INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
TESTING THE TESTING AND GRADING PARADIGM FOR A COURSE IN ENGINEERING ECONOMIC ANALYSIS

A Dissertation
Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

by
Michael John Colburn, M.S.I.S.E.

***

The Ohio State University
1995

Doctoral Committee: Approved by:
David L. Boggs
Susan Imel
George L. Smith

Department of Educational Studies
ACKNOWLEDGMENTS

I would like to thank Dr. David Boggs, Dr. Susan Imel and Dr. George Smith for their wise counsel and support throughout this research project. I am indebted to the faculty team of Dr. Bob Lunquist, Dr. Clark Mount-Cambell, Dr. Jane Fraser, and Dr. George Smith for their sustained dedication in creating and implementing a learning design that reflects an enlightened vision for learning. I wish to thank Connie Snider for her administrative support, patience, and humor in helping me prepare this document. Finally, I would like to acknowledge my good friend, Gabe Reitter for his ‘not so gentle’ nudge that encouraged me to clear a space in my life to complete this project.
VITA

June 13, 1941  Born - Columbus, Ohio

1964  B.I.S.E. The Ohio State University
      Columbus, Ohio

1986  M.S.I.S.E. The Ohio State University
      Columbus, Ohio

1964 - 1967  Industrial Engineer
             The Timken Roller Bearing Company
             Columbus, Ohio

1967 - 1970  Staff Industrial Engineer
             Industrial Engineering Supervisor
             Anchor Hocking Corporation
             Lancaster, Ohio

1970 - 1973  Manager of Industrial Engineering
             Columbus Plant
             Ross Laboratories
             Columbus, Ohio

1973 - Present  President
                Colburn and Associates
                Columbus, Ohio

Major Field of Study: Adult Education
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. iii
VITA ......................................................................................... iv
LIST OF FIGURES ....................................................................... vii
LIST OF TABLES ......................................................................... ix

CHAPTER

<table>
<thead>
<tr>
<th>I. INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Statement of the Problem .................................................................</td>
<td>1</td>
</tr>
<tr>
<td>2. Driving Force - Personal Experiences ................................................</td>
<td>2</td>
</tr>
<tr>
<td>3. Driving Force - The Need to Improve Engineering Education ....................</td>
<td>6</td>
</tr>
<tr>
<td>4. Driving Force - Faculty Learning Inquiry Team .....................................</td>
<td>10</td>
</tr>
<tr>
<td>5. The Testing and Grading Paradigm ......................................................</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II. THEORETICAL FOUNDATIONS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction .....................................................................................</td>
<td>21</td>
</tr>
<tr>
<td>2. Total Quality Management ..................................................................</td>
<td>21</td>
</tr>
<tr>
<td>3. Adult Learning Theory and Principles .................................................</td>
<td>32</td>
</tr>
<tr>
<td>4. Self-Directed Learning ......................................................................</td>
<td>38</td>
</tr>
<tr>
<td>5. Learner-Control and the Role of the Facilitator ................................</td>
<td>41</td>
</tr>
<tr>
<td>6. Grading and Testing - Quality Learning and Student Empowerment ..........</td>
<td>44</td>
</tr>
<tr>
<td>7. High Quality Learning - Theory to Practice ......................................</td>
<td>51</td>
</tr>
<tr>
<td>8. Conclusions .....................................................................................</td>
<td>64</td>
</tr>
</tbody>
</table>
III. RESEARCH METHODOLOGY

1. Subjects and Research Design ................................................ 70
2. Engineering Economic Analysis
   Course Description .................................................................. 71
3. Response Variables .................................................................. 75
4. Data Analysis - Quasi Experimental Design
   Spring Quarter 1994 .......................................................... 77
5. Cross-sectional Data Analysis of Student Performance
   Autumn 1992 to Winter 1995 ............................................. 79
6. Cross-sectional Data Analysis of Student Satisfaction
   Autumn 1992 to Winter 1995 ............................................. 80

IV. RESULTS

1. Introduction ........................................................................ 82
2. Student Test Performance - Spring Quarter 1994 .............. 83
4. Student Projects ................................................................. 97
5. Cross-sectional Analysis of Student Performance .............. 101
6. Student Evaluation of Teaching Rating
   of Course and Instructor ..................................................... 105
7. Content Analysis of Students’ Comments ....................... 112

V. CONCLUSIONS

1. Introduction ........................................................................ 117
2. Summary of Results ............................................................ 117
3. Major Findings and Recommendations ......................... 126
4. Five Fundamental Flaws of Engineering Education ............ 131
5. Reengineering the Engineering Learning Process .............. 135
6. Building a Support System for Students and Faculty ........ 140

BIBLIOGRAPHY ............................................................................. 143

APPENDIX ....................................................................................... 147
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The Extended Production Process</td>
<td>23</td>
</tr>
<tr>
<td>2. Teacher-Student Control Spectrum</td>
<td>63</td>
</tr>
<tr>
<td>3. Student Test Performance Midterm One Spring 1994 Experimental and Control Classes</td>
<td>86</td>
</tr>
<tr>
<td>4. Student Test Performance Midterm Two Spring 1994 Experimental and Control Classes</td>
<td>87</td>
</tr>
<tr>
<td>5. Student Test Performance Midterm Three Spring 1994 Experimental and Control Classes</td>
<td>87</td>
</tr>
<tr>
<td>6. Student Test Performance Final Examination and Midterm Test Average Spring 1994 Experimental and Control Classes</td>
<td>88</td>
</tr>
<tr>
<td>7. Stem and Leaf Diagram Midterm One Test Re-Test Scores Spring 1994 Experimental Class</td>
<td>89</td>
</tr>
<tr>
<td>8. Stem and Leaf Diagram Midterm Two Test Re-Test Scores Spring 1994 Experimental Class</td>
<td>90</td>
</tr>
<tr>
<td>9. Stem and Leaf Diagram Midterm Three Test Re-Test Scores Spring 1994 Experimental Class</td>
<td>91</td>
</tr>
<tr>
<td>10. Student Test Performance Autumn 1994</td>
<td>94</td>
</tr>
<tr>
<td>11. Student Test Performance Winter 1995</td>
<td>95</td>
</tr>
<tr>
<td>12. Stem and Leaf Diagram Midterm Three Autumn 1994</td>
<td>96</td>
</tr>
<tr>
<td>13. Percent of Students Achieving Midterm Average of 85% or Higher</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Title</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>14</td>
<td>Percent of Students Earning a Course Grade of B or Greater</td>
</tr>
<tr>
<td>15</td>
<td>Attribute Control Chart S.E.T. Question 17</td>
</tr>
<tr>
<td>16</td>
<td>Attribute Control Chart S.E.T. Question 5</td>
</tr>
<tr>
<td>17</td>
<td>Percent of Positive Responses Open Ended Questions Testing &amp; Grading</td>
</tr>
<tr>
<td>18</td>
<td>Attribute Control Chart S.E.T. Question 18</td>
</tr>
<tr>
<td>19</td>
<td>Attribute Control Chart S.E.T. Question 19</td>
</tr>
<tr>
<td>20</td>
<td>Attribute Control Chart S.E.T. Question 20</td>
</tr>
<tr>
<td>21</td>
<td>Attribute Control Chart S.E.T. Question 6</td>
</tr>
<tr>
<td>22</td>
<td>Attribute Control Chart S.E.T. Question 7</td>
</tr>
<tr>
<td>23</td>
<td>Attribute Control Chart S.E.T. Question 8</td>
</tr>
<tr>
<td>TABLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1. Conversion of Test Score Averages to Grades</td>
<td>72</td>
</tr>
<tr>
<td>2. Points for Learning Projects</td>
<td>74</td>
</tr>
<tr>
<td>3. S.E.T. Questions, Testing and Grading System</td>
<td>76</td>
</tr>
<tr>
<td>4. Student Characteristics Spring Quarter 1994</td>
<td>84</td>
</tr>
<tr>
<td>5. Distribution of Test Grades for Midterm One Before and After Re-Test Option Spring 1994</td>
<td>93</td>
</tr>
<tr>
<td>6. Distribution of Test Grades for Midterm Two Before and After Re-Test Option Spring 1994</td>
<td>93</td>
</tr>
<tr>
<td>7. Distribution of Test Grades for Midterm Three Before and After Re-Test Option Spring 1994</td>
<td>93</td>
</tr>
<tr>
<td>8. Distribution of Projects Completed at Four Levels Autumn 1994</td>
<td>99</td>
</tr>
<tr>
<td>10. Types of Project Options Completed by Students Autumn 1994</td>
<td>99</td>
</tr>
<tr>
<td>11. Types of Project Options Completed by Students Winter 1995</td>
<td>100</td>
</tr>
<tr>
<td>12. Attribute Control Chart Data S.E.T. Questions 17 through 20</td>
<td>106</td>
</tr>
<tr>
<td>13. S.E.T. Questions 5 through 8</td>
<td>108</td>
</tr>
<tr>
<td>14. Attribute Control Chart Data S.E.T. Questions 5 through 8</td>
<td>108</td>
</tr>
</tbody>
</table>
15. S.E.T. Questions 1 through 4 and 9 through 16 ............................................... 110

16. Attribute Control Chart Data S.E.T. Questions 1 through 4 and 9 through 16 .............................................................. 110
CHAPTER I

INTRODUCTION

1. Statement of the Problem

The traditional coupling of testing and grading is a deterrent to learning. At best it is benign. At worse it is abusive. The purpose of this research is to design and test a model of testing and grading that facilitates learning. The new model of testing and grading is grounded in the principles of adult learning theory and the philosophy of total quality management. This research posits that adult learning theory and total quality management are philosophically compatible bodies of knowledge that are complementary and mutually supportive of the learning process.

In this section I define and contrast the terms testing and grading. Next, I review three forces that have driven this research. The first driving force is my personal experiences as an instructor in an undergraduate engineering course in total quality management. The second driving force is the emergence of five major paradigm shifts identified by a team of engineering faculty who engaged in a mutual learning inquiry process. The third driving force is the participation of the Ohio State University’s College of Engineering, along with nine other universities, in an experiment to reengineer the undergraduate engineering curriculum. Finally, I hypothesize the impact of a new system of testing and grading on student performance and satisfaction.
The term **testing** is used to describe the feedback mechanism to determine the extent to which learning has occurred. The purpose of testing is to provide information for diagnosis of the learning process. Diagnosis leads to decisions and actions for the improvement of the learning process. The term **grading** is used to describe the assessment of a level of competence that the learner has achieved at the “conclusion” of the learning process. Grading is a certification process.

2. **Driving Force - Personal Experiences**

In 1991 I was asked to design and teach a course in total quality management for undergraduate industrial engineering students. The chair of the Industrial and Systems Engineering at the Ohio State University felt it was time to update and improve the department’s course in statistical quality control. The course had a strong emphasis in traditional quality control, but it was lacking in the methods of continuous improvement and in the philosophical foundation of total quality management. The department felt a strong commitment to provide industrial engineering students with the knowledge and tools necessary to be active participants in the quality movement. As a practitioner in quality improvement I accepted this challenge with enthusiasm. The enthusiasm was fueled by the important role I believed that industrial engineers could play in organizational quality improvement. In designing this course my one guiding principle was: **To learn about quality, students need to experience quality in their learning situations.** As I designed the course I asked: What does quality learning look like? How can I build quality into each activity? How can I encourage students to exercise greater
self-direction in their learning? These questions helped me formulate a set of course design criteria:

- Choose a text and create a packet of learning materials that reflected the philosophy and techniques of total quality management ... materials that were cutting edge and that would be of value after the course.
- Invite quality practitioners to participate in this course --- to expose the students to best practices.
- Create learning experiences that bring concepts to life and help the students develop skills by using them right away.
- Establish high quality standards for the course and for each of the learning modules - to help each student achieve high quality performance.
- Encourage students to be self-directed - to pursue their personal and professional interests in quality improvement.
- Establish a grading system that would support high quality performance, continuous improvement, and personal autonomy.

The establishment of a high quality testing mechanism and a high quality grading system challenged me in the course design.

Following are the features of the testing and grading systems used to guide the design of the course in total quality management:

- The course had four basic modules. In the first module the students were asked to demonstrate their understanding of the principles of total quality
management by developing a case based on their own experiences of quality. The other three modules dealt with the demonstration of the use of statistical methods for quality improvement. The students were given comprehensive tests that included problems as well as open-ended questions to demonstrate their competence and understanding.

- All tests were “take-home”. Students were given a few days to complete the tests. A minimum score of 85 was required to demonstrate competence, and a score of 95 was a demonstration of high competence. The students were given specific feedback on their test responses. They were encouraged to study areas that they missed and then take another test to improve their score. Student could continue to be tested until they received the score they desired.

- A demonstration of competence in all basic modules was necessary to complete the course. A grade of “B” was earned for this competence.

- Students could initiate one or more inquiry projects to pursue their own learning interests. They were shown how to create a learning contract and could contract for quality points based on the scope of the project.

- Quality points were the basis for earning an “A” after competency was demonstrated in the basic modules.

This system, a combination of mastery learning and learning contracts, was designed to meet the grading criteria.
I taught this course twice and was generally encouraged by the quality of the students' performance and their satisfaction with the learning process. Following are some of the things I learned:

- Many students required more structure than anticipated. The second time I taught the course I gave them brief quizzes at the beginning of each class to encourage their preparation. First-test performance for the basic modules improved significantly for the second course offering. I attribute the improvement to the change in structure. By the fifth week of the course, quizzes were unnecessary.

- Many students struggled with the inquiry projects. Early in the course, they didn't know what they didn't know. Some had difficulty in exercising the autonomy that I offered. Some had a lot of fun with the project and their creativity added to the quality of the class.

- Although students were required to take statistics as a prerequisite, I was disappointed in their ability to apply statistics to practical problems. I needed to do some remediation in the area of statistics.

- Most students responded well to the grading system. The overall quality of their work improved throughout the quarter. Based on the analysis of student feedback (departmental questionnaire) and informal feedback, I concluded that most students accepted personal ownership of their grade. Our discussion after a test was more on what they needed to learn, not on how many points I
awarded for a problem. A few students, however, questioned the “fairness” of requiring “extra work” (inquiry projects) to get an “A”.

- As the course instructor I struggled a bit in getting the students to take a more active role in the class. Many students expected me to lecture and for them to listen. Introducing experiential learning activities and rearranging the physical arrangement (semi-circular vs. “classroom style”) of the room helped encourage participation.

Overall this experience encouraged me to continue to improve as a facilitator of learning. On a personal basis I have discovered that the application of adult learning principles were congruent with the principles of total quality management. The purpose of this research is to explore the congruence of adult learning theory and total quality management and to design and test a model of testing and grading that would improve the quality of learning for undergraduate engineering education.

3. Driving Force - The Need to Improve Engineering Education

Dr. John White (1993) Dean of the College of Engineering at Georgia Tech is critical of the quality of undergraduate engineering education. He suggests that colleges and universities are facing a three-fold challenge: (1) a decline in the size of college-age population; (2) a reduced attractiveness of science and engineering among the “best and brightest”; (3) demographic shifts to populations (minorities, women, disabled) who have historically not opted for engineering and science. White, criticizing undergraduate education, cites engineering colleges performance (or lack of performance) against four
quality criteria. First, just 65% (35% for minority students) of the students entering engineering obtain a degree in this field. Second, it takes the typical student five years to complete a four year program. Third, less than half of the students who declare an initial interest in engineering are ultimately satisfied with the experience. Fourth, the cost of higher education has risen faster than the cost of living. White concludes that students are questioning the “total cost” (financial and psychological) of majoring in engineering.

White attributes the low quality of engineering education to a number of reasons:

- The lack of understanding of who is the customer of education. He asserts that students are the primary customers, and future employers, secondary.

- Too much dependence on selection of students and not enough on the development of students.

- Lack of a mechanism to cope with inherent variation in the learning process.

- Too great an emphasis on producing Ph.D’s and not enough on undergraduate degrees.

White suggests that academia put more effort in listening to its customers (students, employers and faculty). He challenges educational institutions to pay more attention to those who are at the margin. He suggests that with a developmental approach, these same students could be superstars. He says:

We must develop a mentoring capability within engineering, we must pay more attention to the development of the human resources made available to us; we must acknowledge their scarcity; and we must replace a ‘weeding out’ philosophy with a ‘bringing in’ or ‘cultivating philosophy. (p. 62)
White compares the current quality management of education with the mistake that industry made for years in trying to inspect quality into the product instead of building quality into the process. He asserts that testing and grading practices exemplify this same 'quality by inspection' philosophy. He goes on to say:

In my judgment, when a student fails a course, it is because of a failure in the teaching/learning process; both the student and the professor are jointly culpable. However, under our current system, only the student pays the penalty for the lack of successful education on the part of the professor. (p. 62)

White’s criticisms of the education community is a challenge to its current paradigm of what business it is in; to its priorities, and to the process of education itself.

In the wake of White’s criticisms, Drexel University has taken bold steps in reengineering its curriculum and creating a high quality learning process for its undergraduate engineering students. Quinn (1992) reported on the progress that Drexel University has made in a five year experimental program entitled, An Enhanced Educational Experience for Engineering Students. He cited several national studies that indicated deficiencies in undergraduate engineering education. Among these were the need for integration, synthesis, design, non-technical skills, and preparation for life-long learning. Drexel responded by restructuring its entire program. Some of the major elements include:

- Moving from 25 individual courses to four integrated areas of study:
  - engineering fundamentals, mathematical and scientific foundations, engineering laboratory, personal and professional enrichment.
• Team teaching and integrating content areas.

• Focus on application and design from the beginning of the student’s academic experience.

• Personal and professional enrichment including, journal keeping and self-directed learning.

• Frequent feedback for continuous improvement.

This program continues to use the co-op design that Drexel has successfully used throughout the years. The instructional technology additions (each student had his/her own personal computer and a comprehensive set of software) were made to improve the effectiveness and efficiency of education. For example, students could work on some educational modules while on their co-op assignment.

Although the program is still in its early stages, Quinn indicated that the preliminary evaluations are encouraging. These evaluations include high performance in computer, laboratory, and communication skills. Student retention rate is exceeding that of the traditional engineering education and student satisfaction has been positive.

The Drexel experience has resulted in the National Science Foundation funding of the Gateway Project which includes the implementation of the Drexel Design in ten universities throughout the United States. Ohio State is one of the participating universities where a group of freshman and sophomore students are currently participating in a pilot program. The Drexel Design has a strong adult learning focus and has at its core the improvement in the quality of undergraduate education.
4. Driving Force - Faculty Learning Inquiry Team

In the spring of 1993 I proposed to work with a team of faculty members in the Department of Industrial and Systems Engineering to build and test an education model that integrated the principles of total quality management and adult learning theory and apply this model to the design of a course. Four faculty members and I formed a team to explore the redesign of a course in Engineering Economics. One faculty member was the department chair, one faculty member taught the course in Engineering Economics, and two faculty members served on the Engineering College’s Engineering Economics Committee. This course design would be based on the principles of total quality management and adult learning theory. We decided to begin this exploration by forming ourselves into a learning inquiry team. During the summer each faculty member read two books and several articles on total quality management and adult learning theory. Each team member kept a personal journal to reflect on their personal insights during a period of independent study.

The readings on adult learning included Brookfield’s (1988) *Understanding and Facilitating Adult Learning*, chapters of Knowles (1980) *The Modern Practice of Adult Education*, and a number of articles on learning theories and practices. The readings on total quality management included Walton’s (1986) *The Deming Management Method*, a series of articles on total quality management by D. Scott Sink (1989); and several articles on the application of total quality management in the educational setting.
Knowles’ discussion of assumptions about the learner helped the team members question their assumptions about their students and the implications for course design. The model of learner assumptions that Knowles calls andragogy, is that as individuals mature: their self-concept moves toward self-direction; their experience becomes a rich resource; learning relates to the developmental tasks of their social roles; and their time perspective is immediacy of application, moving towards a performance centered approach to learning. Knowles reviews seven conditions of learning that are superior for the application of andragogy. These conditions provide a set of criteria for designing a learning process. Knowles assumptions of andragogy and pedagogy helped the faculty team examine and question their own learning paradigms.

One of Brookfield’s (1988) contributions is his identification of five themes that are central to self-directed learning. These themes are (1) the centrality of the learning contract, (2) preparation for self-directed learning, (3) peer learning groups, (4) the time commitment required by faculty to support self-direction, and (5) the perceived benefits and frustrations of learners. Brookfield challenges the earlier myths of self-direction as a solitary learning venture and that the facilitator was merely a resource person. In his six principles of effective facilitation he paints a picture of an active partnership between the learner and the facilitator. These principles include (1) participation in learning is voluntary, (2) there is a respect among participants for each other’s self worth, (3) facilitation is collaborative, (4) facilitation involves reflection and meta-learning, (5)
facilitation fosters a spirit of critical reflection, and (6) the aim of facilitation is the nurturing of self-directed, empowered adults.

Clearly, the fostering of self-direction and democratic values is central to adult learning theory and practices. It is based on humanistic principles and the potency of intrinsic need satisfaction. What adult learning does not address as clearly is the legitimate role of multiple stakeholders in the outcomes of higher education. With multiple stakeholders (students, employers, society) how is quality defined and evaluated? Can these answers be reconciled with the humanistic and democratic values of adult learning?

Total quality management is a term to describe a major worldwide emphasis on quality and organizational performance. Deming (1986) criticized traditional American management and proposed fourteen points as a prescription for quality improvement within organizations. Sink (1989) provides an organizational change model including an extended production model (from suppliers to customers) and identifying five key quality checkpoints that must be addressed if quality is to be total. Six recurring themes emerge from the spokespersons of the quality movement. These include (1) Customer Orientation - that the customer defines quality and there are multiple customers, (2) Employee Empowerment - that employees are customers, want to be involved, and have intrinsic as well as extrinsic needs, (3) Leadership - that change needs to be lead from the top and that people can exercise self-direction with skill and good will, (4) Measurement - that quality performance needs to be defined and operationalized so that feedback can drive continuous improvement, (5) Continuous Improvement - that poor systems are the source
of most quality problems and that the system must be analyzed and improved, and (6) Training and Development - that new tools and skills are required to transform business.

