann, D.G., & Price, W.J. (1995, February). Preparing school leaders for the 21st century: Results of a national survey of educational leadership doctoral programs. A paper presented at the National Council of Professors of Educational Administration, New Orleans, LA.

Hall, F.R., & Bailey, M.A. (1992, February). Correlating students' undergraduate science GPAs, their MCAT scores, and the academic caliber of their undergraduate colleges with their first-year academic performances across five classes at Dartmouth medical school. Academic Medicine, 67, 121-123.

Hawkins, B.D. (1993). Educators push for diverse grad school admission criteria. Black Issues in Higher Education, 10, 26-30.

Jenkins, L. J. (1972). Predicting success in graduate school. Unpublished master's thesis, Colorado State University, Fort Collins.

Kezar, A.J. (1999). Higher education trends (1997-1999): Graduate and professional education. Washington, DC: The George Washington University. (ERIC Document Reproduction Service No. ED 435 346)

King, G., Bruce, J.M., & Gilligan, M. (1993). The science of political science graduate admissions. PS: Political Science and Politics,26, 772-778.

Kirnan, J.P., & Geisinger, K.F. (1981). The prediction of graduate school success in psychology. Educational and Psychological Measurement, 41, 815-820.

Lada, P., Brier, M.E., & Zurada, J.M. (1999). Therapeutic drug dosing prediction using adaptive models and artificial neural networks. Proceedings of the International Confernece on Neural Networks, USA, 5, 3669-3673.

Lipschutz, S.S. (1993). Enhancing success in doctoral education: From policy to practice. In LL Baird (Ed.), Increasing graduate student retention and degree attainment (pp.69-80). San Francisco, CA: Jossey-Bass.

Lykins, S., & Chance, D. (1992b). Comparing artificial neural networks and multiple regression for better predictability using correlation and absolute error statistics. Proceedings of the sixth Oklahoma Symposium on Artificial Intelligence. Tulsa, OK.

MacLin, O.H., Chance, D.C., & Lykins, S. (1993). Predictability comparisons between multiple regression, artificial neural networks, and abductive reasoning. Proceedings of the sixth Annual Conference on Applied Mathematics in Edmond OK.

Morgan, G. A., & Griego, O.V. (1998). Easy use and interpretation of SPSS for windows: Answering research questions with statistics. Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, G.A., Griego, O.V., & Gloeckner, G.W.(2001). SPSS for windows: An introduction to use and interpretation in research. Mahwah, NJ: Lawrence Erlbaum Associates.

Morrow, J.M. (1993, October). Revision, validation, and follow-up of admissions standards in counselor education. Paper presented at the Southern Association for Counselor Education and Supervision Convention in Charleston, SC.

Nelson, J.S., & Nelson, C.V. (1995). Predictors of success for students entering graduate school on a probationary basis. Muncie, IN: Ball State University.

Pandya, A.S., Kondo, T., Talati, A., & Jayadevappa, S. (2000). Foreign currency rate forecasting using neural networks. Proceedings of the Society of Photo-Optical Instrumentation Engineers, USA,4055,392-400.

Paolillo, J.G.P. (1982). The predictive validity of selected admissions variables relative to grade point average earned in a master of business administration program. Educational and Psychological Measurement, 42, 1163-1167.

Powers, D.E., & Enright, M.K. (1987). Analytical reasoning skills in graduate study: Perceptions of faculty in six fields. Journal of Higher Education,58, 658-681.

Rajecki, D.W., Lauer, J.B., & Metzner, B.S. (1998). Early graduate school plans: Uniformed expectations. Journal of College Student Development,39,629-632.

Reilly, R.R. (1976). Factors in graduate student performance. American Educational Research Journal, 13,125-138.

Sacks, H.S. (1978). Hurdles: The admissions dilemma in American higher education. Atheneum, N.Y.

Sadler, P.M., & Hammerman, J.K. (1999). Employing quantitative models of a qualitative admissions process: Uncovering hidden rules, saving time, and reducing bias. College and University, 74, 8-24.

Snyder, R.M. (1996, June). Neural networks for the beginner. Paper presented at the Association of Small Computer Users in Education summer conference. Myrtle Beach, SC.

Song, Q., & Chissom, B.S. (1993, April). New models for forecasting enrollments: Fuzzy time series and neural network approaches. Paper presented at the Annual Meeting of the American Educational Research Association in Atlanta, GA.

Statistica neural networks (1998). Tulsa, OK: Statsoft, Inc.

U.S. Department of Energy. (1994). Neural network definitions of highly predictable protein secondary structure classes (Report No. LA-UR-94-110). Los Alamos, NM: Lapedes, A., Steeg, E., & Farber, R.

U.S. Department of Energy. (1988). Application of neural networks and other machine learning algorithms to DNA sequencing analysis. (Report No. LA-UR-89-1788). Lapedes, A., Barnes, C., Burks,C., Farber, R., & Sirotkin, K.

University of California at San Diego (1999) Graduate School Report. [On-line]. Available: http://www-ogsr.ucsd.edu/graddata/99Fall/indices/index.htm#VI

Valafar, H., & Valafar, F. (1999). Prediction of a patient's response to a specific drug treatment using artificial neural networks. Proceeding of International Conference on Neural Networks, USA, 5, 3694-3697.

VanEyden, R.J. (1996). The application of neural networks in forecasting share prices. Haymarket, VA: Finance and Technology Publishing.

Vaseleck, J. (1994). Stop working and put down your pencils: the use and misuse of standardized admissions tests. Journal of College and University Law, 20,405-415.

Weiss, S., & Kulikowski, C.A. (1991). Computer systems that learn: Classification and prediction methods from statistics, neural nets, machine learning, and expert systems. San Francisco, CA: Morgan Kaufmann.

Wendel, F., Joekel, R.G., & Spurzem, C.W. (1991). Do graduate school admissions relate to assessment center dimensions. (Tech. Rep. No. 143). Lincoln, NE: University of Nebraska-Lincoln, Department of Educational Administration.

Youngblood, S.A., & Martin, B.J. (1982). Ability testing and graduate admissions: Decision process modeling and validation. Educational and Psychological Measurement, 42, 1153-1161.

Zhang, G., Patuwo, B.E., & Hu, M.Y. (1997). Forecasting with artificial neural networks: the state of the art. International Journal of Forecasting, 14, 35-62.

Zwick, R. (1993). The validity of the GMAT for the prediction of grades in doctoral study in business and management: An empirical bayes approach. Journal of Educational Statistics, 18, 91-107.