99B001I

99B001I

Project Number: JEM-1082 -51

ANALYSIS AND RECONSTRUCTION OF THE WPI COURSE EVALUATION

An Interactive Qualifying Project Report

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Robert G. Holmes **Matthew S. Gaffney**

Date:
October 26, 1999

Approved: Judith E. Miller

Table of Contents

1.0 INTRODUCTION 4
2.0 LITERATURE REVIEW 7
2.1 What is important to the student? 7
2.2 Characteristics of effective teaching 8
2.3 Important aspects of good class design and good teaching 9
2.4 Factors affecting course evaluation responses 9
2.4.1 Seriousness of the student evaluation 10
2.4.2 Class size 11
2.4.3 Scale 11
2.4.4 Paying back 11
2.4.5 Academic field 12
2.5 Conclusion 12
3.0 OBJECTIVES 13
4.0 METHODOLOGY 14
5.0 RESULTS 17
5.1 The Original Survey 17
5.1.1 Part 1- Individual Questions 17
5.1.1.1 Part 1- Means of Responses 18
5.1.2 Part 2- Category Ranking 19
6.0 DISCUSSION 22
6.1 Students Responses to the Surveys: 22
6.2 Correlation Matrices 23
6.3 Examining the Means for each Question: 24
6.3 Summary of important results: 27
7.0 CONCLUSIONS 28

ABSTRACT

A course evaluation and it's analysis provides information that is very useful to students, to instructors, and to the institution that employs it. Currently, Worcester Polytechnic Institute utilizes a course evaluation for the purpose of recording the performance of instructors, providing information to students, tenure review and future course design. It was the purpose of this Interactive Qualifying Project to improve the current evaluation form, and the process by which it is analyzed, with the students' opinions in mind. This was done via a questionnaire that was sent to a sample of the student population. The questionnaire inquired about the importance of various aspects of course design and the characteristics of the instructor in relation to the satisfaction of the student and to the student's perception of overall amount of learning that takes place. The results of these questionnaires provide a basis for assessing the relative importance of each question on the proposed form. With the students' opinions at hand, we chose to assign a weight to each question and use it in the compilation of the new evaluations form. Data gathered from our surveys, and information concluded on by previous researchers will aid in the development of a new course evaluation form for WPI.

1.0 INTRODUCTION

College students at Worcester Polytechnic Institute and across the nation express concern for the validity of teacher evaluations taken at the end of a given academic term. Many students believe that over the years, several of the questions on the current evaluation have become outdated. A lack of confidence in the importance of these evaluation forms has led to students not taking them as seriously as they should. In turn, a feeling of powerlessness may arise on the part of the student, and they may not put reasonable effort into filling out the form. Students and faculty alike feel that changes need to be made to WPI's current student evaluation in order for it to be easily analyzed. When making such changes, the new evaluation should also be molded to fit the current environment.

At WPI, each student can be represented as a client of a large corporation: WPI. The reason that most students attend WPI, and any university for that matter, is to gain an education, and to prepare them for life after formal education. As no university is free of charge, it is the obligation of the university, the corporation, to keep the student, the client, content. Course evaluations can be a very successful method of ensuring such satisfaction, provided they are used effectively. It is because of these ideas that it only seems right that both the students and the faculty take part in the creation of the course evaluations. This method of producing an evaluation will allow the students to add their input as to how they would like a class to be run, and the faculty will be able to complement this with knowledge from experience.

It is from the above thoughts and ideas that a main goal was extracted. This goal was to create a new course evaluation that will reflect the interests of not only the

individual department, but also the student, and moreover, the entire institute. The proposed evaluation was constructed in collaboration with both faculty and student representatives. Student surveys will obtain opinions on the importance of various questions on the current evaluation, ample space to input new questions, and ideas that may have a place on a new course evaluation. A proposed course evaluation form was developed from the responses from the student surveys, and from the current evaluation form. As the new evaluation neared its completion, each question on the form was given a category and assigned a weighted rank of importance for future analysis.

The information drawn from the students on a course evaluation should be useful and unambiguous. The new format of evaluation will allow accurate conclusions to be drawn about the instructor's performance as well as the overall course format. The overall process of data acquisition from the forms will remain quite similar. Like the current process of course evaluation, the proposed course evaluation should be distributed at the end of the course by the instructor, and collected and deposited by a student. The proposed evaluations will gather more relevant and usable data from the students than the current evaluations, and, like the current evaluation form, will be analyzed and summarized on the WPI web site. The difference between the current system that WPI employs, and that which we propose, is the ranking of each question on the new course evaluation. The evaluation is broken down into categories, giving each question an appropriate weight. The summary of evaluations for a given course/professor combination will take into account the importance of each question, in each category, rather than weighting all questions with equal importance.