The focus of total quality management is on improving organizational performance. This is in contrast, but not necessarily in conflict, with adult learning's emphasis on individual development. What these two areas have in common are the assumptions they share about the nature of people. Both adult learning theory and total quality management are based on humanistic values, the potency of intrinsic needs, and the skill and good will of people to work cooperatively. These are the values that bring congruence to these two bodies of knowledge. Educational institutions, particularly those interested in improving quality, need adult learning and total quality management concepts to accomplish their missions.

In the Autumn of 1993 the learning inquiry team, after a period of individual study about adult learning theory and total quality management, came together and invested two days in intensive dialogue. This dialogue concluded with a commitment to work together to apply the principles of adult learning and total quality management to the education of undergraduate engineering students. Their project mission statement affirms this commitment:

Our mission is to design and implement a course in engineering economics that incorporates the principles and philosophies of total quality management and adult learning theory. Our customers are the students of this course, other faculty and departments in which the course is requisite, and employers for whom our students will begin their careers as engineers. We will actively seek the input from all of our customers and will be sensitive to these needs in our design. We are committed to creating a learning environment which shows respect for the students as responsible adults and active partners in their journey toward high quality learning. We are committed to helping students develop mastery in engineering
economics, grow as human beings and become lifelong learners. Finally, we will work together with diligence, good will and humor as we co-create together in this worthy venture.

During this dialogue the faculty learning team also identified five major paradigm shifts that would be initiated in the design and implementation of the Engineering Economics course. These were:

1. The course design will move from faculty-driven to customer-driven. There will be an explicit recognition and response to the needs of all of the customers. We are committed to building a mutually beneficial partnership among all customers, respecting each unique perspective.

2. Testing and grading will move from an after-the-fact faculty evaluation of student’s performance to a real-time feedback mechanism that results in continuous improvement of the learning system. We will distinguish grading (certification of competency) from testing (feedback for diagnosis). Students use testing to diagnose their own learning needs. The faculty use testing to diagnose effectiveness of learning interventions. Grading is a process to ensure learning mastery.

3. We will move from an individual to a team focus. Cooperative team approaches will be used in the course design as well as in the learning and teaching methods used. Students and faculty will collaborate as fellow learners in an atmosphere of mutual respect and trust. All customer groups will be invited to engage in collaborative learning for the mutual benefit of all.

4. We will move from content mastery to learner mastery. Content mastery focuses on understanding current knowledge. Learner mastery focuses on profound knowledge of the content and self understanding. We will incorporate the activity/reflection cycle in the course design and implementation. Self-knowledge is a requisite for lifelong learning. Learning how to learn is a requisite for personal empowerment.

5. The role of faculty will move from lecturer to learning facilitator. The role of the student will move from passive recipient of information to an active partner in learning. Together this partnership builds a creative and synergistic learning environment.
The faculty team has been using these five paradigm shifts to guide the redesign of this course. This research focuses on the second paradigm shift on testing and grading. The testing and grading paradigm shift is central to high quality learning because it addresses issues that are central to total quality management and adult learning theory. From the perspective of total quality management, a high quality testing and grading design needs to focus on the improvement of processes, the elimination of fear, the encouragement of teamwork, and the achievement of high quality performance. From the perspective of adult learning theory, a high quality testing and grading design needs to focus on the satisfaction of intrinsic needs, the encouragement of self-direction, personal responsibility and lifelong learning. Testing and grading is a leading paradigm shift for education because it changes the nature of the relationship between the student and teacher. There is a shared responsibility for the success of the learning process. The teacher has the content expertise, knowledge of learning standards, and the understanding of the legitimate needs of other stakeholders (other academic units, employers, society). With these the teacher can guide and challenge students in their quest for high quality learning. The students have the greatest personal stake in the quality of the learning process. What is at stake are the fulfillment of the students' need to increase knowledge, improve competence and satisfy their personal needs. A healthy student-teacher relationship is based on the recognition and acceptance of their individual and mutual responsibilities to manage the learning process. This healthy interdependence has the potential to create a positive learning environment that will support the development of the other four paradigms.
5. The Testing and Grading Paradigm

The fundamental premise of this research is that the close-coupling of testing and grading corrupts the learning process. By close-coupling I mean that testing and grading are treated as a single process. The impact on the learning process is that it: (1) diverts attention away from diagnosis; (2) creates an unhealthy power imbalance between teacher and student; and (3) discourages deep-level learning.

It is the premature assignment of a grade that gets in the way of the diagnostic value of a test. After the test the student may argue for points, criticize the teacher for unfair test questions, criticize self for errors, or praise self for matching the teacher’s answers. The test grade with all of its finality severely limits its diagnostic value. The purpose of diagnosis is to gain understanding for the purpose of improving the learning process. Diagnosis should include the analyses of the learning process of the student, the teaching process of the teacher, and any other learning support activities. A “failed” test might create insights for learning that an “A” might miss. It is hard to celebrate a failure when it becomes a source of punishment. The presumption that students bear the full responsibility for test performance limits the teacher from accepting responsibility for deficiencies in the teaching process.

The power imbalance that is created by the close-coupling of testing and grading is both direct and subtle. The assignment of a grade to a student by a teacher is inherently imbalanced. It is an imbalance that the traditional grading paradigm does not even question. It is presumed that the role of the teacher is to assign grades. The fair teacher is
a benevolent dictator. In the traditional paradigm it is the role of the teacher to provide clear expectations and fair evaluation standards. It is the role of the student to anticipate the teacher's needs. In the traditional paradigm the test may be fair, but it is rarely intrinsically satisfying. There is a power imbalance, not just in who judges performance, but in what is being judged, how learning will be demonstrated and when learning will be evaluated. In the traditional testing and grading paradigm all of these are at the discretion of the teacher.

Deep level learning includes the pursuing self-knowledge, understanding the idiosyncratic nature of learning, and appreciating the paradoxes inherent to knowledge. Outstanding teachers help their students achieve deep level learning in a variety of ways. But testing for the purpose of grading is not one of these ways. In the interest of "fairness", tests are constructed for the ease and consistency of scoring. This need to be fair often robs the teacher of a powerful learning tool --- the ambiguous question. The teacher struggles with the assignment of points and the students complain about the fairness of the question. The "fair" instructor replaces the deep, ambiguous question with a trivial, clear one. And what is lost is a dialogue that could stimulate learning for the students and the teacher.

The purpose of testing should not be to assign grades. Testing should stimulate learning. It should provide information to help the student diagnose learning needs -- his or her understanding of the content and of his or her personal learning process. It should provide information to help the teacher diagnose his or her teaching process, and to gain
understanding of the students’ learning process. The assumptions behind this approach to testing is that the student wants to learn and the teacher wants to teach.

The purpose of grading is to certify that the student has achieved competency and mastery. The student has the responsibility to demonstrate competency and mastery. The teacher has the responsibility to ensure that high quality standards are established. This is done in collaboration with the teacher, students and other legitimate stakeholders in the learning process. The teacher mentors and challenges the student to participate in the process and to accept psychological ownership of the standards. An effective grading system should: be based on high quality standards; be non-competitive; and enable the students to exercise high control.

I propose that testing and grading need to be separate activities so that the legitimate purpose of each can be achieved. In chapter two a theoretical framework is laid for a testing process and a grading process that supports high quality learning.

This research evaluates the impact of a such a testing process and a such a grading process on student performance and satisfaction. The hypotheses that this research tests are:

Hypothesis 1 - For Spring 1994 student performance, as measured by scores obtained on tests and final grades, will be greater for the class receiving the treatment of the new testing and grading system when compared with the class using the traditional testing and grading system that had been used in the course.
Hypothesis 2 - For Spring 1994 the variation in student performance, as measured by the variance of test scores obtained on tests and final grades, will be less for the class receiving the treatment of the new testing and grading system when compared with the class using the traditional testing and grading system used by the control class.

Hypothesis 3 - For Spring 1994 student satisfaction with the course and the instructor, as measured by the students' ratings, as measured by the Student Evaluation of Teaching report, will be higher for the class receiving the treatment of the new testing and grading system when compared with the class using the traditional testing and grading system used by the control class.

Student performance and satisfaction will be explored and analyzed using test performance data and S.E.T ratings for four course offerings prior to the Spring 1994 experiment and for two course offerings after the Spring 1994 experiment. This research also analyzes qualitative information concerning student reaction to the course. This is to gain insight into the quantitative data on student performance and satisfaction.

This research study is limited to a single course, Engineering Economic Analysis for Engineers, taught by one professor. The professor teaching the course is a member of the faculty learning inquiry team that has created the five paradigm shifts. This professor has been the only faculty member that has taught the course over the last several years. The generalizability of results is limited to this course as taught by this professor. What is not limited is the opportunity to track student performance, satisfaction and response to a radically different philosophy of testing and grading over multiple course offerings. The
second limitation concerns the operational definition of student performance. Because test performance is used, the impact on long term retention of course content and the ability to perform as an engineer will not be addressed.
CHAPTER II
THEORETICAL FOUNDATIONS

1. Introduction

In this chapter a theoretical foundation will be laid to build a grading and testing system that incorporates the principles of total quality management (TQM) and adult learning theory (ALT). Total quality management will be reviewed with a discussion of what it might look like in an educational setting. Adult learning theory will be reviewed showing its fundamental links to total quality management. Concepts concerning self-direction and personal autonomy will be reviewed and discussed in terms of empowerment and the nature of the student-teacher relationship. TQM and ALT will be integrated into a model of high quality learning, specifically in relation to the assessment process. Exemplary practices of high quality learning within educational institutions will be reviewed and criteria suggested for the design of a high quality assessment process. Finally, hypotheses will be generated to predict the impact of this high quality assessment process within higher education.

2. Total Quality Management

Total Quality Management (TQM) has been an umbrella and a rallying point to describe a fundamental shift in the relationship between the organization and its stakeholders. The term stakeholders include: customers, suppliers, shareholders, government, the community, and their employees. For industry in the United States,
particularly manufacturing, the path to Total Quality Management has been driven by
global competition and the widespread belief that the quality and productivity of American
business was falling behind the rest of the world. In the 1970s and 80s, Japan was
capturing market share. “Made in Japan” was becoming identified with high quality, in
sharp contrast to its image in the 50s and 60s. If there was an incident that was symbolic
of America’s wake up call, it was in November 1979 when the National Broadcasting
System (NBC) aired its white paper, “If Japan Can, Why Can’t We?” This television
program re-introduced W. Edwards Deming to America. Deming, one of the quality
leaders in the United States during the second world war, became a voice unheard in the
post-war era. The United States was the unquestioned economic power in the world in
the 50s and 60s. Deming’s quality principles, largely ignored in this country, were
embraced, however, by the technical and industrial leaders of Japan. The application of
his principles is credited with the development of Japan as a dominant economic power in
the second half of the twentieth century. Japan’s highest recognition for individual and
organizational contribution to quality enhancement are called the Deming Awards.
Deming’s principles are now widely known in this country as America’s wake up call for
high quality organizations is being heard.

But what does this have to do with education? Are the principles of total quality
management applicable to running a classroom? In this section I will review some of the
foundational principles of total quality management and discuss their implications for
higher education.
Quality is a judgment by customers to the extent to which a product or service meets their needs. The management of quality includes all those things that must be done in the design, transformation, delivery and support of the product/service to assure that the customers needs are being met and exceeded. Sink (1989) describes five quality checkpoints that need to be attended to in the creation of a product/service to ensure that the management of quality is total. Figure 1 shows this model.

![Figure 1: The Extended Production Process and Five Quality Checkpoints](image)

**Providers** are the suppliers of the inputs that the organization needs for its value-adding processes. For a higher education course these inputs may include text books, laboratory equipment, prerequisite knowledge, teaching resources, and library services. Providers include people and organizations inside and outside of the organization. The Q1 quality checkpoint includes all of the things that are done to ensure that the providers understand our needs and have the ability to supply us with consistently high quality inputs. At the Q2 quality checkpoint the organization inspects the providers' inputs to ensure their quality specifications are being met. For example, one of the inputs for an
engineering course in quality improvement, might be the knowledge and skills of the
students to apply statistical methods to quality improvement. If the prerequisite statistics
course does not teach the students the necessary statistical concepts or does not help the
students develop an understanding of how to use these concepts, then the quality of the
student’s learning process in the quality improvement course will be jeopardized. The
students may not adequately grasp new principles, or the quality improvement instructor
may not be able to cover all of the course material because of educational rework.

In the traditional management system, the relationship between the organization and
its suppliers is often distant. There may be multiple suppliers for a single resource and the
selection of suppliers may be based on the lowest bid. If there is no activity to ensure the
quality of incoming resources, then the only tactic that can be used to achieve quality is to
inspect out the defects or to rework the resource to bring it up to specifications. This
inspection and rework is a Q2 activity. Deming (1986) criticized the practice of quality by
inspection when he said: “Inspection to improve quality is too late, ineffective, costly.” (p.
28) He was speaking of the inspection of incoming resources as well as inspection within
the organization. He recommended that organizations develop a long term relationship
with their suppliers based on cooperation and trust. Deming recommended ending the
long standing practice of awarding business on the basis of price tag alone and moving
toward single source relationship with key suppliers. Customer and supplier should work
together to improve each other’s processes for higher quality and lower total costs.
Using Deming's approach, the instructor of the quality improvement course would develop a relationship with the instructor of the prerequisite course in statistics. The quality improvement instructor would provide information on the types of statistical knowledge, skills, and applications required for the quality improvement course. The quality improvement instructor might develop example problems that the statistics instructor could incorporate into his course. Each might visit the other instructor's classes. The statistics instructor would see how the application of statistics occurs in another course. The quality improvement instructor could gain insights so that bridges could be built between theory and practice. Based on the increase in quality of student preparation, the quality improvement instructor would encourage students to take this particular statistics course. The impact would be to improve the quality of the learning experience of the students. The management of the learning process extends out from the classroom and the instructor can play a key learning management role. This management role by the instructor could be extended to other resources as well.

The value-adding processes are the transformation processes that are used to create high quality products and services. The Q3 quality checkpoint includes all the things done to ensure that quality is being built into the transformation process. For a manufacturing process, Q3 activities could include preventive maintenance practices, the selection and use of proper tools, development of innovative work methods and processes, training and development opportunities, and information systems. Q3 activities are those things that are done to build quality into the production process. Deming and other quality
improvement practitioners have given industry a variety of tools to analyze and improve their production processes. These include process flow charts, pareto diagrams, cause-effect diagrams, scattergrams, and statistical process control charts. These techniques will not be explained in any detail here, but I will review some basic principles about their use to assure continuous improvement of the production process.

PRINCIPLE ONE - To improve the quality of the process there needs to be an objective presentation of the current reality. The current processes need to be described and analyzed on a factual basis. The quality improvement tools require data that will enable the process owners to understand the nature of the transformation process and to identify the factors that are currently limiting capacity, quality and productivity.

PRINCIPLE TWO - All processes are subject to variation. When performance varies, we want to know “Is this variation within the expected capability of the system, or has something (whether known or unknown) happened that has caused performance to decrease or increase?” This question underscores one of the most significant issues that Deming has raised and championed: the issue of special cause and common cause problems.

Special cause problems are those that can be attributed to something acting on the process. For a manufacturing process, for example, quality defects may be higher than ‘normal’ because of a defective supply of material, or the improper set-up of the equipment, or any number of factors that did not allow the process to produce to its quality capability. A signal of a special cause problem is variation in performance that
cannot be attributed to chance. When this occurs, the process is said to be out of statistical control. This analysis requires what Deming calls statistical thinking along with data display tools to detect problems and determine process capability. To solve a special cause problem, there needs first to be a recognition that there is a special cause problem. Next, through deductive reasoning the root cause of the problem needs to be determined. Finally, an action must be taken to eliminate the cause, now and in the future.

Common cause problems are those that can be attributed to the process itself. For a manufacturing process, for example, a goal may have been set by management to reduce the cycle time of a process, or to increase the quality of the product. If this goal is not within the current capability of the system, then the goal can not be consistently achieved. The system itself must be changed for improvement to occur. Deming soundly criticized American management for confusing common cause problems as special cause problems saying:

A fault in the interpretation of observations, seen everywhere, is to suppose that every event (defect, mistake, accident) is attributable to someone (usually the one nearest at hand), or is related to some special event. (p. 314)

So managers attribute mistakes to careless employees, airlines attribute crashes to pilot error, and professors attribute low grades to unmotivated students. The list goes on. This practice of assigning special causes to system issues breeds an environment of blame finding and frustration. A front line employee becomes frustrated and feels powerlessness because she or he lacks the resources to change the system. No matter how much effort the employee expends, the system stays the same. Deming goes on to say:
I estimate that in my experience most troubles and most possibilities for improvement add up to proportions something like this: 94% belong to the system (responsibility of management) and 6% special. (p. 315)

PRINCIPLE THREE - The people who are directly responsible for the transformation process need to be empowered to improve their own process.

Employees are well aware of common cause problems. They have learned how to do their job in spite of the system. But it has been at a cost . . . to the individual and to the organization. Front line employees have the technical skill to run the process. They just need a better process. But the word ‘empowerment’ does not create empowerment. To become truly empowered front line people need (1) a work environment that supports continuous improvement of systems and processes, (2) training in process analysis and improvement methods, and (3) resources to implement improvements.

The review of Q3 quality management has been made in the context of a manufacturing process. What does it look like for the instructor of the quality improvement course? Before answering this question, more basic questions need to be asked. Who are the customers of the course? What are the value-adding processes? What products/services are created by these processes?

The primary customer of the course is the student. Secondary customers may be industry and society. The term secondary is used to indicate that these customers do not receive the direct output of the course, but that they have a stake in the success of the primary customers, and are therefore a part of the extended production process. The product of the course is the education of the student. Tribus (1993) expands on this in a
most profound way, answering the question: What is the product of schools of business and engineering?

To begin, let us agree that the student is not the product. The product is the education of the student. In the “manufacture” of this product, as with any other product, it is essential that the “worker” (student) be an active participant in the design and creation of the product. The student, who is the person who stays with the learning process longest, should learn to become the “co-manager” of his or her education. This means, according to the tenets of TQM, that the student should be involved, consciously and with skill, in the continuous improvement of the processes which create the product. (p. 1)

Tribus asserts that there are four components to quality education. These are:

1. Knowledge, which enables us to understand what we learn in relation to what we already know.
2. Know-how, which enables us to do..... enables us to put knowledge to work.
3. Wisdom, is the ability to distinguish what is important from what is not.
4. Character, as Stephen Covey has said, is a combination of knowledge, know-how and wisdom, coupled with motivation. (p. 3)

The creation of quality education consists of improving both teaching and learning processes. Teaching processes include curriculum design, learning interventions, coaching and mentoring. The instructor plays a major, but not exclusive, leadership role in designing and implementing these processes. Learning processes are more idiosyncratic. They will vary with the student’s preferred learning style, experiences, and learning objectives. Since the student has the greatest stake in the quality of his or her learning, it follows that the student needs to develop personal insight, self-leadership and lifelong learning skills to sustain the learning past an individual course.

The translation of Sink’s model to the educational process raises questions and issues. The student is both a customer of the process as well as a “front-line worker” in the
process. The instructor is a supplier of the process as well as a “front-line worker” in the process. Both student and instructor are co-managers of the process. This challenges the traditional relationship between student and teacher, and raises questions about the evaluation of learning. There are two types of learning evaluation, formative and summative. The purpose of formative evaluation is to assess the quality of the learning processes. This would include the learning interventions facilitated by the teacher, the learning activities facilitated by the students, and the support systems to support the teacher and the students. The purpose of formative evaluation is to improve the quality of the learning processes through diagnosis and revision. Testing is a process for formative evaluation. It is a Q3 activity because its purpose is to improve the quality of the process. The purpose of summative evaluation is to assess the quality of the learning outcomes. Has the student learned? Has the student demonstrated competency or mastery? The mastery of learning outcomes can be demonstrated in a variety of ways depending on the nature of the outcome and the needs of the customer (student, employer). Summative evaluation is a Q4 activity because its purpose is to certify that learning has been achieved. As education improves, the quality of the learning processes and the quality of the learning outcomes will improve. Sink (1989) says that Deming’s approach to quality management is to shift the emphasis from quality checkpoints Q2 and Q4 to an emphasis on quality checkpoints Q1, Q3, and Q5. For education, a Q2 and Q4 orientation to quality management is a screening process. Q2 management keeps students out of the classroom who are not likely to succeed. Q4 management fails students who do not succeed.
Further, the range of grades that are acceptable, and that include those that are not yet high quality, is tacit approval of learning mediocrity by the education system ... close enough for government work.

Tribus (1993) raises even more fundamental questions and criticisms about what should constitute learning. He argues that the focus of the traditional educational system is on knowledge alone. There is virtually no focus on know-how, wisdom, and character. Tribus asserts:

> It appears that in higher education, attention is given only to the first of the four categories, with the least two not even given lip service. (p. 3)

> Professors often believe, as I once did, that at the level of the university, their sole duty is to develop knowledge and pass it on to the next generation. The development of character is none of their business. ..... Only the football coach seems to care about the development of character. (p. 3)

> The typical Professor in the University would consider it beneath a Professor’s calling to actually teach people to apply their knowledge in a practical way. (p. 3)

Tribus’ observations indirectly challenge the evaluation of learning. How can know-how, wisdom and character be evaluated? How relevant are traditional grades as educators take up Tribus’ challenge. Philosophically, Tribus has support from the proponents of total quality management. Defects or mediocrity are never acceptable.

This review of total quality management is not exhaustive, but is developed to build a framework to explore the congruence with adult learning theory. In the next section adult learning theory and principles will be reviewed as congruent with the principles of total quality management. The meshing of total quality management and adult learning theory
will provide the conceptual framework for designing, implementing and evaluating learning in a higher education environment.

3. Adult Learning Theory and Principles

Darkenwald and Merriam (1982) define Adult Education as...