It is the purpose of the proposed course evaluation to assess the ratings of a given course and/or professor accurately. The evaluation must depict the attitudes of the students towards the class/professor accurately in order to be useful to prospective students of that class/professor. Such an evaluation, developed by both students and faculty, will give students a voice in the current system. If students are involved with such issues, and feel that their opinions are valued, their participation in evaluations will also be more motivated, and will better allow the student to value exceptional teaching.

2.0 LITERATURE REVIEW

For years student evaluations have been a resource for determining the effectiveness of teaching here at WPI, and across the nation. Evaluations serve as a link between the students and their professors, transmitting both positive and negative feedback. This communication between student and professor can measure the professor's performance, and the effectiveness of their teaching skills. Using this information can serve as a institutional standard for tenure, promotion, and improvement. However, evaluations can be ineffective if the information is not drawn from them correctly (Edel, 1989).

2.1 What is important to the student?

While the multiple choice questions on a given evaluation form can be easier to analyze than open ended questions, the open ended questions can provide useful information. Although it is much more difficult to analyze such comments, they are still a valuable source of information. A 1978 analysis of the student comments done by Braskamp, Ory, and Pieper (1978), focused on the responses to various written questions. The study found that, among open ended type questions, comments regarding the instructor's ability and knowledge were most frequent, with those on personality and rapport following closely. Comments made by the students regarding the course itself consisted mostly of remarks on content, material and grading. The study also showed that out of four student responses to general essay type questions regarding the class, half of all comments concerned the instructor's personal characteristics, mainly communication, and a third of every comment concerned the course itself. The high frequency of

comments regarding the characteristics of the instructor shows how important it is for the instructor to communicate clearly. After clarity, knowledge of the subject and interpersonal skills followed. The responses acquired from the Braskamp, Ory, and Pieper survey were sorted and categorized. The most frequent remarks made by students were categorized by frequency. These categories include clarity, knowledge, interpersonal rapport, organization, difficulty/workload and grading.

2.2 Characteristics of effective teaching

One may ask, "What characteristics would best define exemplary teaching?" There are many different aspects of teaching that can define an excellent teacher. For each different teaching style, there are many different techniques, and they vary from instructor to instructor. "Words such as 'stimulating,' 'dynamic,' 'enthusiastic,' 'caring,' 'motivating,' and 'knowledgeable,' are proposed first, followed by elaborations on images of individuals who are both engaging and capable of motivating and supporting student learning" (Svinicki and Menges, 1988).

Lowman, 1984 developed a theory on teaching effectiveness that categorizes most popular characteristics of excellent teaching into two categories. These two divides, intellectual excitement and interpersonal rapport, are each then broken down into several different attributes. Instructors that are engaging, creative and exciting in their classrooms best represent the intellectually exciting professor, while those excelling in interpersonal rapport will succeed in communicating with the students on a more personal level. Such teachers often express concern for the student's well being, and are able to convince them of this concern (Lowman, 1984). Such skills need to be evaluated on a course evaluation.

These skills should also be assigned weights in order to ensure an accurate measurement of the course as defined by the student.

2.3 Important aspects of good class design and good teaching

Further investigation was done by a separate researcher in 1996. Svinicki and Menges utilized more surveys and written comment analysis to develop their own set of categorized teaching attributes. This study showed that most critiques and evaluations of teachers focus on four main categories. The first category deals with how clear the instructor made his/her class, the second with how interesting the classes were. Thirdly, the classes had to be positive, and finally motivating. Through student surveys it has been determined that categories 1 and 2, clarity and interest, are the dominant characteristics of good teaching. However, it is also evident that most students would prefer an instructor to also be highly effective on a more personal level rather than one who solely communicates that which he/she is teaching (Svinicki and Menges, 1996).

2.4 Factors affecting course evaluation responses

A student evaluation can serve a useful purpose only if its questions are thoughtful and also answered in a thoughtful manner. Some professors discredit the value of student evaluations, claiming that many of them are filled out in an improper manner. However, recent studies have shown that non-biased evaluations can be one of the most important resources for determining teaching excellence. These factors will be examined individually in the paragraphs to follow.