...a process whereby persons whose major social roles are characteristic of adult status undertake systematic and sustained learning activities for the purpose of bringing about changes in knowledge, attitudes, values or skills. (p. 9)

Adult education distinguishes itself from traditional education more from its epistemological assumptions, rather than from the age of the learners. To understand the practice of adult education requires an understanding of its philosophy of learning and human development. The two competing psychological theories that have had a major impact on learning in this century are behaviorism and humanism.

To a behaviorist the goal of learning is to change behavior. Summarizing the basic assumptions of behaviorism, Merriam and Caffarella (1991) state:

First, observable behavior rather than internal thought processes is the focus of study; in particular, learning is manifested by a change of behavior. Second, the environment shapes one’s behavior; what one learns is determined by the elements in the environment, not by the individual learner. And third, the principles of contiguity (how close in time two events must be for a bond to be formed) and reinforcement (any means of increasing the likelihood that an event will be repeated) are central to explaining the learning process. (p. 126)

Thorndike’s Stimulus-Response theory of learning and Skinner’s operant conditioning are the cornerstones of traditional education. The desired response by the learner is strengthened by the connection it has to the stimulus created by the teacher.
Reinforcement is central to understanding operant conditioning. Citing Gippen and Peters Merriam and Caffarella said:

In essence, his (Skinner’s) work indicates that since all behavior is learned, behavior can be determined by arranging the contingencies of reinforcement in the learner’s immediate environment. (p. 127)

Behaviorism is linear in its view of causality, is morally neutral and value-free. It has been the dominant learning theory shaping our current educational system.

Humanism offers a contrasting theory of learning. Merriam and Caffarella state that:

... humanist theories consider learning from the perspective of the human potential for growth.” and

... humanists refuse to accept the notion that behavior is predetermined by either the environment or one’s subconscious. Rather human beings can control their own destiny; people are inherently good and will strive for a better world; people are free to act, and behavior is the consequence of human choice; people possess unlimited potential for growth and development. (p. 132)

Maslow and Rogers are two of the leading humanist theorists. Maslow (1970) proposed a theory of motivation based on a hierarchy of needs. As the lower order needs (basic survival and security needs) are satisfied, higher order needs (belonging, love, self-esteem and finally, self-actualization) have greater potency. Maslow viewed the motivation to learn as inherently intrinsic and that learning is a form of self-actualization. Rogers’ (1980) client-centered therapy is the conceptual foundation for student-centered learning. He viewed the primary purpose of education as the development of fully functioning human beings. He said that significant learning was characterized by personal involvement, self-initiation, pervasive, learner evaluation, and the essence of learning is meaning.
Not all the learning methods of the behavioristic orientation are in conflict with humanism. For example, the self pacing of programmed instruction provides the learner a measure of autonomy in the timing of learning, something that is consistent with the humanistic orientation. Both behaviorist and humanist would reject the notion of punishment as an appropriate feedback for mistakes. Where these two major theories diverge is in their orientation towards intrinsic motivation. The behaviorist does not recognize the existence of intrinsic motivation. The humanist would accept that extrinsic factors are environmental influences on human behavior, but that individuals make the ultimate choices for their behavior. A humanist would also challenge the appropriateness of extrinsic motivation to learning, something that they believe is fundamentally intrinsic.

Behaviorism and humanism are not the only theories that are relevant to learning. But it is the fundamental dichotomy between the extrinsic orientation of behaviorism and the intrinsic orientation of humanism that separates adult learning from traditional education practices. It is intrinsic orientation of adult learning theory which provides the common ground with total quality management.

In his famous book, *The Modern Practice of Adult Education*, Knowles (1970) clearly linked adult learning with the philosophical orientation of Maslow, Rogers, and other humanists. Historically, Knowles showed the congruence of early adult education leaders, (Houle, Lindeman, Overstreet), to humanism and third force psychology. Knowles (1980) also showed the linkage between adult education and the ideas of leaders in management
and organizational development, (Argyris, Benis, Likert, McGregor, Schein) who were challenging the management practices of American industry during the 1960s and 70s.

The first edition of *The Modern Practice of Adult Education*, Knowles (1970) used the word, *andragogy* in the book's sub-title, and literally brought it into the dictionary. He defined andragogy as the art and science of helping adults learn. Knowles first used the word in 1968 and credited European adult educators with its first use. He said that andragogy is based on the Greek word “aner” with the stem “andr” meaning “man, not boy” or adult. He asserted that the word pedagogy was derived from the Greek words “paid” (meaning child) and “agogus” (meaning leading). In his early work Knowles used the term andragogy to distinguish adult education from the education of children. In later works, based on feedback from elementary and secondary school teachers (who had told him that they were experimenting with the concepts of andragogy), Knowles (1980) said: “... that andragogy is simply another model of assumptions about learners to be used alongside the pedagogical set of assumptions” (p. 43). Although he described the contrasting assumptions of pedagogy, Knowles did not overtly show pedagogy's philosophical ties to behaviorism.

He contrasted four basic assumptions of pedagogy and andragogy as:

1. The learner is dependent upon the teacher to determine what is to be learned, and how, and when (Pedagogy).

As individuals mature, their self-concept moves from one of being a dependent personality toward being a self-directed human being (Andragogy).
2. The learner's experience is at best a starting point for learning (Pedagogy).

   As individuals mature, they accumulate a growing reservoir of experience that
   becomes an increasingly rich resource for learning (Andragogy).

3. The school (society) is best able to decide when a learner is ready and most
   people are ready to learn the same things at the same age (Pedagogy).

   As individuals mature, their readiness to learn becomes oriented increasingly to
   the developmental tasks of their societal roles (Andragogy).

4. The orientation of learning is to postponed application and subject-
   centeredness (Pedagogy).

   As individuals mature, their time perspective changes from one of postponed to
   immediacy of application, therefore, moving from a subject-centeredness to
   one of performance-centeredness (Andragogy).

As an adult education practitioner, Knowles described seven conditions of learning
that are superior for the application of andragogy. These are:

1. **The learners feel a need to learn.** The teacher does this by exposing the learner
   to new possibilities and by helping them clarify their own aspirations, diagnose personal
   gaps and identify life problems they experience because of these gaps.
2. The learning environment is characterized by physical comfort, mutual trust and respect, mutual helpfulness, freedom of expression, and acceptance of differences. The teacher builds both a positive physical and psychological environment. The teacher accepts the learners, builds cooperative relationships, expresses personal feelings, and is a co-learner.

3. The learners experience the goals of a learning experience to be their goals. Learning goals are set collaboratively in which the needs of the learner, teacher, institution, and society are taken into account.

4. The learners accept a share of the responsibility for planning and operating a learning experience, and therefore have a feeling of commitment toward it. The teacher shares options and involves the learners in deciding among the options jointly.

5. The learners participate actively in the learning process. The teacher helps the learners to organize themselves and to share responsibility in the process of mutual inquiry.

6. The learning process is related to and makes use of the experience of the learners. The teacher helps the learners use their own experiences and integrate and apply learnings to these experiences.
7. **The learners have a sense of progress towards their goals.** The teacher involves the learners in developing evaluation criteria and helps them to evaluate their own performance.

In the next section, the concept of self-directed learning will be explored further. If self-directed learning is important, how is it learned, what is the relationship between student and teacher, and how does the educational institution support it?

4. **Self-Directed Learning**

Knowles (1980) asserted that:

> The primary and immediate mission of every adult educator is to help individuals satisfy their needs and achieve their goals. (p. 27)

> One mission of the adult educator, then, can be stated positively as helping individuals to develop the attitude that learning is a lifelong process and to acquire the skills of self-directed learning. (p. 28)

Adult educators’ identification with Knowles statement is grounded in the humanistic view of the individual. Self-directed learning was typically defined by early adult educators as a process in which the individual exercises a high level of control in determining learning objectives, designing learning activities, locating resources and evaluating learning. The picture of the self-directed learner was a solitary person, pursuing his or her personal goal, independent of outside control. When an educational institution was involved in self-directed learning, it was often in a non-credit offering or in a non-traditional degree setting. The role of the teacher (helper, facilitator) was often that of a resource person --- helpful, but somewhat passive in the learning process.
In the late 1960s Tough (1971) conducted major research on self-directed learning, published in the *Adult’s Learning Project*. Knowles (1980) paraphrased Tough’s research question as: “How do adults learn naturally -- when they are not being taught.” (p. 42)

The nature of the research question and its methodology was to essentially exclude learning within an institutional setting. Tough’s findings included: (1) a high percentage of people (more than 75%) were engaged in at least one (and an average of eight) learning project per year; (2) people prefer informal over formal means; (3) people are motivated by a variety of reasons, but went through their learning project in a similar set of steps; (4) somewhere in their process, people would seek out help from others, usually fellow amateurs. Tough reported that people avoided trained “teachers” as helpers because they would interfere with the learners’ natural process. Other studies extended the research to various populations and generally followed Tough’s research methodology. The findings supported and elaborated on Tough’s conclusions.

Brookfield (1988) also addressed the issue of self-direction within a formal educational setting. He identified five major themes: (1) the centrality of the learning contract; (2) the supportive and coaching role required of the learning facilitator; (3) the importance of peer learning groups to successful self-directed learners; (4) the time commitment required by faculty; and (5) the frustration, ambiguity, and resentment learners may initially experience, when asked to take greater self-direction within the formal educational environment. Brookfield discussed a number of institutional barriers to and conflicts with the facilitation of adult learning. These included untrained faculty, varying student
readiness, faculty time, institutional structural factors, and the tension between institutional mandates and individual control.

In spite of his criticism of the early research, Brookfield (1988) clearly came down on the side of applying and fostering self-directed learning within educational institutions when he wrote:

The aim of facilitation is the nurturing of self-directed empowered adults. Such adults will see themselves as proactive, initiating individuals engaged in a continuous re-creation of their personal relationships, work worlds, and social circumstances rather than reactive individuals, buffeted by uncontrollable forces of circumstance. (p. 11)

Brookfield views the role of the facilitator as an active collaborator with the learner, asserting that to view a facilitator as primarily a resource person is limiting.

To act as a resource person to adults who are unaware of belief systems, bodies of knowledge, or behavioral possibilities other than those that they have uncritically assimilated since childhood is to condemn such adults to remaining within existing paradigms of thought and action. (p. 124)

Brookfield describes a creative tension that will exist in a healthy learner-facilitator relationship. He says that adult education is a collaborative relationship between learner and student in which objectives, methods, and evaluation should be negotiated. In describing the relationship between the learning facilitator and the educational institution, he warns of a tension that may not be so healthy.

Facilitators choosing to work within existing educational institutions that have as their rationale the certification and accreditation of learners according to institutionally set standards must expect to encounter ambiguity, contradiction, and compromise in their efforts to promote self-directed learning. (p. 88)
Brookfield spoke with the words of one who had done battle with the institution — a battle to create a quality learning experience for students. He showed scars of his battle with:

Attempts to help adult learners make their lives more personally meaningful tell a tale of pleasure and satisfaction. The expressions of genuine appreciation offered by adults for whom my efforts produced some kind of joy in learning more than compensated for the institutional neglect. (p. 89)

Brookfield’s plea was for high quality learning institutions - organizations that promote the joy and pleasure of learning.

5. Learner-Control and the Role of the Facilitator

Reviewing the recent research on self-directed learning Kerka (1994) identified three myths: (1) the myth that adults are naturally self-directed; (2) the myth that self-direction is an all-or-nothing concept; and (3) the myth that self-direction means learning in isolation. She indicated that preference for self-direction may be a matter of degree and influenced by “... learning style, exposure to self-direction, familiarity with subject matter, expectations of schooling and learning, motivation, length of time away from formal schooling, social and political context” (p. 1). Kerka stated that self-direction can be thought of as a continuum over multiple learning decisions: learning goals, type of participation, methods, and assessment. The social nature of much learning challenges the myth of a self-directed learner as an “intellectual Robinson Crusoe”. Citing Caffarella (1993) Kerka states: “The assumption that self-direction, individuation, and autonomy are marks of adulthood is also being challenged by research on gender and cultural
differences, from which is emerging support for connectedness, interdependence, and relationships as equally valid ways of thinking and learning” (p. 2).

Candy (1991) argues that the term self-direction refers to four separate, but related, phenomena: personal autonomy (internal attribute), self-management (ability to conduct one’s own education), learner control (organizing instruction in formal settings), and autodidaxy (noninstitutional pursuit of learning opportunities).

In reviewing the literature on personal autonomy, Candy states that “... in order to be regarded as autonomous, a person needs to be free of internal and external constraints and to heave a coherent and robust set of personal values and beliefs ...”. (p. 125). He goes on to say that: “A person may be regarded as autonomous to the extent that he or she: conceives of goals and plans, exercises freedom of choice, uses the capacity for rational reflection, has the will power to follow through, exercises self-constraint and self-discipline, and views himself or herself as autonomous” (p. 125).

What is the role of education in the development of self-direction in learners? Is self-direction an essential characteristic of a high quality education? How should the principles and practices be used within an educational setting? What is the role of the teacher and particularly what is the nature of his or her relationship with the learners? And finally, the paradox of personal autonomy: What if the learner chooses not to be autonomous in his or her education? In spite of the self-criticism of adult educators towards the unquestioned acceptance of self-direction, it appears that their fundamental support has really not wavered. The research has not so much challenged the basic tenet as it has illuminated
our understanding and provided insights into its application. It is the creative tension among individual learner needs, the facilitator’s needs and mission, legitimate stakeholders, and society, that self-direction within educational institutions offers challenge and opportunity.

Candy (1991) uses the words ‘helper’ and ‘assistant’ to describe the person who guides self-directed learners in their pursuit of personal high quality learning. The success of the relationship, he suggests, depends on the extent of assistance and on the personal relationship between student and teacher. Characteristics of a successful helper includes negotiates clear expectations, responds in flexible manner to the learner’s process and content needs, listens and coaches, supports self-reflection, encourages learner ownership, initiates and challenges as appropriate, shares insights and experiences, and provides balanced feedback while helping the learner to be self-correcting. Candy indicates that the assistance provided by the facilitator is not a purely technical or functional one. It is a personal one marked by warmth, empathy, and authenticity. Candy warns, however, that if learning is defined solely as the acquisition of information there is little evidence to support its superiority over traditional teaching methods. But he also says that prolonged and consistent exposure to learner-control shows gains in creativity, critical thinking, enhanced understanding and retention by students. Learner-control is a fundamental shift in the power relationship between student and teacher. This fundamental shift to a genuine learner-control learning environment will challenge the beliefs and practices of student and teacher alike. The learner-control area that will be explored in the next
section is the area of testing and grading, for it is in this area that the purpose of education and the power relationship between student and teacher will be challenged.

6. Grading and Testing - Quality Learning and Student Empowerment

This section deals with a fundamental issue of learning quality, grading and testing, beginning with a review of the criticism of traditional grading. This criticism suggests that educators need to reflect on the purpose of grading and testing and create a system that supports a high quality learning environment. Alternative approaches to grading and testing methods will be reviewed to address the criticisms raised.

In a major literature review of the relationship between grades and adult accomplishment, Stice (1979) concluded that there was virtually no relationship between grade point average and adult performance. Stice cited the extensive review that Hoyt made (46 research studies) that showed that college grades were unrelated to success in scientific research, engineering, teaching, business and medicine. There were three studies that showed mild correlations between Grade Point Hour Average and salary. This relationship, however, disappeared after four or five years.

Stice, discussing the implications, suggested that the impact of higher education grading was not benign. He suggested that teachers are prone to “... consider students with high grades as bright, serious motivated, ambitious, promising, and even moral, while the average student may be seen as dull pedestrian, lazy, unimaginative and lacking in potential” (p. 392). He goes on to say, “A detached look at one’s former students and
classmates probably will show that these comparisons are invidious and inaccurate” (p. 392). Stice suggests that a big part of the problem is that education is not providing students opportunities to learn things that really count: problem solving, interpersonal relations, creativity and common sense. He states: “By far our greatest commitment is to teaching content, and we reward those students who do well with good grades, grades which do not seem to mean very much in their later careers” (p.392). Miller (1966) in connection with a study of grading practices, makes a strong indictment of the impact of grading when he wrote:

   It seems reasonably clear that the grading system at all levels including the graduate one, tends to reward the conforming plodder and to penalize the imaginative student who is likely to make a significant contribution in nearly any field. (citation by Stice p. 392)

Stice suggests: “If we want to have a more positive effect on our students’ lives, perhaps we should make some changes in what we try to teach them” (p. 392).

   Based on a survey of practicing engineers, Maul (1994) reported that engineers believed that they were not well prepared for actual job demands. Courses were not integrated and did not provide adequate application and practice. Tribus (1993) said that the objective of every school or university is to provide opportunities for students to develop knowledge, know-how, wisdom and character. The traditional grading system is limited, at best, to knowledge. Tribus and others cite the United States Department of Labor’s report (1992), *Secretary’s Commission on Achieving Necessary Skills* (SCANS), as an example of the types of competencies that education needs to be providing students. Workplace know-how identified by the SCANS report includes the use of resources,
interpersonal skills, information, systems thinking and technology. Foundational competencies include basic skills, thinking skills and personal qualities. In an addendum to the original SCANS report the Department of Labor has identified the following changes as desirable for K-12 education:

- Teaching should be offered "in context", that is, students should learn content while solving realistic problems. "Learning in order to know" should not be separated from "learning in order to do."

- Improving the match between what work requires and what students are taught requires changing how instruction is delivered and how students learn.

- High performance requires a new system of school administration and assessment.

- The entire community must be involved.

These researchers and practitioners provide insight into a fundamental principle: For grading to be relevant, it must deal with useful competencies. The lack of relevance of what is being graded must be addressed by educators. Meyer (1992) uses the term "authentic assessment" as performance assessed in a context more like that encountered in real life. For the student the process is absent of contrived test taking. The challenge for the teacher is that he or she needs to know what authentic performance is and how to help students experience it in their learning.

Kohn (1993) states that the belief system behind traditional grading is that the purpose of grading is to provide feedback, motivate performance and sort students for admissions and placements. He challenges the belief that grading motivates performance, even -- and particularly -- if a grade is perceived as a reward for performance. Kohn makes three
major points to support his opposition to grading. **Rewards for learning undermine its intrinsic motivation.** Rewards dilute the pure joy that comes from success itself. Kohn cites Deci’s (1978) research on attribution theory that says that the presence of an external motivator reduces a person’s perception of the internal motivation that the activity has. Kohn states that rewards are less effective than intrinsic motivation for promoting effective learning. His second opposition to grading is: **Grading by a teacher is inherently controlling, causing many students to rebel or relinquish their personal autonomy.** When the students are later asked to think for themselves, they may actually resist, insisting that they have the right to be told what to do. External control reduces personal responsibility. Kohn’s third opposition is: **Students’ attention to performance, “How’m I Doin’”, reduces their attention to what they are doing.** He states:

> Students who are encouraged to think about what they are doing (and not about the grade), will likely find more meaning in the process involved in learning content, value mastery, and exhibit pride in craftsmanship. .... the focus is not on themselves, their abilities, how their progress will be perceived by others, or issues of success or failure or reward or punishment. (p. 156)

In a performance orientation, the student’s self-concept is on the line, resulting from a lack of absorption in the task.

Kohn summarizes by saying that grades cannot be justified on the grounds that they motivate students, because they actually undermine the sort of motivation that leads to excellence and quality. Using grades to sort students undercuts our efforts to educate. There are better ways to give feedback on performance than by giving grades. Kohn believes that students who are motivated by grades or other rewards typically don’t learn
as well, think as deeply, care as much about what they are doing, or choose to challenge themselves to the same extent as students who are not grade-oriented. Kohn asserts that grades dilute pleasure, encourage cheating, strain the relationship between teacher and student, and reduces the student’s sense of control.

Glasser (1990) asserts that traditional grading is coercive, but that grades that provide information to improve learning can be empowering. Glasser has a fundamental problem with the stimulus-response approach of behaviorism. He states that control theory offers a better explanation of people’s behavior. Control theory asserts that our behavior is not a result of outside stimulus, but from our desire to satisfy internal needs (survival, love, power, fun, freedom). Glasser states:

From birth, our behavior is always our best attempt at the time to do what we believe will best satisfy one or more of our needs. (p. 44)

What we get from the outside is information; how we choose to act on this information is up to us. (p. 41)

Glasser distinguishes between Boss-Teacher and Lead-Teacher (adapted from his terminology Boss-Manager and Lead-Manager). The Lead-Teacher shows the student that grades will be used to empower, it is a source of information, not a punishment. When grades are used to coerce, Glasser says that students will usually refuse to take ownership for their work. To be empowering, grades must be seen as a part of the learning process. He states:

A low grade does not mean failure; it means that the student has not yet learned enough. Until the final grade, all grades are temporary. Any low grade along the way can and should be raised. By demonstrating that they know more than they did, students will get a higher grade to replace a lower grade. (p. 53)
Glasser says that a quality school would not accept low-quality work from any student. He proposes three grades: “B”, “A”, and “A+”. A grade of “C” could be retained temporarily. A “C” means that the student has not yet demonstrated quality learning. The student would continue to work to raise his or her grade. Only “B” and higher would go on a student’s transcript. Glasser suggests a number of ways to help a student improve including re-tests and tutoring from fellow students. The role of the Lead-Teacher is critical in establishing a warm, supportive environment where there is no failure and always a way and an opportunity to improve.