However, recent studies have shown that non-biased evaluations can be one of the most important resources for determining teaching excellence. These factors will be examined individually in the paragraphs to follow.

repeat

2.4.1 Seriousness of the student evaluation

A major factor in the validity of student evaluations is how seriously the student takes the form to begin with. Deans and faculty across the nation discredit the results of all student evaluations (Miller, 1972). They claim that freshmen are incapable of seriously completing a form because they are in new surroundings and are not familiar with critiquing professors. Freshmen have no previous experience in a college atmosphere, and the style of teaching in college is usually very different from that of their high schools; therefore, the students may not know how to approach it. At the other end of the spectrum are the upperclassmen. Faculty claimed that even though the upperclassmen have better understanding of the importance of the evaluations and how to properly fill them out, they become bored with them and hastily fill them out. Sophomores, juniors, and seniors have all seen the same forms for many classes before and treat them as a hassle, not taking the proper time to give accurate answers (Miller, 1972). An evaluation which is easily understood by the evaluator, and brief enough, could at least aid even the most misguided freshmen towards a better evaluation. An example of this is how the questions are grouped together in relevant topics forming subgroups, and each question uses common wording so its easily understood. This sort of evaluation would also be short enough for the busiest upperclassmen to take time for.

2.4.2 Class size

Class size is another variable that can affect the evaluations of professors. The ability of the professor to communicate to the whole class is directly linked to the number of students attending lectures, conferences, or labs. Research has shown that professors receive higher ratings from small classes, compared to large lectures. This may be due to the better individual attention received by students in a small class (Singh, 1998). Therefore the professor teaching a lecture will not score as high since he cannot meet the demand for the students' attention. In small classes better teacher/student relations can be formed and the students tend to rate these professors higher (Cashin, 1988).

2.4.3 Scale

Here at WPI there is a four-point scale for each question, meaning that the student has a choice of five answers: N/A, strongly disagree, disagree, agree, and strongly agree. The scale of the questions is one of the most important variables. It is known that this type of scale is biased toward an answer of "agree". This is because if the student doesn't have any strong opinion, either positive or negative, they tend to answer "agree." Scales of more choices (5,7), and with a neutral position, offer a wider range of answers, and are much more accurate in their portrayal of the student opinions (Miller, 1972).

2.4.4 Paying back

Helping a professor out with a good evaluation for a good grade is another variable that needs to be considered. The old saying "You pat my back, and I'll pat yours" best describes it. Students are more willing to give a good evaluation to a professor they know they're getting a good grade from (Aubrecht, 1988). Cognitive

psychology also teaches us that if students perform poorly in courses they're more likely to blame the teacher rather than take responsibility for their low effort (Doyle, personal communication).

2.4.5 Academic field

Academic fields also are yet another factor that can alter the performance evaluation of a teacher. Professors in humanities, language, and other liberal arts courses score higher than professors in more technical courses. Students also tend to give higher ratings to a professor that is teaching a class that is related to or within the student's major (Cashin, 1988).

2.5 Conclusion

What changes should be made to the existing course evaluation in order to improve it? Questions should be clear and concise, and be organized into subgroups which all pertain to the same category. The main purpose of a course/teaching evaluation is to identify categories such as organization, communication, difficulty/workload, grading, and rapport, which are important attributes of exemplary teaching. Our research has shown that the students feel certain categories have more importance to them than other categories. Since this is true, these categories receive weights according to their importance.

3.0 OBJECTIVES

The main objective of the project is to develop a new course evaluation form. The goal of the new course evaluation is to accurately measure those characteristics that students feel are most important in the classroom. The next objective is to develop a mechanism to compile the evaluation data in such a manner so that the data received from the students can be used in the most effective manner possible. Our project will utilize student surveys, and develop a weighting scheme in order to achieve a final student evaluation. We will also use statistical analysis software to help us to interpret the data collected by the survey.

4.0 METHODOLOGY

The first step in the development of a new evaluation was to research the current evaluation processes at WPI. There are several ways in which this was done. The first was to read over past IQP's, and develop an understanding of how the evaluation process was researched before. From the past research we determined how the new evaluation must differ from the old, and how it can be made more effective. Second, sources of information on all related topics were searched for, and used for as much pertinent information as possible. This included talking with professors, and reading books related to developing surveys. The main goal of the background research was to determine what previous studies have said about students' opinions of course evaluations. To complete the project effectively, it was also necessary to research survey and interview design, and processes to propose a new course evaluation.

Throughout the research process, a majority of authors spoke of five categories that summed what students believe are most important in a course. These categories are organization, communication, rapport, difficulty/workload, and grading. It was determined early on that these five categories are going to be the basis for the new course evaluation, and that any student surveys or faculty interviews will at minimum touch upon each of these categories.

A survey was developed to distribute to a cross-section of the student body (Appendix A). The purpose of the survey was to gain an understanding of the general opinions of the student body on the importance of various elements of course structure

and instructor behavior. The surveys were developed using tips and strategies from Salant and Dillman's How to Conduct Your Own Survey (1994). Some important tips that were acquired from the text pertained to the wording of the questions, and how the questions were asked. For example, each question must only have one question in it. It would be wrong to ask a question like, "Did the professor's personality seem concerned about the students, and did he stimulate your interest in learning?" This question has two parts to it and cannot properly be answered with a simple "yes" or "no." If this question were to be used in a survey it must be split up into its separate components. Questions that are answered with numerical responses, "yes" or "no," or True or False are the easiest to form conclusions from. They do not contain any gray areas or open-ended responses which are hard to interpret.

Since the return rate of surveys is low when placed in campus mailboxes, they were distributed during D term, 1999. With the instructors' permission, surveys were distributed with the usual course evaluations. The survey asked questions that determined what attributes in the evaluation are most important to the students. After the surveys were distributed and collected, each was reviewed, and the results tallied. A summary of the totals for each question was developed and the final results were recorded.

The results from surveys and interviews were then subjected to basic statistical analyses to determine what course characteristics were most important to the students. Each evaluation was entered into a spreadsheet in SAS. One spreadsheet for "Your Learning," and another for "Overall Course Satisfaction." The 1-5 integer answers for each question were used to form frequency histograms, and correlation matrices. The Histograms showed the frequency of responses for each question, and the correlation

matrices were used to determine if there were any correlation between student answers for "Overall Course Satisfaction" and "Your Learning" columns.

The results from the analyses were then used to develop a more suitable student evaluation. Rather than starting completely from scratch, the current course evaluation was used as a foundation to build on. Since the current evaluation was thoroughly thought over in development, and has many good ideas, the research done in this project looked for new things that needed to be added to the current evaluation, and clearer grouping of questions to form categories.

Once a new form was developed, a method of using the data was produced. Data analysis was presented in more meaningful and effective manners. The evaluation was split into categories (including the five mentioned previously) and each different category holds a varying weight by opinion of students. These weights are conducted like many professors weight course homework to the final exam.

5.0 RESULTS

After observing the results of our student surveys, and performing some preliminary statistical analyses on them, it appears that the overall number of student responses are skewed towards 1, (very important,) and 2, (fairly important) responses. This generalization is made in regards to the entire number of responses gathered from all methods of data acquisition. In short, it seems that most students believe that all of the questions on the questionnaire are very important. The majority of the responses for all of the questions were 1's and 2's, and make it difficult to make accurate judgments about the sentiments of the students in the sample about the relative importance of organization, communication, rapport, difficulty/workload, and grading.

5.1 The Original Survey

5.1.1 Part 1- Individual Questions

The first process in the search for data began with a simple survey that was distributed, and collected from 106 students. Seen in appendix A, the survey asked questions on the importance of various teaching characteristics. The survey was divided into two parts, individual questions regarding various aspects of course design and teaching style, that were answered via a multiple choice scale, and the ranking of the 5 categories of teaching characteristics listed previously.

After collecting the completed surveys, our first step in analysis was to calculate the frequency of the responses of the multiple choice questions. This was done using simple frequency histograms. Using SAS, we found the frequency of occurrence of each possible numerical response over the range of 1-5, 1 being most important, and 5 being least, for each question on the questionnaire. This procedure was done for both the "Overall course

satisfaction" responses, and for the "Your learning" responses. The frequency histograms can be seen in **Appendix B**. The frequency histograms provide a good graphical representation of the number of responses, however they are ambiguous due to the large number of similar responses. This makes it difficult to draw conclusions about which particular characteristic of teaching or class design is most important to the student.

5.1.1.1 Part 1- Means of Responses

The next step in analyzing our results was to determine the mean of the responses for each of the individual questions. The data returned from the 106 surveys collected was in the form of a number from 1-5. In this case, a 1 response represented a student's strong sentiment of importance, while 5 indicated that the student felt that the particular characteristic was less important. This method gave us the average response to each question, and was more useful in comparing questions that are in the same category (organization, communication, etc.) The mean response for each question can be seen below.