Tribus (1993), like Kohn and Glasser, clearly comes down on the side of internal motivation as the basis for creating high quality learning. Like Kohn, he sees threats for poor performance and special honors for good performance as unhealthy external motivators. He believes that cooperation will be reduced if students compete for grades. Like Glasser he advocates rigorous performance standards. Tribus quotes a phrase coined by the students at Mt. Edgecumbe High School: “If it isn’t perfect it isn’t done.” Like Kohn and Glasser he believes that tests should be used to gather information for learning. He states that: “The only legitimate purpose of an examination in the classroom is to help the teacher and learner to decide what to do next” (p. 7). Suggesting the use of tools like Quality Characteristics Evolution Diagrams and Quality Function Deployment Charts as ways for teacher and learner to collect and diagnose learning information for the purpose of improvement. Tribus summarizes his view of high quality learning by saying:
Internal motivators are called into play when a learner understands what it means to do something very well, has had a hand in setting the rules whereby an excellent job is to be recognized, knows that there is someone who shares the joy of knowing the job was well done and is taught to self-assess the work as it is ongoing. The key is not just to make the student responsible, it is to make them response-able. (p. 17)

The following lessons can be taken from these authors and practitioners:

A. **Grades need to reflect meaningful learning outcomes.** To be meaningful the content needs to relate to a practical context, be integrated with other content areas, and support the develop of the whole person.

B. **High quality standards should be the norm for learning for every student.** The learning process and learning environment must support the achievement of these standards.

C. **Grades and tests should be used as an information source that drives continuous improvement.** Their focus should be on formative evaluation to improve learning.

D. **The inducement and satisfaction of intrinsic motivation should replace the practice of external inducements.** The presence of internal motivation encourages high quality learning and joy in learning. It also supports the development of personal autonomy and responsibility.

E. **Students and teachers are co-managers of the learning process.** The system of grading recommended by these authors suggests a fundamental change in this relationship. Decisions traditionally seen as the prerogative of the teacher, are now seen as a joint
process. The teacher is a mentor and a coach. The relationship is both professional and personal.

7. High Quality Learning - Theory to Practice

One of the recurring themes emerging from the literature is that educational institutions need to promote learning that goes beyond information and algorithms. High quality learning stimulates in-depth understanding of the subject and self. This section will open with a number of definitions of high quality learning by many of the authors that have been cited. After these definitions I will review a number of assessment tools that have been used to support high quality learning. Finally, I will review exemplary practices of educational institutions in the application of high quality learning.

A. Definitions of High Quality Learning

The term deep-level learning has been used by adult educators to address a fundamental purpose of education. Candy (1991) wrote:

Deep-level learning involves an attempt to delve beneath the words or symbols to the underlying ideas or "the things signified". It requires a critical and analytical disposition, a deliberate search for the meaning of the subject, and an attempt to identify the relationship between ideas already held and those newly encountered. (p. 291)

Rogers (1983) in writing about whole-person learning said:

Significant learning combines the logical and the intuitive, the intellect and the feelings, the concept and the experience, the idea and the meaning. When we learn in that way, we are whole; we will use all our masculine and feminine capacities. (p. 37)

Biggs (1986) in writing about performance said:

To perform well, one needs to be aware not only of knowledge and algorithms required by the task, but of one's own motives and resources, the contextual
constraints, and to plan strategically on that knowledge ... Learning is no exception. (p. 143)

Tribus (1993) described four levels of competence of quality education:

Knowledge enables us to understand. Know-How enables us to do something. Wisdom enables us to set priorities. Character defines how we behave towards ourselves and others. An education that does not pay heed to all four categories is not a quality education. (p. 12)

Tribus continues with his definition of quality in education, by talking about the affect on the student and in the quality of the process used to achieve the four levels of competence.

He says that:

You know that you are providing quality in education when you find your students working diligently, and with enjoyment, in independent study and discussing what they have learned, in an animated way, eager to engage you in debate or to show you what they have discovered for themselves. This is the kind of joy I have in mind. It is based on doing a quality job because a quality job feels good. (p. 4)

A consistent value and a basic tenet of adult learning is the development of self-directed lifelong learners. This belief is grounded in democratic values, self-knowledge and personal autonomy. Quality learning postulates that most students are valuable, educable, want to learn and are willing to take responsibility for their learning. Tribus (1993) asserts, “The basic philosophy of quality management is that schools can and should make every learner an autonomous team player and that students have the internal motivation to so become” (p. 9). High quality learning requires that students learn more than content.

Students must also increase their own self-knowledge. Finally, high quality learning includes a healthy relationship between student and teacher and between student and student. As Covey (1989) asserts, when people grow to claim their independence
(personal autonomy), they can grow to claim and nurture their interdependence with others.

B. Tools for Assessment

One criticism that Deming (1986) made of performance measurement and goal setting in the business environment was that the wrong things were measured, things that really don’t relate to quality. He asserted:

One of the main effects of evaluation of performance is nourishment of short-term thinking and short-time performance. A man (sic) must have something to show. His superior is forced into numerics. It is easy to count. Counts relieve management of the necessity to contrive a measure with meaning. ... people that are measured by counting are deprived of pride in workmanship. (p. 105)

Knowles (1980) advocated the use of learning contracts as a way to help learners structure their own learning. The learning contract consists of five major sections:

1) Learning Objectives; 2) Learning Resources and Strategies; 3) Target Date for Completion; 4) Evidence of Accomplishment; and 5) Criteria and Means for Validating Evidences. The first three sections are guides to help the learner create his or her learning plan. The last two sections are used for assessment. The learning contract can be used to develop an independent study plan, or to individualize a course. As an instructor in an institutional setting, Knowles would invite his students to use the learning contract to develop their learning plan and to contract for a specific grade, typically an A or a B. Once the learner and the teacher (facilitator) agreed to the contract and the criteria for judgment, the learner was in control of his or her learning and grade. Knowles acknowledges that the evaluation part of the process was the most difficult. The learner
often feels anxious about accepting responsibility for their learning. Knowles emphasizes
the need to provide a strong support for the students. Although the learning contract
provides an opportunity for maximum learner control, it provides a format for evaluation,
rather than a method.

Portfolios provide a more specific framework for assessing the quality of learning.

Paulson, Paulson, and Meyer (1991) define portfolio:

A portfolio is a purposeful collection of student work that exhibits the student’s
efforts, progress and achievements in one or more areas. The collection must
include student participation in selecting contents, the criteria for selection, the
criteria for judging merit, and evidence of student self-reflection. (p. 60)

The authors state that portfolios can be a powerful method for student control of their
learning. They see the portfolios as an intersection of instruction and assessment. The
authors describe a number of portfolio guidelines. They provide an opportunity to learn
about learning through self-reflection, offer a way for students to own and value their
work, and convey the student’s activities, thus giving insights into the learning process.
They show progress towards goals and examples of performance. The authors conclude
with:

... portfolio assessment offers the opportunity to observe students in a broader
context: taking risks, developing creative solutions, and learning to make
judgments on their own. ... A portfolio, then, is a portfolio when it provides a
complex and comprehensive view of student performance in context. (p. 63)

A portfolio, the authors assert, encourages independent self-directed learning.

Although portfolios have a strong focus on the needs and interests of the individual
learner, Paulson and Paulson (1991 discuss the role of stakeholders in creating and
evaluating a portfolio. The interests of secondary stakeholders must be negotiated. Different stakeholders can review the contents of the portfolio from a personal set of intents and standards. In a college course in engineering, the secondary stakeholders may be potential employers. The instructor can give guidance to the students in identifying and responding to the needs of these stakeholders in their portfolios.

Stice (1979) reviews two models of learning, Personalized System of Instruction (PSI) and Mastery Learning, that take a similar approach in helping students achieve mastery in a subject. Using PSI, also known as the Keller Plan, the teacher makes an analysis of what is to be learned in a course and develops terminal and intermediate objectives. These objectives are the basis for developing units of instruction. Each unit includes "... a reading assignment, study questions, collateral references, study problems, and any necessary introductory or explanatory material." The student uses this material for self-instruction and completes it at his or her own pace. When the student feels the unit has been mastered, a proctor gives a readiness test to see if the student is ready to move on to the next unit. The proctor "grades" the test with the student present. The proctor may ask follow up questions for minor errors. A grade of 100 is required to move on to the next unit. If the student has not demonstrated mastery, he or she is given feedback and asked to go back and restudy the material missed. The student continues to retake different readiness tests until mastery. The student is never penalized. A student's grade may be determined by the number of units completed or by a score on a final examination.
Mastery learning is similar to PSI in that the teacher develops learning units and uses testing for diagnoses rather than grades. In mastery learning the pacing is established by the teacher in a more conventional instructional style. After a unit is completed a diagnostic-progress test is given, which may be scored by the teacher or self-scored by the student. As in PSI students are given alternate learning sources to help them master the content. The course grade is based on the score on a final examination.

Koen, Himmelblau, and Jensen (1985) discusses the application of the Keller Plan (PSI) within the engineering curricula at a major university. The authors report high student achievement and satisfaction with the Keller plan. The Keller plan requires an upfront time commitment by instructors to prepare course materials and multiple tests. The authors reported that there are some problems with student procrastination, affecting the completion of the courses within the university’s schedule. Consequently, the university has moved away from a completely self-paced implementation and has lowered the “unit perfection” requirement. Hereford (1979) in a ‘meta-analytic’ study reported on the use of the Keller Plan within a conventional academic environment. The study generally supported the assertion of higher student performance and satisfaction. Problems of student procrastination, proctor time, and institutional requirements have resulted in the addition of instructional structure and a reduction of mastery performance standards.

PSI and Mastery Learning contain some elements to be considered for high quality learning. Tests are used for diagnostic purposes. A level of student autonomy is provided
for pacing. There is an intention to have each student perform at a mastery level and to provide the support so that can occur. But PSI and Mastery Learning fall short in other areas of high quality learning. It is a teacher-centered learning strategy. The student has no input into the course structure or assessment method. The issue of authenticity in the assessment method is problematic. In improving the students’ performance on mastery of course content, it does not challenge that the validity of the mastery criteria for performance.

In describing his version of competency based education, Tribus (1993) states:

At the beginning of the semester, the teachers should discuss with the entire class the list of competencies, and the level of mastery expected, for each competence. The students should participate in the discussion of each competency, how they, themselves, will know their level of competency, how they will demonstrate it, how the teacher will assess it and what the teacher will do to help them achieve it. (p. 7)

Tribus advocates a process in which the teacher and the students create the evaluation criteria together. He asserts that such a process results in both high standards and high commitment. Tribus describes a number of tools that can be used for both planning and assessment. He proposes the use of the Nominal Group Technique to assist students to explore their own learning objectives and priorities. This begins a process of negotiation and collaboration between the students and the teacher. Deployment Flow Charting is used as a means of charting progress and displaying “customer supplier” relationships in the learning process. The Quality Characteristics Evolution Diagram is a tree diagram that shows the detail and relationships of competencies. In addition to identifying competencies within a course, it can and should be used to show the teamwork required
among courses. **Quality Function Deployment** can be used to see how the institution is deploying its resources against professed objectives. It is also a tool for student evaluation of how well courses and teaching strategies meet their learning objectives. Tribus approach to assessment goes well beyond the assessment of student performance. In the spirit of total quality management he believes that assessment tools should include the evaluation of the teaching and learning processes, with the focus on continuous quality improvement.

All of the assessment tools reviewed here offer alternative views of evaluation to be considered for application within institutional learning. Effective evaluation is part of learning, not a separate entity. In the next section, examples of exemplary practices in high quality learning will be reviewed to show how some of these tools have been used.

**C. Exemplary Practices**

Following are examples of exemplary practices in the application of high quality learning in a formal educational setting.

Kalonji (1992) reviews a radical change in the implementation of an introductory materials science course taught at the University of Washington. She said that the course “... changed from lecture, problem sets, and exams to a heavy emphasis on student-initiated projects, which are done individually and in large and small groups” (p. 28). The grading scheme was based primarily on journals and portfolios. Students were asked to identify problems they would solve if they could. They were also asked to develop solution strategies. The portfolio-based assessment matched the project orientation of the
course. The course has nine modules and the students must choose three projects (one individual, one in small group, one in large group). One project is assigned to the students. The students can choose the remaining two. Kalonji indicated:

The journals allowed students a chance to reflect on their own learning styles, and gave instructors an ongoing chance to see how our course reforms were working. It was a way to involve students in transforming the educational process. (p. 29)

The course professors also keep a journal which were open to the students to read. TA’s and undergraduate teaching interns were used as module consultants. The grading system was not explicitly described, but Kalonji said that grading strategies were decided as a group. She indicated that the one of the challenges of the course: “... is balancing the tremendous amount of freedom students have to pursue their own ideas with enough structure so they feel secure” (p. 29). Preliminary feedback showed students responded positively to this approach and they particularly liked the journals. Informal feedback from downstream professors indicated students who had taken this course are more likely to take the lead and get other students involved in future courses. Kalonji concluded the interview with: “The last time I taught this course was by far the most fun I’ve had as a professor” (p. 29).

Green and Winn (1992) report on a redesign of an engineering thermodynamics course based on the principles of total quality management. Learning improvements were made in four areas: testing methods, cooperative learning, mentoring, and teaching TQM to the students. The testing process moved from a traditional process that the authors call “Learning by Inspection” (exams given strictly to evaluate) to one that gave the students
more than one chance to pass. An 80% grade or better was required to pass an exam. Students were assigned to design project teams of four or five (assignment was made to make the combined GPA’s of all the groups equivalent). Students on a team were given bonus points if everyone on the team passed the exams on the first try, thus encouraging support and cooperation. Volunteer mentors (the stronger students) would help students who needed help. The mentor’s grade would rise in proportion to the struggling student’s improvement. Finally, all students were taught the TQM philosophy and continuous improvement methods. These methods were used by the design team for their projects. They developed process flow charts, defined measures and developed milestone charts. The groups also used tools for self assessment.

Student performance improved significantly based on comparison with previous offerings of the course. Student test performance improvement was greatest with the low to average students, but the better students stayed about the same. Student satisfaction increased significantly in the areas of knowledge, thinking skills and enjoyment. The authors felt the major course improvement was in its shift from a “High Stakes” (based on evaluation and competition) learning system to a “Help” (based on diagnosis and improvement and cooperation) learning system. They are working to improve the system to benefit the high performing students.

Smith (1993) reports on an eleventh grade American Studies (integrating history and literature) course in a Chicago suburban high school. The two teachers who team teach this course begin the first day of class with the question: “What is your idea of a perfect
This question begins a student-negotiated curriculum that is based on a shared belief system of the teachers. The authors statement of beliefs include:

First, we believe students want to learn, their own questions will be at least as in-depth as those that teachers pose to them.

Second, because planning their own learning is outside the experience of most students, we must proceed in a way that stretches their ability while maintaining a comfort level.

Third, the power of the group greatly determines the success or failure of a given course in a given situation. If group norms are positive, the course will succeed.

Next, we need to look beyond our written curriculum to life-based outcomes, shaping the former based on the latter.

Finally, if students are actively engaged in thinking and problem solving, we believe that they will be able to meet the challenge of any state, local, or state-generated test. (p. 36)

The teachers reported a process in which the students developed a class constitution, including group guidelines and norms. Students were asked to keep journals about what they already know about a class topic. They used a K-W-L process as part of their journaling. K-W-L means: what I already Know, what I Want to know, and what I Learned. Throughout the course the students took greater responsibility for the course.

The teachers reported their primary roles in the class were:

We instructed students on writing journals, taking notes, conducting research, and summarizing material. We also brought resource materials, contacted speakers, and arranged time in the library for student research. However, most of our time was spent discussing the students' ideas for class content and structure. (p. 37)

Students worked independently and in small and large groups. They were given wide latitude in choosing the assessment process. The creativity and depth of the assessment
process was a major departure from the traditional teacher driven testing. The teachers reported that the process was not always smooth and it challenged the strength of their belief. They quoted a student who asked early in the semester: “it would be a lot easier if you two just taught and we just obeyed!” (p. 37). Reflecting, they admitted it might be easier, but not better. Quoting one of the students at the end of the first semester: “I particularly enjoyed the interaction between ‘teacher’ and ‘student’. These terms (required criterion-based tests) to me are artificial in that nobody stops learning in their life, and everyone is a teacher of somebody else” (p. 37).

Tribus (1990) writing on a school-wide application of total quality management, states “To my knowledge, this is the only school in the world which is attempting to apply Deming’s ideas to the totality of education” (p. 1). The school is Mt. Edgecumbe High School, Sitka Alaska. David Langford, one of the early pioneer teachers, reported how he experimented with different approaches to overcome student indifference to learning. He begins the semester in dialogue with the students around a number of fundamental questions: “Why are we here?” “What do we want to get out of this course?” What are the barriers to success?” “What does it mean to do this course with quality?” These questions stimulated personal insight that “... caused the students to examine their own objectives and thereby alter their attitudes” (p. 2). The insight and dialogue is a beginning of a process to help the students become co-managers of their education. In a course on entrepreneurship, the students prepared a plan to sell smoked salmon in Japan (They received an order for $140,000). The students take responsibility to ensure their
work is of high quality. Two student mottos are part of the school's culture: "If it isn't perfect, it isn't done." and "No excuses." There are no grades or incompletes. The task is not complete until the work is done. Tribus reports: "The faculty, the students, and the administration, all agree that evaluation is destructive to the learning process" (p. 5). The results from MEHS are impressive even using traditional measures. Figure 2 illustrates a Control-Empowerment spectrum developed by the students. This chart shows their understanding and commitment to lifelong learning.

![Control Spectrum Diagram]

As the movement in education moves towards a more learner-centered approach, students accept and eventually seek empowerment and the personal responsibility that it demands. The details of the Mt. Edgecumbe High School transformation is not the focus of this review. What is the focus is the principles and values behind the process. This
small high school and its pioneers, students, staff, faculty, and administrators give us a model and it touches something deep within us. They are nurturing lifelong learners who will be dissatisfied that the colleges they are entering are not “learning environments”.

8. Conclusions

Adult learning theory and total quality management share principles that are a source of attraction. Adult learning theory and total quality management offer each other complementary principles that are a source of creative tension. It is the blending of the congruencies and the balancing of the tension that provide the greatest opportunity for the application of these two bodies of knowledge within a formal educational setting.

The common ground that is shared between adult learning theory and total quality management is in the common roots of humanism. Human beings are fundamentally good, are able to exercise self-direction, strive to satisfy intrinsic needs, and are capable of self-actualization. Maslow, often called the father of third force psychology, was a leader in the study of psychology from a perspective of emotional health rather than the perspective of emotional sickness. He believed that human beings had a hierarchy of needs from basic survival needs (lower order) to self-actualization needs (higher order). As human beings satisfied their lower order needs their higher order needs became more potent. Maslow believed that the satisfaction of higher order needs was related to personal growth and psychological health. Management theorists like McGregor believed that the fundamental flaw of traditional management (which McGregor called Theory X management) was that
organizational leadership did not recognize the existence or strength of intrinsic motivation in the life of workers. Consequently, the tactics of traditional management was to use extrinsic factors alone to motivate human performance. Traditional education, grounded in behaviorism, treats students in the same way. The focus of traditional education is on teaching and extrinsic factors to motivate human learning. Performance and learning does occur, but often in spite of these extrinsic factors. There are four areas of congruency for adult learning theory that flow from the humanism they share. These are empowerment, process, motivation and quality.

ALT and TQM share an empowerment orientation. Brookfield (1988) said that all learning is voluntary. Individuals exercise their own personal power in the learning process. They make choices and decisions on how new knowledge is to be assimilated into their life. The knower and the knowledge are inseparable. The goal of teaching is to help people to grow and become empowered as human beings. In an organization that embraces total quality management, decisions are made by the person most able to make them. Employees are encouraged to exercise greater self management in their job. Both adult learning theory and total quality management share a common tension in the process of empowerment. In the learning process, there is tension between the need of the learner to exercise self-direction and the need for the teacher to share his/her knowledge and to challenge the beliefs of the learner. Within organizations there is tension between individual self management and the legitimate leadership and guidance that managers exercise.
ALT and TQM share a process orientation. In traditional education (pedagogy) there is a strong content orientation. In ALT content is subordinate to process. Knowledge is not separate from the knower. Content has no meaning outside of context. The goal of learning is to help people learn how they learn, and to enhance understanding, self-knowledge and lifelong learning. In a total quality management organization there is an emphasis on continuous process improvements. Improving fundamental processes has an intimate quality about it. In manufacturing, for example, the machine's rhythm and sound reveal more about the production process than does the attribute measures of the object being produced.

ALT and TQM share the same orientation to motivation. Intrinsic factors have greater potency in the individual's motivation to learn and to perform. The role of the teacher is not to motivate students to learn. The role of the teacher is to provide an environment in which people can motivate themselves through satisfying intrinsic needs. The role of the manager is not to motivate people to perform. The role of the manager is to help satisfy individual needs that are aligned to organizational needs. The use of external rewards can even have a detrimental affect on learning and performance because it draws attention away from the inherent intrinsic satisfaction to learning and working.

ALT and TQM share an orientation to high quality. The term deep-level learning relates to a higher level of understanding that Tribus called wisdom and Deming called profound knowledge. Tribus speaks of learning as a joy and Rogers speaks of learning as the development of a fully functioning life. Total quality management does not accept any
level of defects as acceptable. The goal is to delight the customer in the never ending process of improvement. High quality learning and performance has a spiritual quality about it. High quality enhances life.

There is also a creative tension between ALT and TQM. The source of this tension is adult learning theory’s orientation to the individual and total quality management’s orientation to the organization. These orientations are in tension with one another. The tension provides a basis for finding a healthy balance as we strive to honor the integrity of both.

Although both ALT and TQM value personal autonomy and teamwork, personal autonomy is more central to ALT and teamwork is more central to TQM. There is a natural tension between these values. Does teamwork mean that for the sake of the organization that I subsume my needs? Does consensus decision making cause me to lose something? How is personal autonomy balanced with personal responsibility and commitment? It is the process of balancing these tensions that offer an opportunity to create something better. Bohm (1992) defines dialogue as a stream of meaning that is created when people in a group learn to listen to one another and to themselves in a new way. Through the process of dialogue shared meaning can be discovered -- and something is created. Dialogue is a meta-communication process which Bohm says that can lead to a transformation of consciousness for individuals and groups. Covey also speaks of interdependence as a natural development in which people can attain after they have achieved a level of independence (personal autonomy).
The beliefs that ALT and TQM share are:

- A belief in a humanistic orientation with a view of people as capable and willing to exercise self-direction.
- A belief in empowerment and that people will exercise responsibly in its use.
- A belief that the role of institutions is to create an environment that will support high quality work and learning.
- A belief that assessment should focus on continuous improvement - to fix the problem, not find blame.
- A belief that quality relates to process and not just to product.