Question	Mean for "Overall Course Sat." Category	Standard Deviation	Mean for "Your Learning" Category	Standard Deviation
The instructor arrives well prepared for lectures.	1.8	1.0	2.0	1.2
The class objectives were defined.	2.0	1.1	2.0	1.1
The instructor used class time efficiently.	2.0	1.1	1.9	1.0
The instructor spoke clearly and comprehensibly.	1.8	1.0	1.7	1.0
The instructor gave good explanations of the material.	1.9	1.1	1.8	1.0
The instructor shared his/her class agenda.	2.4	1.0	2.3	1.1
The instructor seemed really concerned about the students.	2.2	1.1	2.1	1.1
The instructor showed enthusiasm for the subject.	2.3	1.1	2.0	1.1
The instructor was sensitive with the students, and was willing to work with them.	2.2	1.1	2.0	1.1
The instructor made difficult topics easier to understand.	1.7	1.1	1.9	1.0
The homework/exams were good measures of the material.	2.0	1.1	1.9	1.0
The workload was appropriate for the course.	1.5	1.0	1.9	0.9
The instructor graded fairly and consistently.	2.0	1.2	1.8	1.0
The textbook was helpful in this class.	2.2	1.0	1.9	1.0
The instructor was knowledgeable in the subject matter.	1.9	1.0	1.7	1.0
The class room was acceptable.	2.8	1.1	2.6	1.2

Table 1: WPI Students' Perceptions of Importance of Course Attributes to Satisfaction and Learning

5.1.2 Part 2- Category Ranking

The second portion of the student survey inquires about the relative importance of the five characteristics of teaching we identified. The students were asked to rank the five categories listed previously, relative to each other, in order of importance (1-5, 5 being most important.) This information will assist in determining the weights of each of the questions on the revised evaluation form.

During the analysis of the collected data, a source of error was found in the survey. The scale provided for rank ordering the course characteristic categories on the survey was completely the opposite of that provided for the individual questions in part 1. Although the instructions on the survey state the proper scales for each set of questions, a possibility existed that some students did not completely read the instructions given on the survey. If this happened, then some students' answers to the last question on the survey may have been the opposite of what the student intended. There was no way of determining which surveys were correct, so a second survey was needed to decided whether the data was tainted or not.

The same survey was distributed to a convenient sample of thirty students, and each was asked to fill out the survey in our presence. As the student finished filling out the survey, we then reminded them that the second scale was opposite to the first. The students were asked if they had correctly filled out each part of the survey or not, and the results are as follows. Out of thirty students, eight had filled out the survey incorrectly. The only mistake found in all 30 surveys, was that these eight students used the same scale for the category ranking as they did for the first part. Unfortunately, this proves that our original data for the second part of the original survey, the category ranking, was tainted. Although the data found in the second part of the original survey was incorrect, the data collected in the second survey was usable. The twenty-two students that did fill out the survey correctly provided useful information, and the responses on the other eight surveys were reversed to provide more useful data. This data is displayed via the mean response for each category below:

	Organization	Communication	Rapport	Diff/Work	Grading
Mean	2.5	2.4	3.2	3.5	3.3

Table 2: A Corrected Version of the Category Ranking of the Original Survey-Students' View of the Importance of Various Characteristics of Teaching.

6.0 DISCUSSION

6.1 Students Responses to the Surveys:

Unfortunately the students feel that everything is important to them. By the histograms in Appendix B alone it would be hard to come to any conclusions on the data. However, the means of each question yield a more detailed result by being able to compare two different means for each question. Although the difference in the means is small, conclusions can be made from them. If there were a more substantial difference between the statistical means we would be more confident in our interpretations.

However, by looking at the means for each answer for both "Course Satisfaction" and "Your Learning" it is clear that with the questions we asked in the student's view personal learning was more important to them than overall course satisfaction. This is evident by taking the column total average from each side and comparing the two. "Your Learning" has an overall average of 1.86, and "Course Satisfaction" has an average of 2.05 (**Table 3**). We feel that the questions are not biased toward "Your Learning" consequently this suggests that "Your Learning" received lower scores because the students felt that their personal learning is more important. This can be seen in **Table 3** on the next page.