Total quality management and adult learning theory are both holistic and long term.

The grading system to support these beliefs include the following criteria:

- Evaluation, both formative and summative, must relate to meaningful outcomes that relate to in-context learning and the development of the whole person.
- Evaluation should first focus on diagnosing learning and teaching methods for the purpose of improving the quality of learning for each student and teacher.
- High Quality standards should be the norm. Mediocrity must be eliminated as acceptable performance.
- Evaluation should recognize the potency of internal motivation and the potential harm of external motivation.
• Evaluation should support the establishment of a respectful and egalitarian relationship between learner and teacher.

For purposes of this study it is hypothesized that a grading and testing system built upon these criteria will:

• result in higher student and teacher performance
• result in higher student and teacher satisfaction
• result in a greater demonstration of self-direction
• result in a more positive relationship between student and teacher.
CHAPTER III

RESEARCH METHODOLOGY

1. Subjects and Research Design

The research study was conducted to determine the effectiveness of a new testing and grading system design on student performance and satisfaction in a course in engineering economics. The subjects were junior and senior engineering students from all engineering departments except Civil and Chemical Engineering (who taught their own engineering economics courses). The instructor has been teaching this course for several years and taught all of the classes in which data were collected for this research analyses. This instructor was also a member of the faculty learning inquiry team that created the new course design.

A quasi-experimental research design was used to compare the impact of the new testing and grading system to that of the testing and grading system that had been used for a number of years. This research design was implemented during the Spring Quarter of 1994. The experimental group (7:30 a.m. section) consisted of 96 students and the control group (3:30 p.m. section) consisted of 98 students. The students selected these sections based on the university bulletin and were unaware of the research design when they enrolled in the course. The students in the experimental section were told that they were participating in a research study during the first day of class.
A cross-sectional pre-experimental research design was used to analyze student performance and satisfaction for four quarters (Autumn 1992, Winter 1993, Autumn 1993, Winter 1994) before the experiment and two quarters (Autumn 1994, Winter, 1995) after the initial experiment. Additional changes were made to the testing and grading system in the autumn quarter, 1994 and continued in the winter quarter, 1995. This pre-experimental design was implemented to explore the impact of the changes in the testing and grading processes on student performance and satisfaction over a period of time.

2. Engineering Economic Analysis Course Description

Engineering economic analysis is an introductory course that deals with the use of analytic techniques in making decisions relevant to engineering design. The content includes topics such as: the time value of money, benefit/cost ratios, rate of return, depreciation modeling, income taxes, inflation, replacement, decisions under uncertainty, and the use of spreadsheets in calculating economic cash flows.

The testing and grading system used in the spring quarter, 1994 and previous quarters consisted of three midterm exams, two spreadsheet exercises and five pop quizzes. A final examination was required for all students who had a grade of B- or less, based on their performance up to the final. A student’s grade would be determined by the percentage he or she would obtain from all exams, quizzes and exercises. This percentage would be converted into a letter grade based on the conversion chart illustrated in Table 1.
TABLE 1  CONVERSION OF TEST SCORE AVERAGES TO GRADES

<table>
<thead>
<tr>
<th>Grade</th>
<th>Score Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Less than 60</td>
</tr>
<tr>
<td>D</td>
<td>60 to 65</td>
</tr>
<tr>
<td>D+</td>
<td>66 to 69</td>
</tr>
<tr>
<td>C-</td>
<td>70 to 72</td>
</tr>
<tr>
<td>C</td>
<td>73 to 75</td>
</tr>
<tr>
<td>C+</td>
<td>76 to 79</td>
</tr>
<tr>
<td>B-</td>
<td>80 to 82</td>
</tr>
<tr>
<td>B</td>
<td>83 to 85</td>
</tr>
<tr>
<td>B+</td>
<td>86 to 89</td>
</tr>
<tr>
<td>A-</td>
<td>90 to 92</td>
</tr>
<tr>
<td>A</td>
<td>93 to 100</td>
</tr>
</tbody>
</table>

In the spring quarter, 1994, the final exam was required for both experimental and control groups because the final exam score was one of the response variables for the research design.

The testing and grading system for the experimental group for the spring quarter, 1994 included the following changes.

- The students were given an option to be re-tested on up to three (out of six) test questions for each midterm examination. This re-test was administered in a group at a pre-set schedule. The re-test was given in the same week that the original tests were returned. The test scores received on the re-test exam questions were substituted for the scores received on the original test and the student’s exam score was recalculated. The re-test questions were different from the exam questions but tested the same concepts.

- A student could elect to do a learning inquiry project using a modified learning contract form. The students were given examples of topics they could choose, but were given wide discretion in choosing a project. To elect this option students could propose a learning project up to the end of the fourth week of
the course. They could obtain feedback from the instructor any time during
the course on their project, but had to complete the project by the last class
period. A completed project was worth 150 points. The student could use this
project assessment as a substitute or replacement for a midterm examination.

The determination of the course grade, converting points to percentages and percentages
to a grade was the same as previous quarters.

The testing and grading system used for the autumn quarter, 1994 consisted of the
following changes:

- Pop quizzes were eliminated.
- The re-test option was maintained for all three midterm examinations.
- The final examination remained required.
- The grades of C-, D+ and D were eliminated.
- A grade of C was achieved when a student received a minimum test score of
  70 % on each midterm (after re-test option) and the final examination, and an
  average test score of 75% for all four examinations. If the necessary
  improvement was not achieved during the re-test students were allowed to
  repair (open-book take-home test)
- A grade of B was received when a student received a minimum test score of
  80% on each midterm (after the re-test option) and the final examination, and
  an average test score of 85% for all four examinations.
A student could improve his or her grade (for example, a C to B or a B to A) up to a full letter completing one or more learning projects. The options are illustrated in Table 2.

**TABLE 2 POINTS FOR LEARNING PROJECTS**

<table>
<thead>
<tr>
<th>Project</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet Exercises (up to 2)</td>
<td>50 points each</td>
</tr>
<tr>
<td>Case Studies (up to 2)</td>
<td>60 points each</td>
</tr>
<tr>
<td>Personal Journal of Course Experience</td>
<td>60 points</td>
</tr>
<tr>
<td>Write-up on Guest Speaker</td>
<td>30 points</td>
</tr>
<tr>
<td>Problem Analysis (up to 5)</td>
<td>10 points each</td>
</tr>
<tr>
<td>Interview with instructor</td>
<td>10 points</td>
</tr>
<tr>
<td>Self Initiated and Negotiated Projects</td>
<td>points negotiable</td>
</tr>
</tbody>
</table>

Spreadsheet exercises, case studies and self-initiated projects could be done in teams or individually. Other projects were individual. All learning projects were evaluated on a go/no-go basis. Students could continue to revise their project and improve it to achieve a "go". A student earning 50 Learning Project points would raise his or her grade by one mark (for example, a B to a B+). A student earning 100 Learning Project points would raise his or her grade by two marks (for example, a B to an A-). A student earning 150 points would raise his or her grade by three marks (for example a B to an A).

The testing and grading system used for the Winter quarter, 1995 was the same as the Autumn quarter, except that the final examination was optional. (This option was proposed by the students during the quarter and accepted by the instructor.)
3. Response Variables

There are two categories of response variables that were used in this research. The first category dealt with student performance. The second category dealt with student satisfaction. In this section the operational definition of each response variable will be reviewed.

A. Student Performance

Students performance is defined as test score achievement on midterms, final examinations and final grades. The tests and the scoring keys were created by the course instructor. For the control groups (autumn 1992, winter 1993, autumn 1993, winter 1994, and the spring 1994 control class) the following data were collected:

- the median and Inter Quartile Range (IQR) of the test scores on the midterm examinations and the final examination.
- the percentage of students for each class achieving a three-test midterm test score average of 85% or higher.
- the percentage of students for each class receiving a course letter grade of B or greater.

For the experimental groups (spring 1994 experimental class, autumn 1994, and winter 1995) the following data were collected:

- the median and Inter Quartile Range (IQR) of the test scores on the midterm examinations and the final examinations.
• the median and the IQR of the test scores on the midterm examinations after
the re-test.
• the percentage of students for each class achieving a three-test midterm test
score average of 85% or higher after the re-test option.
• the percentage of students for each class receiving a letter grade of B or
greater.

B. Student Satisfaction

Student satisfaction is defined by the students’ rating of the course and instructor
using the Student Evaluation of Teaching (S.E.T.) Report. This survey form has twenty
items in which the student may respond: Strongly Agree, Agree, Neutral, Disagree,
Strongly Disagree. Agreement responses indicate student satisfaction with the item.
Overall satisfaction will be measured by the percentage of responses that were satisfactory
(Strongly Agree and Agree). A list of the twenty questions included in the S.E.T. is in the
appendix. Table 3 shows the S.E.T. questions that relate to the students satisfaction with
the testing and grading system.

TABLE 3  S.E.T. QUESTIONS, TESTING
AND GRADING SYSTEM

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17.</td>
<td>Exam questions fairly reflected course content.</td>
</tr>
<tr>
<td>18.</td>
<td>Exams tested well my learning/understanding of course material.</td>
</tr>
<tr>
<td>19.</td>
<td>Exams emphasized important or major topics.</td>
</tr>
<tr>
<td>20.</td>
<td>Exams were graded fairly.</td>
</tr>
</tbody>
</table>
4. **Data Analysis - Quasi-Experimental Design Spring Quarter 1994**

This section describes the data analysis techniques used to test the hypotheses for the experiment conducted in the spring quarter 1994. In addition, data analyses are presented to explore the impact of the re-test option on variation reduction, and conduct exploratory research on the data.

A. **Hypotheses Testing**

**HYPOTHESIS 1** - The experimental class in the spring quarter 1994 will have a higher median and a lower IQR for the final examination test score distribution than will the control class in the spring quarter 1994.

Notched box plots will be used to test the hypothesis for a difference in medians. A box plot is a summary display of a data distribution. At the top of the box is a horizontal line that represents the 75th percentile value of the distribution. At the bottom of the box is a horizontal line that represents the 25th percentile value of the distribution. Within the box a horizontal line is drawn to represent the 50th percentile value (the median). The interquartile range (IQR) is a measure of the distribution’s variation. The IQR = the 75th percentile value - the 25th percentile value. The IQR displays the middle 50 percent of the distribution’s data. To display the tails of the distribution, two additional values are determined. The upper adjacent value is defined to be the largest observation that is less than or equal to the upper quartile plus 1.5 x IQR. The lower adjacent value is defined to be the smallest observation that is less than or equal to the lower quartile minus 1.5 x IQR. These values are displayed on the box plot by extending a vertical lines up and down to
these values from the center of the horizontal lines representing the upper and lower quartiles. A small horizontal line is drawn to show both of these values. Any other data points that may fall outside of the upper adjacent or the lower adjacent value is shown by an individual dots extending away from the upper and lower adjacent value.

Notches can be used to show an approximate confidence interval for the comparison of two or more distributions. The following formula can be used to calculate the confidence interval:

\[ M = +/- 1.57 \times \text{IQR} / \sqrt{n} \quad (1) \]

where \( M \) is the median, IQR is the interquartile range and \( n \) is the number of observations in the distribution. Chambers, Cleveland, Kleiner, and Tukey (1983) state that the notched box plot is an informal yet statistically sound method for comparing the means of two or more distributions. They state: “if the notches for two boxes do not overlap, we can regard it as strong evidence that a similar difference in levels would be seen in other sets of data collected under similar circumstances” (p. 62). They go on to say “if the two data sets are independent and identically distributed random samples from two populations with unknown medians but with a normal distribution shape in the central portion, then the notches provide an approximate 95% test of the null hypothesis that the true medians are equal; if the two notches overlap, then we fail to reject the null hypothesis with (approximate) 95% confidence” (p.62). The authors also indicate that notched box plots can be useful guides for comparing medians even when the data conditions are not completely met.
Although there will be no formal statistical test to compare the changes in variation, the changes in the IQR will be reported and discussed.

HYPOTHESIS 2 - The experimental class in the spring quarter 1994 will have a higher median and a lower IQR after the re-test for each midterm test score distribution than will the control class in the spring quarter 1994.

Notched box plots will be used to test for the difference in the medians.

HYPOTHESIS 3 - The experimental class in the spring quarter 1994 will have a higher median and a lower IQR for the three-midterm average after the re-test than will the control class in the spring quarter 1994.

Notched box plots will be used to test for the difference in the medians.

5. Cross-sectional Data Analysis of Student Performance Autumn 92 to Winter 1995

Box plots are used to show changes in student performance based on design revisions to the testing and grading system implemented in the autumn 1994 and winter 1995 quarters on student test performance. Specific hypotheses are not tested. The purpose of this data analysis is to explore the impact of these changes to gain insight into the effect that the incremental changes in the testing system and the grading system had on student performance. The data from the first four quarters establishes base line of student performance for the old testing and grading system used prior to this research. The data from the spring 1994 quarter shows the impact of the changes in the testing system that was used in the quasi-experimental design and its relationship to the base line performance. The data from the autumn 1994 and winter 1995 shows the impact of the
second set of changes that were made to the grading system on student performance. The significance of this cross-sectional analysis is that we are able to explore the changes in the testing and grading process over time. The three testing and grading conditions are (1) the traditional testing and grading system used prior to this research, (2) the change in the testing system that was used in the experimental design in spring 1994, (3) the changes in the testing and grading system that was implemented in the autumn 1994 and winter 1995 quarters.

Bar charts are shown to show changes in student performance over the three testing and grading systems: the traditional testing and grading system; the re-test option to the testing system implemented with the spring 1994 experimental class; and the testing and grading changes implemented in the autumn 1994 quarter and continued in the winter 1995 quarter.

6. Cross-sectional Data Analysis of Student Satisfaction Autumn 92 to Winter 1995

HYPOTHESIS 4 - The experimental class in the spring quarter 1994 will have higher satisfaction with the testing and grading system as measured by the S.E.T questions 17 through 20 than will the control class in the spring quarter 1994.

HYPOTHESIS 5 - The autumn 1994 class and winter 1995 class will have a higher satisfaction with the testing and grading system as measured by the S.E.T. questions 17 through 20 than will classes that used the old testing and grading system.

An attribute control chart is used to analyze student satisfaction, as measured by SET scores, over eight offerings of the course. These eight offerings are Autumn 1992, Winter

A content analysis was made of students written response to questions on the Industrial and Systems Department Course evaluation form. This was done to explore student’s reaction to the course changes and develop a deeper understanding of the hypotheses tests and data displays.
CHAPTER IV
RESULTS

1. Introduction

The purpose of this chapter is to present the results of the analyses of the data to:

- test the hypotheses for differences in student test performance between the experimental class and the control class for the spring quarter 1994.
- determine the impact of the re-test option on test scores performance and variation in performance within the experimental class for the spring quarter 1994.
- determine the impact of the testing and grading options used in the autumn quarter 1994 and the winter quarter 1995 on test score performance and variation in performance.
- test the hypotheses for differences in student satisfaction between the experimental class and the control class for the spring quarter 1994.
- determine the impact of the testing and grading options used in the autumn quarter 1994 and the winter quarter 1995 on student satisfaction.

Notched box plots were used to test the hypotheses for the differences in student test performance for the spring quarter 1994. Side-by-side stem and leaf diagrams were used to explore the impact of the re-test option for the experimental class for the spring quarter 1994 on student test performance and variation in performance.
Box plots were used to explore the impact of the testing and grading options used in the autumn 1994 and winter 1995 quarters on student test performance and variation in performance.

Bar charts were used to display overall class test performance and grades for four quarters prior to the spring 1994 experiment, the control and experimental classes for the spring quarter 1994, and the autumn 1994 and winter 1995 quarters.

Attribute control charts (p charts) were used to test for the differences in student satisfaction for four quarters prior to the spring 1994 experiment, the control and experimental classes for the spring quarter 1994, and the autumn 1994 and winter 1995 quarters.

Content analyses were performed on students responses to an open ended question on their reaction to testing and grading for two quarters prior to the spring quarter 1994, the control and experimental classes for the spring 1994 and the autumn 1994 quarter.

2. **Student Test Performance - Spring Quarter 1994**

   Table 4 shows the student characteristics for the control class and the experimental class for the spring 1994 experiment. The characteristics include the average grade point hour average for the students prior to the spring quarter, class rank, gender, and nationality. This data supports the assertion of the equivalence between these two classes.
TABLE 4  STUDENT CHARACTERISTICS
SPRING QUARTER 1994

<table>
<thead>
<tr>
<th>CHARACTERISTIC</th>
<th>CONTROL CLASS</th>
<th>EXPERIMENTAL CLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students</td>
<td>98</td>
<td>96</td>
</tr>
<tr>
<td>G.P.A. Average</td>
<td>2.953</td>
<td>2.839</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.497</td>
<td>0.483</td>
</tr>
<tr>
<td>Class Rank</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>36%</td>
<td>46%</td>
</tr>
<tr>
<td>4</td>
<td>60%</td>
<td>52%</td>
</tr>
<tr>
<td>Other</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>86%</td>
<td>82%</td>
</tr>
<tr>
<td>Female</td>
<td>14%</td>
<td>18%</td>
</tr>
<tr>
<td>Nationality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American</td>
<td>88%</td>
<td>91%</td>
</tr>
<tr>
<td>International</td>
<td>12%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Figure 3 shows notched box plot displays for student test performance for the first midterm for the spring quarter 1994. The overlap of notches between the test scores of the control class and experimental class indicates that the null hypothesis in the differences between the medians is not rejected. The absence of the overlap of the notches between the test scores for the experimental class after the re-test and the experimental class before the re-test indicates that the null hypothesis in the differences between the medians is rejected at an approximate Alpha = .05. The Inter Quartile Range (IQR) reduced from .23 for the ‘before re-test’ test scores to .15 for the ‘after re-test’ test scores.

Figure 4 shows notched box plot displays for student test performance for the second midterm for the spring quarter 1994. The overlap of the notches between the test scores of the control class and experimental class indicates that the null hypothesis in the differences between the medians is not rejected. The absence of the overlap of the notches
class before the re-test indicates that the null hypothesis in the differences between the medians is rejected at an approximate $\alpha = .05$. The Inter Quartile Range (IQR) reduced from .19 for the 'before re-test' test scores to .11 for the 'after re-test' test scores.

Figure 5 shows notched box plot displays for student test performance for the third midterm for the spring quarter 1994. The absence of overlap of the notches between the test scores of the control class and experimental class indicates that the null hypothesis in the differences between the medians is to be rejected. The absence of overlap of the notches between the test scores for the experimental class after the re-test and the experimental class before the re-test indicates that the null hypothesis in the differences between the medians would be rejected at an approximate $\alpha = .05$. The improvement in the test scores for the experimental class after the re-test is still not significantly different from the control group that did not have a re-test option. The Inter Quartile Range (IQR) for the experimental class reduced from .33 for the 'before re-test' test scores to .26 for the 'after re-test' test scores. The IQR for the control class for the third midterm was .15.

Figure 6 shows notched box displays for the final examination scores and the midterm test performance average for the control class and experimental class. The overlap of the notches between the final examination scores of the experimental class and the control class indicates that the null hypothesis in the differences between the medians is rejected.
The box plot notches for the midterm test score average indicate that the null hypothesis in
the difference between the medians are narrowly rejected at an approximate Alpha = .05.

FIGURE 3  STUDENT TEST PERFORMANCE MIDTERM ONE
SPRING 1994 EXPERIMENTAL AND CONTROL CLASSES
FIGURE 4  STUDENT TEST PERFORMANCE MIDTERM TWO
SPRING 1994 EXPERIMENTAL AND CONTROL CLASSES

FIGURE 5  STUDENT TEST PERFORMANCE MIDTERM THREE
SPRING 1994 EXPERIMENTAL AND CONTROL CLASSES
Figure 7 is a side-by-side stem and leaf diagram that displays the test performance data for the first midterm for the experimental class. The left side displays the distribution of test scores before the re-test. The right side displays the distribution of test scores after

FIGURE 6  STUDENT TEST PERFORMANCE FINAL EXAMINATION AND MIDTERM TEST AVERAGE SPRING 1994 EXPERIMENTAL AND CONTROL CLASSES
the test scores. Test scores that are bold represent the scores for the people who chose the re-test option. Test scores that are bold italic represent the scores for the people who chose the re-test option and did worse on the re-test. 63 out of 94 students took the re-test option. The test scores for these students went from a range of 37% to 93% (first test) to 55% to 97% (after re-test). 31 students chose not to take the re-test option. The range of test scores for these students was 66% to 100%.

<table>
<thead>
<tr>
<th>TEST SCORES BEFORE RE-TEST</th>
<th>TEST SCORES AFTER RE-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 21</td>
<td>5</td>
</tr>
<tr>
<td>932</td>
<td>6</td>
</tr>
<tr>
<td>995431</td>
<td>7</td>
</tr>
<tr>
<td>9 977663311</td>
<td>8</td>
</tr>
<tr>
<td>9876443 3 11</td>
<td>9</td>
</tr>
<tr>
<td>9 9 43 3 0</td>
<td>10</td>
</tr>
<tr>
<td>8766 6 544 33 2221 1 100 0</td>
<td>11</td>
</tr>
<tr>
<td>9 988 8 77 77 6666 55443 333 11 0 0</td>
<td>12</td>
</tr>
<tr>
<td>9876322100</td>
<td>13</td>
</tr>
<tr>
<td>00</td>
<td>14</td>
</tr>
<tr>
<td>00</td>
<td>15</td>
</tr>
</tbody>
</table>

FIGURE 7 STEM AND LEAF DIAGRAM TEST RE-TEST SCORES
MIDTERM ONE SPRING 1994 EXPERIMENTAL CLASS
150 Points = 100%

Figure 8 is a side-by-side stem and leaf diagram that the displays the test performance data for the second midterm for the experimental class. 68 out of 95 students took the re-test option. The test scores for these students went from a median of 77% (first test) to 91% (after re-test) and the test score range went from 37% to 95% (first test) to 59% to
100% (after re-test). 27 students chose not to take the re-test option. The median of their test score was 93 and the range of test scores for these students was 71% to 100%.