Means for individual questions:	Means for "Overall Course Satisfaction"	Means for "Your Learning"
Organization: Category Average	**1.95**	**1.94**
The instructor arrives well prepared for lectures.	1.78	1.99
The class objectives were defined.	2.04	1.97
The instructor used class time efficiently.	2.02	1.87
Communication: Category Average	**2.02**	**1.92**
The instructor spoke clearly and comprehensibly.	1.81	1.67
The instructor gave good explanations of the material.	1.87	1.76
The instructor shared his/her class agenda.	2.38	2.34
Rapport: Category Average	**2.22**	**1.99**
The instructor seemed really concerned about the students.	2.17	2.08
The instructor showed enthusiasm for the subject.	2.33	1.95
The instructor was sensitive with the students, and was willing to work with them.	2.16	1.95
Difficulty/Workload: Category Average	**1.77**	**1.90**
The instructor made difficult topics easier to understand.	1.75	1.94
The homework/exams were good measures of the material.	2.01	1.91
The workload was appropriate for the course.	1.55	1.87
Grading:	**2.04**	**1.76**
The instructor graded fairly and consistently.	2.04	1.76
Other: Category Average	**2.31**	**2.10**
The textbook was helpful in this class.	2.25	1.94
The instructor was knowledgeable in the subject matter.	1.87	1.71
The class room was acceptable.	2.82	2.64
COLUMN TOTAL AVERAGE:	**2.05**	**1.86**

Table 3: Means of Survey Responses Including Category Averages

6.2 Correlation Matrices

One of the questions that needed to be answered was if there were any correlation between how the students answered in "Overall Course Satisfaction," and "Your Learning." Using the statistical analysis program SAS, we conducted correlation matrices between questions in both categories (**Appendix C**). The purpose of the correlation matrix was to compare how the student answered each question for "Overall Course Satisfaction," to their answer to the same question but for the "Your Learning" category. The matrices were easily understood: positive 1 meant that there was a positive

correlation between answers (if the student answered "1" in "Your Learning", they would answer "1" in "Overall Course Satisfaction," negative 1 meant that there was a negative correlation, zero meant that no correlation could be made. Of course you could have many numbers in between negative 1 and positive one, or even less than -1, and greater than +1. When this happens you have to determine if the integer is close enough to -1, 0, or +1 to make any conclusions. In our study every correlation matrix yielded a correlation number that was between -0.06 and 0.16. Professor Petruccelli confirmed that our results were to close enough to 0 that there is no correlation between the answers. This can mean that for the survey in general students answered randomly, instead of having substantial thought, or logic behind their answers. Or their answers for one column were not influenced by their answer for the previous column. This might be an indication of substantial thought for each column. We believe that the second ideology is true.

6.3 Examining the Means for each Question:

Categories with the lowest averages will clearly be some of the most important questions to the students. From the "Your Learning" column in **Table 3** individual categories can be examined. Where there is more than one question in a category the average is in **bold**. Directly from the data set in **Table 3** the order of importance of categories to the students is seen on the next page in **Table 4.**

	"Your Learning"	"Overall Course Satisfaction"
#1	Grading	Difficulty/Workload
#2	Difficulty/Workload	Organization
#3	Communication	Communication
#4	Organization	Grading
#5	Rapport	Rapport
#6	Other	Other

Table 4: Order of Importance for each Column

The orders are almost very similar aside from the position of Grading.

Students feel very strongly about how the professors grade for "Your Learning." We conjecture that this is why it received the lowest overall average for any category. Clearly students are concerned with their final grades, and want to receive a grade that rewards their hard work in the class.

Difficulty of the workload could be linked to grading, which may be why it was the second most important category according to the students. Students may feel that if the workload is too demanding for the 7-week term, they might not get enough out of the class, or have learned anything. Too much material covered in the class results in a broad knowledge of the subject matter, or in some cases "utter confusion." However, it some cases it may be more beneficial to know half of the material in a book extremely well, instead of knowing the whole book in less detail.

Another important category is communication. As students we also agree with the importance of this category. Nothing can be learned by the students if the professors,

TA's, and other assistants cannot verbally express their knowledge of the subject matter. We personally remember many courses that were taught by professors of lacking verbal (English) skills. The result was an audience looking around at each other in a confused state. If professors know that students need to be able to understand them, then accommodations can be made. For example, visual aids to help illustrate theories, graphs, or homework problems could be brought to class to show the situation in simpler terms. "A picture is worth a thousand words" holds true in trying to visualize acceleration vectors, how stress acts on rigid bodies, or the mitochondria within the body's cells.

Organization is a key to a well developed course. Without proper organization a professor might waste time looking for answers, or going over material that isn't necessary for the course. A well organized lecture is easier for the students to understand, and helps students to organize their own notes. For example, in Calculus it wouldn't make sense to learn second and third derivatives without covering the basics of first derivatives. A course organized in this manner would surely confuse, and set back the course in schedule.

Good rapport complements all the other attributes above. A professor that excites the students about the subject matter will stimulate their learning. The result is usually students that are excited to go to class, and eager to hear him speak on the subject. A personal example of mine was when one of us took ES 2503. Professor Grandin seemed to enjoy teaching Stress Analysis, and stimulated us with practical applications and examples. As a result I personally couldn't look at anything with out thinking about what kind of stresses the example was under. This surely had an effect on the effort I put into the class.

The category entitled *Other* is comprised of questions related to generalized questions about the course. For example the category asks about text books, lab equipment, and other related subjects of the course. This category is important but not as important to the students as *Grading*, *Communication*, or *Organization*.

The individual questions within the categories tell us what's important, and ranking the categories in an order of importance is also helpful. Unfortunately for us we identified a flaw within our survey, and the ranking of each categorization from the first survey was determined not useful. A second survey was administered to obtain useful data in order to determine any relationship between category ranking, and the averages from **Table 3**. Since the "Your Learning" column was more important to the students and was also close in comparison to "Overall Course Satisfaction," we used the results from the "Your Learning" column in **Table 3** to draw our conclusions from.

6.3 Summary of important results:

i). The responses from the surveys yielded skewed answers toward 1's and 2's.

ii). No correlation was found between "Overall Course Satisfaction" and "Your Learning."

iii). There is no relationship between the category ranking from T**able 2**, and the top 3 most important categories from **Table 3**.

7.0 CONCLUSIONS

The ranking of each question is important in developing a new course evaluation. Since statistical analysis shows mathematical differences between categories, we can conclude that some are more important than others. A new course evaluation should have weighted responses to certain categories. From the "Your Learning" column of **Table 3** the question with the lowest average is 1.76, and the highest 2.1. This doesn't seam like much of a difference, but if you take the difference between the two numbers and divide it by the highest ((2.1-1.76)/2.1) there's a 16% difference. Since %16 is a significant percentage, this is substantial enough to draw conclusions from. If we received these surveys and all categories had the same averages they would have a 0% difference. Then it wouldn't matter and an evaluation shouldn't have weighted questions. But since there are clear mathematical differences, these categories such as Grading, Workload/Difficulty, and Communication should have weighted importance over less important categories such as Rapport and Other. The categories ranked 3-5 don't have much of a difference between so we depend on our logic to separate them from the others, and draw conclusions from. This would be the same kind of principle as professors counting a students' exam more than they count their homework (example: final exam is 30% of final grade, and homework is 10% of final grade). The weighting of the six separate categories must add up to 100%. The information from the ranking in **Table 4** showed that in both columns Rapport, and Other were always ranked last, and in that order. The remaining four categories were given a weight of 20%, and Rapport and Other were given a weight of 10%.

Using questions from the existing course evaluation, and questions from our survey we have created a new course evaluation which we feel is a more precise tool for

measuring course and teacher effectiveness. The new evaluation can be seen in **Appendix D**. Changes from the survey to the our proposed evaluation include an enlarged scale, and a few added questions. We also wanted to try and maintain at least 3 questions for each category. We don't feel that this creates an evaluation which is too long, but one that is compete, and has categories that are complimentary toward each other. We used an eight point scale to try to clearly distinguish a strong opinion from one that is not really neutral, but not very strong either. The scale starts with "1" (strongly agree) to be consistent with our survey. The categories appear in the order of importance that was derived from **Table 3**.

The question "*The instructor used evaluations that were good measures of the material covered,"* was added to the Grading category not only to add more questions, but to also cover an important issue. Other changes were made to exclude or include questions that relevant to the categories. We had to do this to clearly define each category. These extra questions can from the existing evaluation so students are familiar with them, and we assume that they are correctly worded.

Although we feel that our revised evaluation is an improvement over the current WPI course evaluation, other improvements should be made before proposing the adoption of the new course evaluation. Further surveys should be administered to the faculty, and the student body should be resurveyed about the proposed changes in the evaluation.

BIBLIOGRAPHY

Svinicki, M and Menges, 1996 R, Honoring Exemplary Teaching Jossey-Bass Publishers, San Francisco, No. 65.

Salant and Dillman. How to Conduct Your Own Survey. 1994

Aubrecht, J.D., 1981 Reliability, Validity, and Generalizability of Student Ratings of Instruction (Idea Paper No. 2). Manhattan, KS: Center for Faculty Evaluation and Development.