<table>
<thead>
<tr>
<th>TEST SCORES BEFORE RE-TEST</th>
<th>TEST SCORES AFTER RE-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>60</td>
<td>7</td>
</tr>
<tr>
<td>9764300</td>
<td>8</td>
</tr>
<tr>
<td>99843</td>
<td>9</td>
</tr>
<tr>
<td>9887 7 55210</td>
<td>10</td>
</tr>
<tr>
<td>6 6 55433110 0</td>
<td>11</td>
</tr>
<tr>
<td>9 9 98 8 87 66 4 332111000</td>
<td>12</td>
</tr>
<tr>
<td>977655 5433 3 22211000</td>
<td>13</td>
</tr>
<tr>
<td>85443 2 11000 000</td>
<td>14</td>
</tr>
<tr>
<td>00000</td>
<td>15</td>
</tr>
</tbody>
</table>

FIGURE 8 STEM AND LEAF DIAGRAM TEST RE-TEST SCORES
MIDTERM TWO SPRING 1994 EXPERIMENTAL CLASS
150 Points = 100%

Figure 9 is a side-by-side stem and leaf diagram that displays the test performance data for the third midterm for the experimental class. 44 out of 86 students took the re-test option. The test scores for these students went from a median of 53% (first test) to 73% (after re-test) and the test score range went from 11% to 85% (first test) to 21% to 95% (after re-test). 42 students chose not to take the re-test option. The median of their test score was 86% and the range of test scores for these students was 12% to 100%.
Discussion of Results:

One of the objectives of the re-test option was to improve the quality of learning by giving students the opportunity to learn from their mistakes without penalty. The hypothesis was that this improved quality would be observed through the increased median value of the test scores and through the reduction of the Inter Quartile Range (IQR). In language relevant to the students there should be fewer grades of C, D, and E. Tables 5, 6, and 7 show the change in the distribution of grades for the three midterm tests for the experimental class. For the first two midterms, the median re-test grades for students getting Ds and E's went up almost two letter grades (C to B+). The median re-

<table>
<thead>
<tr>
<th>TEST SCORES BEFORE RE-TEST</th>
<th>TEST SCORES AFTER RE-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>86</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>75</td>
<td>4</td>
</tr>
<tr>
<td>665210</td>
<td>5</td>
</tr>
<tr>
<td>997400</td>
<td>6</td>
</tr>
<tr>
<td>77765532110</td>
<td>7</td>
</tr>
<tr>
<td>54100</td>
<td>8</td>
</tr>
<tr>
<td>876666109</td>
<td>9</td>
</tr>
<tr>
<td>9987442321000</td>
<td>10</td>
</tr>
<tr>
<td>65422</td>
<td>11</td>
</tr>
<tr>
<td>833311</td>
<td>12</td>
</tr>
<tr>
<td>86410</td>
<td>13</td>
</tr>
<tr>
<td>96554220</td>
<td>14</td>
</tr>
<tr>
<td>000</td>
<td>15</td>
</tr>
</tbody>
</table>

FIGURE 9 STEM AND LEAF DIAGRAM TEST RE-TEST SCORES
MIDTERM THREE SPRING 1994 EXPERIMENTAL CLASS
150 Points = 100%
test grades for students getting B and C went up about one letter grade (B- to A). The students getting an A or A- typically did not take the re-test. There was apparently a problem with the third midterm. The test scores for the experimental class were particularly poor. More than 60% (54 out of 88) scored a D or E on the midterm. (The percentage of students scoring a D or E on the first two midterms was about 25%.) In addition, a low percent of the students opted to take the re-test, even after a low score. For the first two midterms, however, more than 95% of the students receiving a grade of E, D, or C chose to take a re-test. For the third midterm only 60% of the students receiving a grade of E, D, or C chose to take a re-test.

Using Deming’s terminology for cause attribution, the test score data suggest the existence of common cause (system) factor that might explain the reason for low test performance and the reason for the low percentage of students taking the re-test. A possible common cause explaining the low test performance was that the test problem construction for the third test was at fault (difficulty, ambiguity). A possible common cause explaining the low percentage of students taking the re-test was that the third midterm was returned to the students with only two days left in the quarter. This gave the students less than two days to prepare for the test. This lack of time, particularly at the end of the quarter, was very possibly a factor for the low number of students taking the re-test. In all likelihood the test construction and the lack of adequate response time may have negated the potential benefits of the re-test option for the experimental class on the third midterm.
TABLE 5  DISTRIBUTION OF TEST GRADES FOR MIDTERM ONE
BEFORE AND AFTER THE RE-TEST OPTION

<table>
<thead>
<tr>
<th>GRADE</th>
<th>NUMBER</th>
<th>NUMBER RE-TESTED</th>
<th>E</th>
<th>D</th>
<th>C</th>
<th>B</th>
<th>A</th>
<th>MEDIAN GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>13</td>
<td>13</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>16</td>
<td>14</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>C+</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>B -</td>
</tr>
<tr>
<td>B</td>
<td>29</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>14</td>
<td>13</td>
<td>B+</td>
</tr>
<tr>
<td>A</td>
<td>27</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>26</td>
<td>A</td>
</tr>
</tbody>
</table>

Median Grade Before Re-test = B -  
Median Grade After Re-test = B+  

TABLE 6  DISTRIBUTION OF TEST GRADES FOR MIDTERM TWO
BEFORE AND AFTER THE RE-TEST OPTION

<table>
<thead>
<tr>
<th>GRADE</th>
<th>NUMBER</th>
<th>NUMBER RE-TESTED</th>
<th>E</th>
<th>D</th>
<th>C</th>
<th>B</th>
<th>A</th>
<th>MEDIAN GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>B+</td>
</tr>
<tr>
<td>C</td>
<td>18</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>B+</td>
</tr>
<tr>
<td>B</td>
<td>31</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>21</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>26</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>A</td>
</tr>
</tbody>
</table>

Median Grade Before Re-test = B  
Median Grade After Re-test = A -  

TABLE 7  DISTRIBUTION OF TEST GRADES FOR MIDTERM THREE
BEFORE AND AFTER THE RE-TEST OPTION

<table>
<thead>
<tr>
<th>GRADE</th>
<th>NUMBER</th>
<th>NUMBER RE-TESTED</th>
<th>E</th>
<th>D</th>
<th>C</th>
<th>B</th>
<th>A</th>
<th>MEDIAN GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>35</td>
<td>24</td>
<td>22</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>E</td>
</tr>
<tr>
<td>D</td>
<td>19</td>
<td>9</td>
<td>0</td>
<td>11</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>D+</td>
</tr>
<tr>
<td>C</td>
<td>9</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>C+</td>
</tr>
<tr>
<td>B</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>A</td>
</tr>
</tbody>
</table>

Median Grade Before Re-test = D+  
Median Grade After Re-test = C
3. **Student Test Performance - Autumn 1994 and Winter 1995**

Figure 10 shows box plot displays for student test performance (midterm one, midterm two, midterm three, and midterm average) for the autumn quarter 1994. Figure 11 shows box plot displays for student test performance (midterm one, midterm two, midterm three, and midterm average for the winter quarter 1995. These figures show that the median test performance, before re-test, varied from 69% to 85%. The IQR for these test scores varied from 16% to 27%. After the re-test, the median test performance varied from 89% to 93%. The IQR for these test score varied from 7% to 13%.
Figure 12 is a side-by-side stem and leaf diagram of student test performance for the third midterm. This diagram is illustrative of the pattern of test score improvement for all of the midterms in the autumn and winter quarters. All students below a score of 80% on the first test took the re-test. Almost all of the students achieving a score between 80%
and 85% took the re-test. Some of the students took the re-test when their test performance exceeded 85% and 90%.

<table>
<thead>
<tr>
<th>TEST SCORES BEFORE RE-TEST</th>
<th>TEST SCORES AFTER RE-TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>7642</td>
<td>8</td>
</tr>
<tr>
<td>9776420</td>
<td>9</td>
</tr>
<tr>
<td>9765400</td>
<td>10</td>
</tr>
<tr>
<td>9987764420</td>
<td>11</td>
</tr>
<tr>
<td>88 887 6666655544 4 3300</td>
<td>12</td>
</tr>
<tr>
<td>4 4322222 211 21</td>
<td>13</td>
</tr>
<tr>
<td>999 977 7755 5</td>
<td>55 55 6677777 77 899 999</td>
</tr>
<tr>
<td>77655422 2 00000</td>
<td>00000 0122222 22 3344 4</td>
</tr>
<tr>
<td>00000</td>
<td>00000 00000</td>
</tr>
</tbody>
</table>

FIGURE 12 STEM AND LEAF DIAGRAM TEST RE-TEST SCORES MIDTERM THREE AUTUMN 1994 150 Points = 100%

Discussion of Results:

The role that test performance played in the grade that a student earned was fundamentally different in the autumn 1994 and winter 1995 quarters as compared to the traditional testing and grading system or the first testing system revision. In these two quarters there were threshold performance standards that were required for a student to achieve a C or a B grade. The C grade threshold was achieved when a student achieved a minimum test score of 70% on each midterm and an average test score of 75% for all midterms. This threshold performance was required to pass the course since the grades of
C-, D+, and D were eliminated. The B grade threshold was achieved when a student achieved a minimum test score of 80% on each midterm and an average test score of 85% for all midterms. This threshold performance guaranteed a grade of B, but higher test performance would not result in a higher grade. Students’ reaction to this grading policy was overwhelmingly positive, based in their written comments to an open-ended question on the grading and testing system. Just one student’s wrote that it was frustrating that a high test score would not receive a higher grade than a score of 85%. Even this criticism was couched in an overall favorable reaction to the system. To achieve a grade of A the student had to achieve the B threshold and then perform projects. Achievement of projects earned the students points. Each increment of 50 points earned the student one grade level increase (for example, from a B to a B+). A student could earn up to three level grade increases (a full letter grade) through projects. Projects were evaluated on a pass/fail basis. The professor would provide feedback to the students on project drafts throughout the quarter so the student could revise the project to meet the performance criteria. In the next section the student project choices will be reviewed in greater detail.

4. Student Projects

Students in the spring 1994 experimental class, the autumn 1994 class and the winter 1995 class were given an opportunity to use projects as a means for learning and demonstrating their competence for grading purposes.

In the spring 1994 experimental class students could choose to propose and complete a project that demonstrated an application of engineering economics or demonstrate a
learning inquiry project of their interest. The students had a wide latitude for choosing projects, but had to submit a proposal to the instructor for approval. When students satisfactorily completed their project they got a 100% grade that could be used to substitute for a midterm grade.

In autumn 1994 and winter 1995 students were given specific project options (each with point values assigned) as well as given the opportunity to propose a learning inquiry project to the instructor. One of the options, the spreadsheet assignments, was a required and scored assignments prior to autumn 1994. In autumn 1994 and winter 1995 the completion of projects were used to raise the students grades, not to substitute for a midterm test.

In the spring 1994 experimental class, only ten out of ninety six students completed a learning inquiry project.

Tables 8 and 9 show the distribution of students in the autumn 1994 and the winter 1995 classes completing projects at four levels corresponding to (1) No Grade Level Increase, (2) One-Grade Level Increase, (3) Two-Grade Level Increase, and (4) Three-Grade Level Increase. Tables 10 and 11 show the types of options these students completed in the autumn 1994 and the winter 1995 classes.
### TABLE 8  DISTRIBUTION OF PROJECTS COMPLETED AT FOUR LEVELS - AUTUMN 1994

<table>
<thead>
<tr>
<th>PROJECT LEVEL</th>
<th>PERCENTAGE OF STUDENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Than 50 Points</td>
<td>17</td>
</tr>
<tr>
<td>50 &amp; Less Than 100 Points</td>
<td>5</td>
</tr>
<tr>
<td>100 &amp; Less Than 150 Points</td>
<td>27</td>
</tr>
<tr>
<td>150 Or More Points</td>
<td>51</td>
</tr>
</tbody>
</table>

### TABLE 9  DISTRIBUTION OF PROJECTS COMPLETED AT FOUR LEVELS - WINTER 1995

<table>
<thead>
<tr>
<th>PROJECT LEVEL</th>
<th>PERCENTAGE OF STUDENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Than 50 Points</td>
<td>8</td>
</tr>
<tr>
<td>50 &amp; Less Than 100 Points</td>
<td>7</td>
</tr>
<tr>
<td>100 &amp; Less Than 150 Points</td>
<td>21</td>
</tr>
<tr>
<td>150 Or More Points</td>
<td>63</td>
</tr>
</tbody>
</table>

### TABLE 10  TYPES OF PROJECT OPTIONS COMPLETED BY STUDENTS - AUTUMN 1994

<table>
<thead>
<tr>
<th>TYPE OF PROJECT</th>
<th>PERCENTAGE OF STUDENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet I</td>
<td>68</td>
</tr>
<tr>
<td>Spreadsheet II</td>
<td>41</td>
</tr>
<tr>
<td>Interview</td>
<td>59</td>
</tr>
<tr>
<td>Journal</td>
<td>37</td>
</tr>
<tr>
<td>Guest Write Up</td>
<td>49</td>
</tr>
<tr>
<td>Cartoon/CBS</td>
<td>10</td>
</tr>
<tr>
<td>Problem Analyses</td>
<td>36</td>
</tr>
<tr>
<td>Self-Initiated Project/Case</td>
<td>10</td>
</tr>
</tbody>
</table>
TABLE 11 TYPES OF PROJECT OPTIONS COMPLETED BY STUDENTS - WINTER 1995

<table>
<thead>
<tr>
<th>TYPE OF PROJECT</th>
<th>PERCENTAGE OF STUDENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spreadsheet I</td>
<td>72</td>
</tr>
<tr>
<td>Spreadsheet II</td>
<td>46</td>
</tr>
<tr>
<td>Interview</td>
<td>67</td>
</tr>
<tr>
<td>Journal</td>
<td>47</td>
</tr>
<tr>
<td>Guest Write Up</td>
<td>56</td>
</tr>
<tr>
<td>Cartoon/CBS</td>
<td>9</td>
</tr>
<tr>
<td>Problem Analyses</td>
<td>11</td>
</tr>
<tr>
<td>Self-Initiated Project/Case</td>
<td>4</td>
</tr>
</tbody>
</table>

Discussion of Results:

In autumn 1994 and winter 1995, students chose the more clearly defined (by the instructor) projects (spreadsheets, problem analysis, guest write-up) over self-initiated projects. The expectations and evaluation criteria for spreadsheet assignments, problem analysis and guest write-ups were clearer for these projects as compared with the self-initiated projects, which required the students to propose and generate the expectations and evaluation criteria. The percentage of students choosing self-initiated projects was 10% for the spring 1994 experimental class, 10% for the autumn 1994 class, and 4% for the winter 1995 class. Perhaps the inherent uncertainty required by self-initiated projects was a deterrent to students.

In autumn 1994 and winter 1995, a sizable percentage of students (37% and 47% respectively) of the students kept journals throughout the quarter. A sizable percentage of students (59% and 67%) also participated in an interview with the instructor. Both of these activities were personal and self-disclosing.
5. **Cross-sectional Analysis of Student Performance**

Figure 13 shows student performance based on the percentage of students achieving a three-midterm test score average of 85% or higher. This performance level, a threshold level required for a grade of B was communicated by the instructor as the quality standard for the course for the autumn 1994 and winter 1995 classes. The percentage of students achieving this quality standard can therefore be taken as one measure of the quality of the course. This quality standard can be viewed as the joint responsibility of the instructor and students (collectively). The percentage of students achieving a three-midterm test average of 85% or higher went from approximately 30% for the old testing system (based on five-quarter average from autumn 1992 to spring control class 1994), to a little over 50% for the first ‘test-re-test’ testing system (spring 1994 experimental class), to approximately 85% for the second ‘test-re-test’ testing system (autumn 1994 and winter 1995).

Figure 14 shows student performance based on the percentage of students earning a course grade of B or greater. The percentage of students earning a grade of B or greater went from approximately 50% for the old grading system (four quarter average from winter 1993 to spring 1994 control class) to 60% for the first grading system design (spring 1994 experimental class) to approximately 90% for the second grading system design (autumn 1994 and spring 1995).
FIGURE 13 PERCENT OF STUDENTS ACHIEVING MIDTERM AVERAGE OF 85% OR HIGHER

FIGURE 14 PERCENT OF STUDENTS EARNING A COURSE GRADE OF B OR GREATER
Discussion of Results:

The increase in the percentage of students achieving a midterm test score average of 85% or better in student performance from the old testing system to the first ‘test-re-test’ testing system (spring 1994 experimental class) may or may not be significant. There are many possible sources of variation that may be masking or confounding the results for the five quarters prior to the spring 1994 experiment. The possible sources of variation include the students grade point average, test construction, grader judgments, and teaching methods. The apparent common cause problem experienced in the third midterm exam in the spring 1994 experimental class may have reduced the positive impact of the ‘test-re-test’ option. The increase in the percentage of students achieving a three-midterm test average of 85% or higher for the ‘test-re-test” testing system in the autumn 1994 and winter 1995 classes appear to be significant, both statistically and practically. Possible explanations for this improvement include:

- The instructor continued to refine and improve the quality of the test construction and the time for students to prepare for the re-test in the autumn 1994 and winter 1995 classes.

- A test score quality standard of 85% was clearly communicated by the instructor to the class.

- The establishment of minimum threshold performance levels and the elimination of C - and D grades increased the importance of achieving threshold performance in the autumn 1994 and the winter 1995 classes. In the
spring 1994 experimental class, the improvement in test-re-test scores only had an impact on part (450 points out of a total of 800 points). Quizzes, spreadsheet assignments, and the final examination were tested using the traditional no-re-test option.

The increase in the number of students achieving a grade of B or greater is also more pronounced for the autumn 1994 and winter 1995 classes. The spring 1994 experiment directly evaluated only a new testing system. There was not a fundamental change in the grading system. In the spring experiment, grades were indirectly affected because of the students’ ability to improve their test scores. In the autumn 1994 and winter 1995 classes there was a fundamental change in the grading system. Some of these changes include:

- After the threshold level of 85% for midterm performance was achieved test score performance was not factored into the grade.
- All projects were evaluated on a ‘go/no go’ basis. The instructor provided ongoing feedback to the students to assure that the projects met the ‘go’ criteria. This ‘no punishment’ feedback is a significant de-coupling of formative and summative evaluation.
- The instructor communicated clearly that every student could achieve a grade of A. The tone of the communication, along with the personal connection between the instructor and student, also communicated the instructor’s belief and desire that all students achieve a high grade. Students were given a variety
traditional and non-traditional ways to demonstrate their competence in engineering economics.

6. Student Evaluation of Teaching Rating of Course and Instructor

Student satisfaction with the testing and grading system was measured by using the ratings made by the students on questions 17 through 20 on the S.E.T. These questions are shown in Table 3 in Chapter III. These ratings are displayed on attribute control charts (p-charts) in which the value of $p$ is the percent of favorable (strongly agree + agree) responses to total responses. The value of $n$ is the average number of students in a class for the baseline period. This number is 100. The control limits are plus and minus three standard deviations based on the binomial distribution formula:

$$\text{One standard deviation} = \sqrt{\frac{p \times (1-p)}{n}}$$

(2)

A chance occurrence that a value would fall outside the control limits is approximately .005. If two out of three consecutive points fall outside the two standard deviation limit on the same side of the centerline (mean value) the probability that this is a chance occurrence is also approximately .005. For both of these conditions the null hypothesis would be rejected and we would search for an assignable cause. For purposes of this research we are interested in the impact of the changes in the testing and grading system on student satisfaction in the spring 1994 experimental class and the autumn 1994 and winter 1995 classes.
Table 12 shows the attribute control chart data for the S.E.T. questions concerning students' reaction to the course's testing and grading system. The Base Line Mean is the mean rating of the S.E.T. question for the four base line quarters (autumn 1992, winter 1993, autumn 1993, winter 1994). The Lower Control Limit (LCL) is the value that is three standard deviations below the mean. The Upper Control Limit (UCL) is the value that is three standard deviations above the mean. To reject the null hypothesis that there is no difference in the students' reaction to the testing and grading system the value for the question would need to be higher than the UCL.

### Table 12 Attribute Control Chart Data

<table>
<thead>
<tr>
<th>S.E.T. QUESTION</th>
<th>BASE LINE MEAN</th>
<th>LCL</th>
<th>UCL</th>
<th>SP. 94C SCORE</th>
<th>SP. 94E SCORE</th>
<th>AUT. 94 SCORE</th>
<th>WI. 95 SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>.69</td>
<td>.55</td>
<td>.83</td>
<td>.69</td>
<td>.81</td>
<td>.91</td>
<td>.92</td>
</tr>
<tr>
<td>18</td>
<td>.73</td>
<td>.60</td>
<td>.86</td>
<td>.71</td>
<td>.84</td>
<td>.94</td>
<td>.93</td>
</tr>
<tr>
<td>19</td>
<td>.86</td>
<td>.75</td>
<td>.96</td>
<td>.84</td>
<td>.90</td>
<td>.97</td>
<td>.97</td>
</tr>
<tr>
<td>20</td>
<td>.75</td>
<td>.62</td>
<td>.88</td>
<td>.57</td>
<td>.78</td>
<td>.98</td>
<td>.92</td>
</tr>
</tbody>
</table>

Figure 15 is the attribute control chart for the S.E.T. score for Question 17. The values for this question for both the spring control and experimental classes are within the control limits. Consequently, the null hypothesis is not rejected. The S.E.T. value of question 17 for the spring 1994 experimental class is 2.6 standard deviations higher than the value for control class, suggesting a possible positive impact of the first testing and grading revision on student satisfaction. The S.E.T. values of question 17 for both the autumn 1994 and winter 1995 classes are higher than the UCL. The null hypothesis is
rejected. The additional revisions to the testing and grading system in the autumn 1994 and the winter 1995 are asserted to be the most likely cause for this difference. This assertion is strengthened by the responses made by the students' responses to an open-ended question on the testing and grading system. These responses will be reviewed in the next section. The attribute control charts for questions 18, 19, and 20 are in the appendix. Each of these control charts show a similar pattern as Figure 15.

![Attribute Control Chart](image)

**LEGEND**
- Quarter 1 - Autumn 1992
- Quarter 2 - Winter 1993
- Quarter 3 - Autumn 1993
- Quarter 4 - Winter 1994
- Quarter 5C - Spring 1994 (Control)
- Quarter 5E - Spring 1994 (Experimental)
- Quarter 6 - Autumn 1994
- Quarter 7 - Winter 1995

**FIGURE 15** ATTRIBUTE CONTROL CHART S.E.T. QUESTION 17.
Exploratory data analyses were conducted on the students' ratings of other aspects of the course and course instructor. Values for four additional S.E.T. questions (questions 5 through 8) show a similar pattern as the values for S.E.T. questions 17 through 20. Table 13 shows these questions.

**TABLE 13 S.E.T. QUESTIONS 5 THROUGH 8**

<table>
<thead>
<tr>
<th>S.E.T. QUESTIONS</th>
<th>BASE LINE MEAN</th>
<th>LCL</th>
<th>UCL</th>
<th>SP. 94C SCORE</th>
<th>SP. 94E SCORE</th>
<th>AUT. 94 SCORE</th>
<th>WI. 95 SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.44</td>
<td>.29</td>
<td>.59</td>
<td>.26</td>
<td>.34</td>
<td>.76</td>
<td>.70</td>
</tr>
<tr>
<td>6</td>
<td>.59</td>
<td>.44</td>
<td>.74</td>
<td>.45</td>
<td>.50</td>
<td>.86</td>
<td>.82</td>
</tr>
<tr>
<td>7</td>
<td>.72</td>
<td>.58</td>
<td>.85</td>
<td>.63</td>
<td>.66</td>
<td>.86</td>
<td>.85</td>
</tr>
<tr>
<td>8</td>
<td>.65</td>
<td>.51</td>
<td>.79</td>
<td>.49</td>
<td>.56</td>
<td>.88</td>
<td>.87</td>
</tr>
</tbody>
</table>

Table 14 shows the attribute control chart data for the S.E.T. questions 5 through 8. The values for these questions for the spring 1994 experimental class are well within the control limits. There is no evidence to suggest that the first revision in the testing and grading system had any impact on student satisfaction as measured by these questions. The values for these questions, however, for both the autumn 1994 and winter 1995 classes exceed the UCL. This suggests an assignable cause that was not predicted in the research design.

**TABLE 14 ATTRIBUTE CONTROL CHART DATA S.E.T. QUESTIONS 5 THROUGH 8**
Figure 16 is the attribute control chart for question five. The additional revisions to the testing and grading system in the autumn 1994 and the winter 1995 are asserted to be the most likely cause for this difference. The attribute control chart data for questions six, seven, and eight in the appendix show a similar pattern as the one for question five.

Table 15 shows the remaining twelve questions of the S.E.T. Table 16 shows the attribute control chart data for these 12 S.E.T. questions. The values for each of these
questions are within the control limits for the spring 1994 classes. The values for each of these questions are generally within the control limits for the autumn 1994 and winter 1995 classes. The values for four of the questions (3, 4, 12, 14) are suggestive of higher student satisfaction, but inconclusive.

TABLE 15  S.E.T. QUESTIONS 1 THROUGH 4
AND 9 THROUGH 16

<table>
<thead>
<tr>
<th>S.E.T. QUESTION</th>
<th>BASE LINE MEAN</th>
<th>LCL</th>
<th>UCL</th>
<th>SP. 94C SCORE</th>
<th>SP. 94E SCORE</th>
<th>AUT. 94 SCORE</th>
<th>WI. 95 SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>88</td>
<td>.78</td>
<td>.98</td>
<td>.83</td>
<td>.96</td>
<td>.96</td>
<td>.96</td>
</tr>
<tr>
<td>2</td>
<td>94</td>
<td>.87</td>
<td>1.00</td>
<td>.94</td>
<td>1.00</td>
<td>.98</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>.75</td>
<td>.62</td>
<td>.88</td>
<td>.63</td>
<td>.61</td>
<td>.83</td>
<td>.90*</td>
</tr>
<tr>
<td>4</td>
<td>.67</td>
<td>.53</td>
<td>.81</td>
<td>.59</td>
<td>.54</td>
<td>.84*</td>
<td>.80</td>
</tr>
<tr>
<td>9</td>
<td>82</td>
<td>.71</td>
<td>.93</td>
<td>.74</td>
<td>.72</td>
<td>.57</td>
<td>.86</td>
</tr>
<tr>
<td>10</td>
<td>82</td>
<td>.71</td>
<td>.93</td>
<td>.71</td>
<td>.76</td>
<td>.92</td>
<td>.86</td>
</tr>
<tr>
<td>11</td>
<td>69</td>
<td>.55</td>
<td>.83</td>
<td>.70</td>
<td>.63</td>
<td>.86</td>
<td>.80</td>
</tr>
<tr>
<td>12</td>
<td>81</td>
<td>.69</td>
<td>.93</td>
<td>.75</td>
<td>.82</td>
<td>.91**</td>
<td>.90**</td>
</tr>
<tr>
<td>13</td>
<td>.71</td>
<td>.58</td>
<td>.84</td>
<td>.65</td>
<td>.63</td>
<td>.70</td>
<td>.74</td>
</tr>
<tr>
<td>14</td>
<td>84</td>
<td>.73</td>
<td>.95</td>
<td>.73</td>
<td>.73</td>
<td>.94**</td>
<td>.91**</td>
</tr>
<tr>
<td>15</td>
<td>.74</td>
<td>.61</td>
<td>.87</td>
<td>.78</td>
<td>.61</td>
<td>.54</td>
<td>.74</td>
</tr>
<tr>
<td>16</td>
<td>.78</td>
<td>.66</td>
<td>.90</td>
<td>.78</td>
<td>.70</td>
<td>.80</td>
<td>.85</td>
</tr>
</tbody>
</table>

* Value above UCL
** Two values above Base Line Mean + two standard deviations
Discussion of Results:

The impact of the first revision of the testing and grading system on student satisfaction is inconclusive. All values for the S.E.T. questions for the spring 1994 experimental class that were concerned with testing and grading were within the control limits. There is a suggestion of higher student satisfaction when the spring 1994 experimental class is compared with the control class. For example, the students in the spring 1994 experimental class gave an average rating of .81 for question 17 and the control class gave an average rating of .69 for this question. The difference (.81 - .69 = .12) is 2.6 standard deviations (.12/.046) higher. The difference for questions 18, 19, and 20 is similar with the experimental class giving higher ratings (approximately 2 standard deviations) than the control class. This data alone is inconclusive, but with the significant rise in values in student satisfaction for these four questions for the autumn 1994 and winter 1995 classes, suggest a possible incremental improvement as the testing and grading system continued to be revised.

The data suggests that the second set of revisions for the testing and grading system for the autumn 1994 and winter 1995 classes had a positive impact on student satisfaction. All S.E.T. values for the four testing and grading system questions were above the UCL for both classes. The consistently high scores for two consecutive quarters is suggestive of a significant positive impact of the revisions to the testing and grading system that the instructor made.
The data suggest that the second set of revisions to the testing and grading system had a positive impact on the students' reaction to the course and to the instructor (S.E.T. Questing 5 through 8). These four questions did not deal with specific instructor classroom behaviors. Instead, the questions indicated an overall reaction to the course and the instructor. Additional insights into the students' positive reaction to the course and the instructor is discussed in the next session.

7. Content Analysis of Students' Comments

The Industrial and Systems Engineering Department asks students to provide feedback on courses by responding to a number of open ended questions. The question used for testing and grading is ASPECTS OF THE EXAMS (FAIRNESS, .... ETC.). The comments made by the students were typed and sorted for the autumn 1993 class, the winter 1994 class, the spring control 1994 class, the spring experimental 1994 class, and the autumn 1994 class. The winter 1995 comments were not available. The comments for each class were sorted into two categories: those comments reflecting a negative or critical reaction to testing and grading; and those comments reflecting a positive or supportive reaction to testing and grading. Figure 17 summarizes the results of this first sorting.

For the autumn 1993 class and the winter 1994 class the percent of responses that were positive were 27% and 40% respectively. The positive comments reflected the students opinion that the tests were generally fair. The negative comments were of two types - a perception of unfairness (difficulty of questions, relationship to lecture, clarity of
questions and length of test), and a perception that the length of time to return test was too long.

For the spring 1994 control class the percent of responses that were positive were 56%. As in the previous two classes the positive comments were that the tests were generally fair. The perception of difficulty, relationship to lecture, clarity and length were also similar to the other quarters. The improvement in the percentage of positive comments can be attributed almost entirely to the students reaction to promptness of test return. There were no negative comments on promptness and there were a few positive comments on promptness. There were a few comments complaining of the unfairness that the morning section (the experimental class) had a re-test option and they didn’t. These negative comments were not included because they were created by the experiment.

The percentage of positive comments increased to 77% for the spring 1994 experimental class and to 84% for the autumn 1994 class. In addition to generally positive comments about the fairness of the exams, students gave generally positive remarks to the re-test option. A number of students stated that the re-test option reduced the feeling of time pressure. Students also appreciated the exam packet that was available for them. There were a few mildly critical comments of the re-test policy. They felt that there should be some consideration for the students who didn’t require a re-test.

The students responded to the open ended questions, THINGS THE INSTRUCTOR HAS DONE WELL, and, THINGS THE INSTRUCTOR COULD IMPROVE. Prior to the new testing and grading system the themes that emerged were that the instructor was
a good lecturer and was accessible to them. They felt that improvements could be made
in quality of the visuals, pacing of the lectures and the fairness of the exams. Personal
comments about the instructor were about even - from "he shows concern for students" to
"he responds critically to my questions".

The spring 1994 experimental class comments were similar except that testing and
grading comments were positive. The autumn 1994 class had very positive comments
about the testing and grading system. The most significant new theme that emerged was a
high personal regard for the instructor. The students expressed that the instructor was
concerned about students, created a positive learning environment, and was open minded.
There was not one critical comment about the instructor’s interpersonal style.

FIGURE 17 PERCENT OF POSITIVE RESPONSES OPEN ENDED QUESTIONS TESTING AND GRADING
Additional Discussion of Results:

The S.E.T. data and the written comments suggest that the students' experienced high satisfaction with the second set of revisions to the testing and grading system in the autumn 1994 and winter 1995 classes. Possible explanations for this apparent improvement are proposed. These explanations will be based on the data and the literature in ALT and TQM.

From the literature on self-directed learning the positive student reaction could be that the students felt a higher level of personal autonomy because of the new testing and grading system. This personal autonomy could have been experienced through the students' ability to improve their performance without penalty and through their ability to choose alternative methods to demonstrate competence through projects. A possible explanation from the quality literature for this positive reaction is also proposed. The new testing and grading system may have reduced the fear and anxiety typically associated with learning under a traditional testing and grading system. Performance on a specific test or on a first iteration of a project had very little impact on the student's grade in the autumn 1994 and winter 1995 classes. The students' attention, as Kohn suggests, may have been focused on learning rather than their performance. Finally, the adult learning literature on the role of the facilitator may provide a possible explanation for the high regard that the students appeared to have for the course instructor. This instructor's facilitative role may have been strengthened by his personal interviews with students. In addition, the instructor openly and personally communicated his belief and desire that all students
could achieve a high level of proficiency in the course content. A strong theme in the adult learning literature is the importance of a warm, collaborative relationship between student and teacher. These personal interviews and the genuinely positive nature of the instructor’s communication may have facilitated such a relationship. This was particularly notable because of the large class size.

The data on student satisfaction and performance does not provide conclusive evidence for these explanations, but the explanations are consistent with the data and with the literature. They offer a rich source for future research.
CHAPTER V
CONCLUSIONS

1. Introduction

The purpose of this chapter is to present a summary of results of the research study, review the major findings and propose questions for research and practice. Five fundamental flaws in the education of engineers are proposed from both the research findings and the literature on adult learning and total quality management. A vision for a high quality learning system design for an undergraduate engineering department is presented. This vision is not limited to the testing and grading paradigm. The vision integrates the other four paradigms that the faculty learning inquiry team developed for the re-design of the course in engineering economic analysis. Finally, the implications of this vision to the preparation of students and faculty are reviewed.

2. Summary of Results

The following results are inferred from the data analyses made in Chapter IV:

A. The Spring 1994 Experiment Test Re-Test Option

- No significant difference was found in the final exam test scores as measured by the median and the Inter Quartile Range (IQR) between the experimental class and the control class.

- The median test scores for midterm one and two after the re-test option for the experimental class when compared to the initial test scores for midterms
one and two was significantly higher. There was a reduction in the IQR of test scores of 38% for the re-test scores for midterm one and two when compared with the IQR for midterms one and two. The median improvement for students with initial grades of D and E was approximately two letter grades. The median improvement for students with initial grades of C and B was approximately one letter grade. Students with initial grades in the A range typically did not take a re-test.

- There was a marginal improvement, at best, in the median test score after the re-test for midterm three for the experimental class as compared with the initial test scores. Similarly, the reduction in the variation in test performance as measured by the IQR was also minimal. The initial test performance for the experimental class was significantly lower than control class and significantly lower than the experimental class's initial test performance on midterms one and two. Approximately 40% of the students with an initial grade of C, D, or E on midterm three chose not to take a re-test. This compares with less than 10% not taking a re-test on midterms one and two.

- Students’ rating of the four S.E.T. questions dealing with testing and grading did not significantly improve for the experimental class as compared with the four base line classes prior to spring 1994 using the old testing and grading system.
• Students’ rating of the four S.E.T. questions dealing with testing and grading
and their responses to an open-ended question on the first testing and grading
system revision suggests that the experimental class had a positive response to
the test-re-test option when compared with the spring 1994 control class.

Discussion of Results:

The changes made in the testing and grading system with the spring experimental class
did not appear to result in significant improvement in student performance or satisfaction.
The data infers a common cause problem that had a major impact on class performance of
the third midterm examination for the experimental class. It is not the purpose of this
study to determine this common cause. But, whatever the cause was, the testing and
grading system design did not adequately respond so that the problem could be diagnosed
and corrected. The students in aggregate were apparently expected to correct the problem
as if it were a number of independent special cause problems. I propose that the lack of
time for the students, individually and in a group (with the instructor), to diagnose the
problem and take corrective actions, was the immediate cause of low test performance and
the low incremental improvement in test performance. I further propose a more
fundamental root cause of the problem ... that an adequate diagnostic process was not in
place to deal with common cause problems in the learning process. This lack of a
diagnostic process may reflect a mind set that is a residual of the traditional testing and
grading paradigm. The theory-in-action that may be inferred from the lack of a response
to the third midterm may be: test performance reflects only student ability,
motivation, and effort. I suggest that this inferred theory-in-action is in conflict with the espoused theory that: the instructor and the students have a shared responsibility for the success of the learning process. This espoused theory of shared responsibility is supported directly in the testing and grading paradigm ("Testing and grading will move from an after-the-fact faculty evaluation of students' performance to a real-time feedback mechanism that results in continuous improvement of the learning system" ..... "Students use testing to diagnose their own learning needs. The faculty use testing to diagnose effectiveness of learning interventions."). the team paradigm ("Students and faculty will collaborate as fellow learners in an atmosphere of mutual respect and trust."). and the facilitator paradigm ("Together this partnership builds a creative and synergistic learning environment."). I propose this mind set as a possible root cause not to make excuses nor to find blame. This glitch in the application of the new testing and grading system is an opportunity to learn and grow.

I propose two additional factors that may have contributed to the disappointing results of the spring 1994 experiment. I propose that the institution’s practice of limiting a course to a fixed period of time (in this case, a ten-week quarter) is a performance barrier to the instructor and to the students. This time barrier forces the instructor to give tests at an interval to meet the university’s scheduling needs. This interval may be at the detriment of the learners’ needs and sound teaching practices. For example, in the previous problem with the third midterm, the students had only two days left in the quarter to prepare for the re-test. Their attention was likely divided among other end-of-quarter requirements of
other courses. In addition, the instructor had little time to analyze the test performance data in total and there was no class time left for additional remediation and instruction.

Finally, I propose that the first testing and grading revision was a mixture of the new and old testing and grading paradigms. For example, the traditional testing and grading system was used with the experimental class on all quizzes, spreadsheet assignments and the final examination. The positive impact of the re-test option was limited, therefore, to about half of the course. This may have diluted the impact of the testing and grading revisions, thus affecting overall performance and satisfaction.

B. Student Performance and Satisfaction - Autumn 1994 and Winter 1995

Although a pre-experimental research design was used for the autumn 1994 and winter 1995 classes, the data provide an opportunity to analyze student test performance and satisfaction for three different testing and grading systems spanning eight class offerings. The autumn 1992, winter 1993, autumn 1993, winter 1994, and spring 1994 (control) classes provide data on the old testing and grading system. The spring 1994 experimental class provides data on the first revision to the testing and grading system. The autumn 1994 and winter 1995 classes provide data from the second revision to the testing and grading system. Following is summary of the analyses from the autumn 1994 and winter 1995 data.

- The median test scores for all three midterms after the re-test option was significantly higher when compared with the initial test scores. The median test
performance after the re-test for both classes varied between 89% and 93%.

The IQR was reduced by approximately 50% for all midterms.

• The percentage of students achieving a three-midterm test average of 85% or higher was 80% for the autumn 1994 class and 86% for the winter 1995 class. As a comparison, the percentage of students achieving an 85% average for the five classes using the old testing and grading system was about 30%. The percentage of students achieving an 85% for the spring 1994 experimental class was a little over 50%. Without claiming statistical significance, this data suggest that the second set of revisions to the testing and grading system may have had a positive impact on student test performance.

• Students’ ratings for each of the four S.E.T. questions dealing with testing and grading showed a significant increase when compared with the four base line classes for both the autumn 1994 and winter 1995 classes. Although this was not a controlled experiment, the data suggest high student satisfaction with the second set of revisions to the testing and grading system. This suggestion is strengthened by the content analyses of the student comments from the autumn 1994 class. 80% of the students’ remarks to an open ended question on testing and grading were positive. This compares to less than 35% for the same open-ended question by students from two base line classes (autumn 1993 and winter 1994). In the students’ comments (autumn 1994) they
specifically mentioned their approval of the re-test option and reduced fear and test anxiety.

- Students' ratings for each of four questions dealing with their general reaction to the course and to the instructor showed a significant increase for the autumn 1994 and winter 1995 classes when compared to the four baseline classes and the two spring 1994 (control and experimental) classes. The size of the increase (all points were above the Upper Control Limit of the attribute control charts) was significant. These questions did not deal with the testing and grading system. The questions dealt with the general student reaction to the course and the instructor. The students indicated that they felt the instructor was genuinely concerned for them as individuals, created a positive learning environment and was committed to the continuous improvement of the learning process.

Discussion of Results:

Student test performance and satisfaction improved significantly in both the autumn 1994 and winter 1995 classes. Although experimental conditions were not strictly adhered to, the results are compelling enough to suggest that further revisions to the testing and grading system had a significant impact on performance and satisfaction. In this section I will pose possible reasons why the testing and grading system revisions provided a positive impact, and specifically how this improvement is supported by the literature in
adult learning and total quality management. These reasons are not posed as evidence of statistical significance, but to support further research into the phenomena discovered.

The instructor communicated high performance standards and reinforced these standards by eliminating the grades of C- or lower. Although a grade of C was accepted, I propose that the instructor and the vast majority of students tacitly agreed tacitly that a grade of B was performance norm for the class. The instructor communicated this directly and indirectly. In his direct communication, through the course syllabus and through his personal explanation of the testing and grading system, he stated that the grade of B was the performance norm. To support this norm, students were provided a number of ways to achieve this performance standard. They could elect to be re-tested on up to three questions per test. If necessary, students could repair their tests after the re-test. In addition, students could improve their grade by up to one letter by electing to complete one or more pass/fail projects. The students showed their tacit approval of this performance standard by the test performance level in which they chose to take a re-test. Over 90% of the students scoring below 85% (the B threshold) on a midterm chose to take a re-test. I propose that the establishment of this threshold performance standard and the support of its achievement aided the establishment of a positive learning environment that is consistent with adult learning literature. Specifically, it supports the seven conditions of learning proposed by Knowles (1980) and the collaborative learning relationship between student and teacher discussed by Brookfield (1988) and Candy (1991). The testing and grading system implemented by the course instructor closely
parallels Glasser’s (1990) proposed grading system that only accepts a B or an A as acceptable grades. Glasser further recommends that students be given an opportunity to improve upon (repair) previous C performance. I propose that the improvement in the students’ rating on selected S.E.T. questions for the autumn and winter classes may be evidence of the reduction/elimination of coercion that Glasser suggests is inherent to traditional grading systems. The students’ positive comments concerning the reduction of test anxiety and fear supports the Deming’s (1986) admonition to eliminate fear as a requisite to high quality performance.

The inclusion of learning projects into the course design may be considered as the instructor’s attempt to provide an opportunity for the students to exercise greater self direction in their learning. The relatively low percent (less than 10%) of students that selected their own learning projects may support Candy’s (1991) assertion that there is a situational component to self-directed readiness. I propose that the traditional undergraduate engineering course design, along with traditional testing and grading, does not presently encourage self-directed learning. In traditional testing and grading systems the instructor chooses what is to be tested, the evaluation criteria, and the way the student will present evidence of achievement. This does not encourage the students’ active involvement or initiation. This lack of encouragement may be a strong situational deterrent to student self-direction, and may be a factor in students not choosing their own learning inquiry projects when given an opportunity. Freshmen and sophomore students currently involved in the Gateway project (1992) may be experiencing greater
encouragement to be self-directed than previous engineering students. These students may be more open to the freedom that this course offers.

Finally, the instructor’s personal interaction with the students through personal interviews and through opportunities for feedback on projects throughout the quarter appeared to have established a warm, caring and empathic environment between the instructor and students. It is uncertain what direct impact this may have had on student performance, but it appears to have had a direct impact on student satisfaction. This relationship between performance and satisfaction would be an interesting and useful area for further research.

3. Major Findings and Recommendations

The major findings and recommendations will be presented in three categories: (1) Diagnosis of the learning process; (2) Evaluation criteria for learning outcomes; and (3) Institutional barriers to learning.

The purpose of testing is to diagnose the learning process. Students need to be able to diagnose their own needs --- to discover what they know and don’t know about the subject --- to discover how they learn --- to discover how their interests, beliefs and behaviors affect their performance as learners. The re-test option provided some information about what they knew and it gave them some time to respond. To achieve the full benefit of the new testing paradigm will require that the student be trained in how to diagnose his/her learning needs. This diagnosis needs to be both subject-specific and learner-specific, so that the student analyzes both content and process issues. Subject-
specific diagnosis can be facilitated by the use of instructional technology processes. The overuse of the lecture method for teaching engineering courses is a learning detriment. From a diagnostic perspective, the process of lecture - classroom testing -- scoring by the instructor -- return to the student -- diagnosis -- relearn -- re-test is too long. Students are already engaged in a new learning module before they have successfully completed a prior one.

Mastery learning and self-instructional learning modules can be a useful part of a strategy for high quality learning. The research suggests that these techniques can facilitate high performance in learning while giving the students greater control over the learning process. These techniques are useful for learning the fundamental building blocks for a course. These learning methods may be less useful, however, for the development of the higher order skills that Tribus calls know-how and wisdom.

Students also need to develop skills for diagnosing their own learning process. In the quality classroom described by Langford (1983) and Tribus (1990) students used a variety of statistical methods to help students diagnose these needs. Paulson (1991) indicated that one of the key elements of the portfolio method of evaluation is the display of the process of learning as well as the outcomes of learning. Journal keeping can also be used as a means of personal reflection and self analysis. The use of journals by students in the autumn 1994 class and the winter 1995 class is encouraging and suggests a method for helping students to develop self understanding of their own learning process and their personal reflection on the course content. In these two class offerings the students were
invited to make a journal entry concerning their experience with the course. I hypothesize that the value of this activity would further increase if the students were (1) given some insights into the process of keeping a journal and dialogue, (2) given open-ended and sentence-ending prompts to help them reflect more deeply on the content and their personal experience with it, and (3) asked to review their journal and write a reflection paper. This might stimulate greater self-understanding and critical thinking. Further, the instructor could also keep a journal and openly share his or her own personal insights with the students.

The diagnosis of learning system issues also appears to be a need to ensure high quality learning. I propose that the typical feedback loop for the instructor is also too long to diagnose the teaching processes. If a particular teaching strategy is not working, the instructor needs to diagnose and revise it so that the teaching strategy can be improved for the current class. The instructor in this course in engineering economics used an outside facilitator to come in and conduct a mid-quarter evaluation. This gave him an opportunity to make improvements during the quarter. The experiences reported by Smith (1993) and Langford (1990) also suggest that the instructor should invest significant time at the beginning of the course learning about the students — their needs, interests and learning preferences. As co-managers of the learning process the instructor and students need to engage each other in meaningful dialogue about the process of learning as well as the content of learning. The three questions (What’s working? What’s not working? What can be done to improve the process?) used by the outside facilitator for the engineering
economics class can be used by all instructors as a normal part of the learning process. In addition to providing useful information for diagnosis, these questions also support a collaborative relationship between the instructor and students. The student would be treated as a co-manager with and as a customer of the instructor.

The purpose of evaluation is to ensure the quality of learning outcomes. This research study did not deal with the issue of what should be the learning outcomes in a course in engineering economics. The identification of meaningful learning outcomes is a major educational issue. Tribus (1992) strongly asserts that learning outcomes should include know-how, wisdom and character. He cites methods to identify and organize learning objectives for three of these four levels of learning outcomes. Drexel University’s (1992) pioneering efforts to re-design its engineering curriculum by including design projects, integrating courses, using co-op experiences, and teaching lifelong learning skills challenges the more narrow course-orientation of most engineering departments’ curricula. Portfolio assessment provides a framework for both developing and evaluating a more holistic set of learning outcomes. Portfolio assessment can be used to include the legitimate interests and needs of multiple customers to the learning process. Although this research study did not directly address the issue of learning outcomes, the instructor introduced a variety of projects and the use of journals in the course design. Projects and journals are two building blocks for portfolios. Their effective use may, however, be limited because of institutional barriers. Some of these barriers are discussed in the following paragraphs.
Institutional systems and structures are barriers to high quality learning. The rigid adherence to a calendar system to frame the learning experience for students is a barrier to high quality learning. Stice (1979) indicated that the 100% performance standard of mastery learning and the Keller method was lowered in a number of educational institutions to accommodate the school's schedule. In the spring 1994 experiment in the course in engineering economic analysis, the end of the quarter pressures frustrated the diagnostic value of the re-test option. The problem statement in which this research study was built is that the close-coupling of testing and grading is a deterrent to learning. I suggest that the university's fixed course schedule may not allow adequate time for diagnosis. The primary reason for the de-coupling of testing and grading is to allow diagnosis and re-learning.

The single course structure of the current undergraduate curricula can be a barrier to high quality learning. Courses are currently owned by individual faculty members. Meaningful learning outcomes may cross the boundaries of many courses. The instructor in the course in engineering economic analysis encouraged students to choose self-initiated projects that would apply techniques to problems in other courses. It is possible to cross course boundaries, but I assert that this is the exception rather than the rule. Drexel University’s success with “multi-course” design and team teaching suggests that there are workable alternatives to the single-course structure.

The instructor-centered culture of the university environment discourages self-directed learning. Brookfield (1988) states that faculty training and time, institutional structural
factors, and the tension between institutional mandates and individual control are barriers to self-directed learning. Kohn’s (1993) asserts that traditional grading creates an environment in which students focuses on performance to the detriment of learning. He further states that this focus on performance discourages risk taking and creativity. Self-directed learning requires risk taking and creativity. I would also suggest that the lack of a sustained instructor-student relationship is a deterrent to self-directed learning. When the student has gained enough knowledge to initiate a learning project, insufficient time may be available to pursue it in an individual course. The relationship with the instructor also typically ends, so there is not an ongoing mentoring relationship between student and teacher.

4. **Five Fundamental Flaws of Engineering Education**

The major findings reviewed in the previous section suggest five fundamental flaws in the education of engineers. The term flaw is used because each is imbedded within the system of education. I propose that to make a breakthrough improvement in the quality of learning, we must inspect our educational system with a new lens. It will not be enough to train faculty to be better teachers, to upgrade facilities, or to redesign the curriculum. I propose that the recognition of the flaws is a major step in creating a new vision.

**Flaw One - The Absence of Context.** The attempt to separate the acquisition of knowledge from the context in which it can be understood and transformed is a fundamental flaw. At best, traditional education provides a latent seed that germinates at some later time. Higher education can do better. When knowledge is experienced it
becomes transformed and the learner becomes transformed. It is in the transformation that
the learner experiences the joy of learning. Tribus' (1992) emphasis on know-how requires
a context. Drexel University’s (1992) curriculum re-design has freshman engineering
students engaged in a design project in their first course. Knowles (1980) asserts that
adults are problem-centered in their learning and have an immediacy of application in their
learning orientation. The teacher needs to provide leadership in the learning process by
asking, “How can I help create a learning experience in which the learner can experience
this knowledge within a meaningful context?” Evaluation methods that are based on
context (for example, learning contracts and portfolios) offer an opportunity for more
authentic representations of learning.

Flaw Two - The Absence of Integration. The practice of organizing education
around individual courses is a reductionist approach to learning. It is built for the
convenience of the teacher, grade processing, and administrative policies, not the student.
The student’s learning path is disjointed. Each course is a new experience, unrelated to
the last one. The student wonders, “How does it all fit together?” “Why am I always
starting over?” “So what?” A pile of bricks is a pile of bricks. The student carries the
load of knowledge bricks to his or her first job and then education begins. It does work.
It is not even bad. But, how much better prepared could learners be if they were taught
how to lay the bricks and maybe even build a cathedral. Education must be built around
competencies, not courses. Langford and Tribus (1993) suggest the development of
competency matrices which identify meaningful learning objectives. These matrices
provides a basis for students to see the whole as they gain knowledge and know-how from
the individual pieces. The teacher and the institution need to ask, "How can we help
students integrate their learning experience so that it becomes seamless and whole?"

**Flaw Three - The Lack of Mentoring.** In years past an aspiring artisan would enter
into a period of apprenticeship with a master craftsman. It was the master’s mission to
help the young apprentice learn and grow to become a master craftsman. The mentor is a
special person to the learner. The mentor cares about the success of his or her learner.
The mentor guides, advises, counsels and challenges. The undergraduate learning
experience is largely a solitary and impersonal one. Most undergraduate students do not
have the opportunity to establish a relationship with a faculty member that is truly a
mentoring one. Brookfield (1988) and Knowles (1980) emphasizes the importance of the
relationship between the learner and the helper. The relationship is a personal one which
is warm and caring. If the teacher and the student are co-managers of the learning
process, there needs to be some consistency and constancy in the relationship. What
would happen if every undergraduate student had a mentor that would guide and
encourage the student to grow to be a high quality member of his or her profession?

**Flaw Four - The Acceptance of Mediocrity.** Teachers want their students to learn,
but they often inadvertently send a different message. Teachers give tests that are
superficial. Students prepare for tests instead of pursuing knowledge. Teachers accept
performance that is not high quality. The concern for cheating says more about the
learning process than it does about the character of the student. Tribus (1990) states that
the learning quality standard at Mt. Edgecumbe High School is “perfect”. Glasser (1990) recommends that only grades of B or A should be accepted. This research study suggests that the establishment of high performance standards results in high performance.

Instead of asking students to answer (actually remember somebody else’s answer) questions, the teacher should challenge and encourage them to demonstrate their understanding in a way that is appropriate to each of them. For example, instead of testing their recall by answering the teacher’s questions of a management theory, perhaps students could write a case about a personal work or life experience and analyze it in relation to the management theory. The analysis could include a reflection piece to help the student gain personal insight into his or her own belief system. I suspect that many teachers would believe this is beyond the capability of the undergraduate student. But we get what we expect, and when students are challenged and encouraged, quality emerges. As we challenge students to deep-level learning, grades seem almost trivial. The concern for grade inflation is misguided. Teachers should be more concerned about learning deflation. What if every demonstration of knowledge evoked understanding and insight by the learner?

Flaw Five - The Avoidance of Character Development. Tribus’ (1992) four components to quality education, particularly character development, are an awesome challenge. Knowledge is not value free. Know-how, the application of knowledge, always raises issues of ethics and values. Wisdom requires a deep understanding of subject and self. And there is a spiritual quality to character development. A few courses
in the humanities won’t cut it. Each of these components must be experienced in the totality of the learning experience. We all can remember that one teacher (maybe more) in which we learned more than the content of the course. Something in the person or in the experience touched us deeply. What would the learning process be like if that once in a lifetime experience was the norm? What would be the impact on the learner and to society?

Are these flaws so predominant? Is it that bad? Maybe not. There are enlightened and dedicated teachers. There are resourceful students. And sometimes things fall together. But to recognize a flaw is to recognize a vision. There is power in vision. In the next section I present a vision of a high quality learning experience. The context, an Industrial and Systems Engineering Department, helps me communicate a vision that has emerged through the creation of this dissertation.

5. Reengineering the Engineering Learning Process

Joel Barker (1990), futurist and author of The Business of Paradigms, often uses the phrase, “When a paradigm shifts, everyone goes back to zero.” To help the paradigm shift, we need to suspend our vision of the past. To suspend is not to discard, but to clear a space so that a new vision can be enacted. In this section I will present a vision for the education of a high quality engineer. This vision is created by our ability to recognize the flaws reviewed in the previous section. My purpose is to be a positive force in the transformation of how higher education functions in creating high quality learning for its students, faculty and staff.
High quality learning postulates that all students are valuable, educable, want to learn and are willing to take responsibility for their learning. Tribus (1994) visualizes every classroom as part of a large learning community. I agree and in addition assert that the community is also the classroom. Following are some of the components for the creation of a high quality learning system.

Create a learning experience among students, faculty and staff to learn and engage in dialogue about high quality learning. Many of the readings provide a foundation to explore together so that a vision for the future in the Industrial and Systems Engineering Department can be co-created. An outside facilitator could help the department design and implement a process among the learning and dialogue team. Through reading, self-reflection, and dialogue a shared vision would be created and a commitment would be made to transform the learning process within the department.

Create a comprehensive competency profile of a high quality industrial and systems engineer. This profile would be the driving force for the new curriculum design for the department. The competency profile would be developed by a panel of exemplars who have an in-depth (collectively) understanding of what a high quality industrial and systems engineer knows and does. The panel would include people from the academic and business community, practicing engineers and knowledgeable customers. The process of selecting the panel is important, so that the product of the panel can be of high quality. The competency profile would be used to develop learning competency matrices that would form the basis for reengineering the design of the overall learning system for the
department. As recommended by Langford and Tribus (1993), the learning competency matrices would include the components of knowledge, know-how and wisdom.

Use learning competency matrices to help guide the reengineering of the curriculum design. What does future perfect look like? What resources does the department have that will support the redesign? What are the gaps? The outcome of this process is not just to change the curriculum but to change the vision of how faculty, students and legitimate stakeholders will know when the student has achieved competence as an industrial and systems engineer. The basic building block of the student’s education is not the course. The basic building block is demonstrated competence of the learning outcomes identified in the profile. The goal is to provide a process in which all students can achieve a high level of competence as an engineer and as members of society. The learning process would also help students develop skills for self-directed and cooperative learning. A variety of learning and teaching strategies would be designed to enable students achieve competency and reach their personal goals. These would likely include traditional classes, self-instructional modules, cooperative learning, and projects.

Match each student with a faculty mentor to guide him or her through the process of becoming a high quality engineer. The mentor is the co-manager of the learning process with the students. The mentor helps the student create a learning plan, assess his or her progress and develop high quality learning skills. This is both a professional and a personal relationship. Throughout the student’s learning journey, he or she will interact with all members of the faculty team, but the mentor is the guide and
anchor for the student. Additional mentor roles are explained as the other components of the learning system are reviewed.

Use portfolio assessment as the primary mechanism for the demonstration of student competency. The mentor will help the student build his or her portfolio. The portfolio will be reviewed by other faculty members and business professionals. The students will be given feedback so that they can achieve high quality performance. Testing and other diagnostic tools will be built into the various learning strategies. They will not be used for grading purposes, but for helping the student and the faculty assess the learning and teaching processes. When the faculty is in the role of content expert, she or he is able to concentrate in that role. The students are participating in a class to learn. It is the students’ responsibility to demonstrate competence. Faculty members will be involved in reviewing portfolios, but not in the context of their teaching role. Because the student demonstrates competency by using portfolio assessment, it is possible that he or she may select a learning strategy that does not include a traditional class. For example, a student may be able to develop his or her competency for a particular area through self-directed learning projects and co-op assignments. Although high freedom of student choice is encouraged, there may be times when the mentor may advise the student to pursue a more structured learning activity over a self-structured one. It is the role of the mentor to challenge the student. The healthy student-mentor relationship will not avoid creative conflict.
Establish a continuous co-op work assignment for each student. The co-op assignment will be continuous in that the student will be working and pursuing their learning plan at the same time. There will be a partnering relationship between the department and local businesses to establish meaningful work experiences for the students. Work assignments and learning assignments will be integrated. The faculty member will work hand in hand with his or her business colleague to create a win-win relationship. They will meet jointly with the student on a regular basis to ensure progress is being made. As in other learning activities the portfolio will be the mechanism to evaluate the students' learning. The benefit to the student in the co-op assignment is the ability to apply skills to real world problems in real time. It would also be a source of income that would not be a trade off to the student’s college experience. The benefit to the faculty member would be an increasing level of understanding of business expectations for engineers. This would help keep the teaching relevant to what the student needs to know. The relationship between faculty and business could also be supportive of the research and service mission of the university. The benefit to the business would be an opportunity to improve their processes through the work of the student engineer and the expertise of the faculty. Enlightened business people are interested in supporting high quality learning. The partnership offers opportunity for synergy.

Teaching is an important process in this high quality learning system, but it is more flexible and responsible to the needs of the students. A class, for example, does not have to be ten weeks long. It may be spread over a year or completed in a month. It may
include one or team of faculty members. Students earn credits, not by completing classes, but by achieving the objectives in their learning plan. Teaching is not a solitary process. Faculty members team up so that their areas of expertise can be integrated. One learning project can span two or more learning modules. The learner may complete a learning module in engineering economics as part of a design project in ergonomics.

6. Building a Support System for Students and Faculty

This high quality learning system will challenge the students’ ability to take responsibility for their own learning. I envision an Introduction to High Quality Learning "class" to help the students develop the skills necessary to succeed. Some learning modules may include:

- The competencies of a high quality engineer
- Building a learning plan and portfolio
- Co-managing your learning process with the faculty team
- Working and learning with your mentor
- How to journal - its benefits to you professionally and personally
- Exercising self-direction in learning
- Cooperative learning
- Coping with the traditional learning environment
- Learning and working with your co-op assignments

This preparation be done in a variety of ways -- workshops, one-on-one coaching, dialogue groups, and self-instructional materials. This preparation should be done in a
way that the students can experience high quality learning as they learn about high quality learning.

Faculty will be challenged in this vision of a high quality learning system. In addition to their professional expertise they will be asked to take on new roles. The faculty’s overall teaching mission is to play a leadership role in the co-management of the learning process with the students. Some of the skills that this will require include:

- Mentoring the student towards high quality learning and self-direction.
- Working with the business team member in the co-op experience.
- Using portfolios for learning and certification of competency.
- Team teaching with other faculty members.
- Designing and using experiential learning strategies.

Although the literature indicates that a learner-centered environment is highly satisfying to the teacher, it also presents a challenge. The teacher does not have absolute control over the process. She or he is accountable to the students and to the business community. Current “tried and true” teaching modules may require significant revision in light of the learning competencies identified. The skills, some of them listed above, will be new to many faculty members. It will take time and training to build these skills. A higher education department, like Industrial and Systems Engineering, could use the full time services of an educational professional (instructional designer, adult educator) to provide the support and tools for developing high quality learning processes. Most importantly the
enactment of this vision will require a new way of thinking about learning and the will to bring it to life.

This vision is not a detailed road map towards a high quality learning system. It points a direction, however, for a fundamental shift in higher education. Tribus (1994) challenges the urgency of beginning the transformation by saying,

When asked “When should I begin?”, Dr. Deming has been heard to say, “It doesn’t matter, so long as you start at once.” (p. 21)


40. Tribus, Myron. TQM in Education The Theory and How To Put It To Work. EDU ERIC Microfiche ED370168 [1993].


APPENDIX

ATTRIBUTE CONTROL CHARTS
PERCENT AGREE

OLD TESTING & GRADING SYSTEM

FIRST REVISION

SECOND REVISION

UCL

Mean

LCL

LEGEND
Quarter 1 - Autumn 1992  Quarter 5C - Spring 1994 (Control)
Quarter 2 - Winter  1993  Quarter 5E - Spring 1994 (Experimental)
Quarter 3 - Autumn 1993  Quarter 6 - Autumn 1994
Quarter 4 - Winter  1994  Quarter 7 - Winter  1995

FIGURE 18 ATTRIBUTE CONTROL CHART
S.E.T. QUESTION 18
PERCENT AGREE

FIRST REVISION
SECOND REVISION
OLD TESTING & GRADING SYSTEM

UCL
Mean
LCL

LEGEND
Quarter 1 - Autumn 1992  Quarter 5C - Spring 1994 (Control)
Quarter 2 - Winter 1993  Quarter 5E - Spring 1994 (Experimental)
Quarter 3 - Autumn 1993  Quarter 6 - Autumn 1994
Quarter 4 - Winter 1994  Quarter 7 - Winter 1995

FIGURE 19 ATTRIBUTE CONTROL CHART
S.E.T. QUESTION 19
FIGURE 20 ATTRIBUTE CONTROL CHART
S.E.T. QUESTION 20

LEGEND
Quarter 1 - Autumn 1992  Quarter 5C - Spring 1994 (Control)
Quarter 2 - Winter 1993  Quarter 5E - Spring 1994 (Experimental)
Quarter 3 - Autumn 1993  Quarter 6 - Autumn 1994
Quarter 4 - Winter 1994  Quarter 7 - Winter 1995

PERCENT AGREE
150
317
FIRST REVISION
OLD TESTING & GRADING SYSTEM
100
90
80
70
60
50
40
30
20
10
0
UCL
Mean
LCL
QUARTER
0 1 2 3 4 5 6 7
SECOND REVISION

Quarter 2 - Winter 1993 Quarter 3 - Autumn 1993 Quarter 4 - Winter 1994 Quarter 5C - Spring 1994 (Control) Quarter 5E - Spring 1994 (Experimental) Quarter 6 - Autumn 1994 Quarter 7 - Winter 1995
FIGURE 21 ATTRIBUTE CONTROL CHART
S.E.T. QUESTION 6
FIGURE 22  ATTRIBUTE CONTROL CHART
S.E.T. QUESTION 7
LEGEND
Quarter 1 - Autumn 1992  Quarter 5C - Spring 1994 (Control)
Quarter 2 - Winter 1993  Quarter 5E - Spring 1994 (Experimental)
Quarter 3 - Autumn 1993  Quarter 6 - Autumn 1994
Quarter 4 - Winter 1994  Quarter 7 - Winter 1995

FIGURE 23 ATTRIBUTE CONTROL CHART
S.E.T. QUESTION 8