Cashin, W.E., 1988 Student Ratings of Teaching: A summary of the Research (Idea Paper No. 20). Manhattan, KS: Center for Faculty Evaluation and Development.

Miller, R., 1972 Evaluating Faculty Performance. Jossey-Bass Publishers, San Francisco.

Lowman, J., 1984 Mastering the Techniques of Teaching. Jossey-Bass Publishers, San Francisco.

Appendix A: Student Course Evaluation Survey

Student-

Please take the time to fill out the following survey to help our IQP research on the development of a new course evaluation. The purpose of the following survey is to gain an understanding of students' opinions of important characteristics of a given course. The responses shown will be used to develop a revised course evaluation as a requirement for our IQP. Rate the importance of the following topics in relation to overall course satisfaction, and to your learning.

1: Very Important	2: Fairly Important	3: Somewhat Important	4: Not Very Important	5: Not at All Important

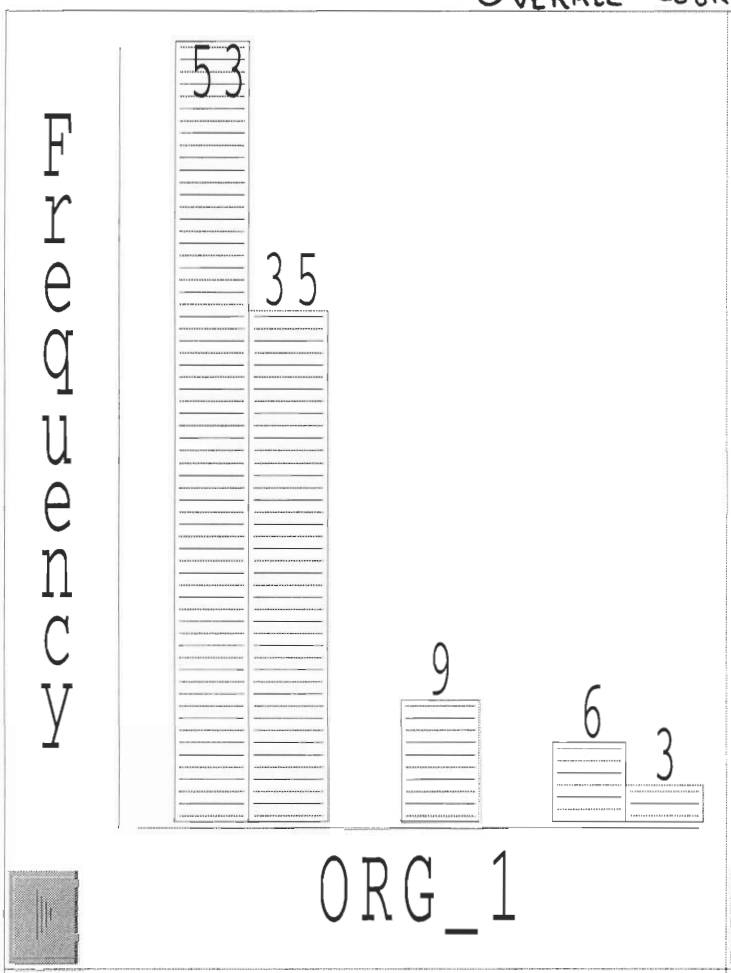
	Overall Course Satisfaction	Your Learning
Organization		
1) The instructor arrives well prepared for lectures.	1 2 3 4 5	1 2 3 4 5
2) The class objectives were defined.	1 2 3 4 5	1 2 3 4 5
3) The instructor used class time effectively.	1 2 3 4 5	1 2 3 4 5
Communication		
1) The instructor spoke clearly and comprehensibly.	1 2 3 4 5	1 2 3 4 5
2) The instructor gave good explanations of the material.	1 2 3 4 5	1 2 3 4 5
3) The instructor shared his/her class agenda/schedule.	1 2 3 4 5	1 2 3 4 5
Rapport		
1) The instructor seemed really concerned about the students.	1 2 3 4 5	1 2 3 4 5
2) The instructor showed enthusiasm for the subject.	1 2 3 4 5	1 2 3 4 5
3) The instructor was sensitive to the students, and was willing to work with them.	1 2 3 4 5	1 2 3 4 5
Difficulty/Workload		
1) The instructor made difficult topics easier to understand.	1 2 3 4 5	1 2 3 4 5
2) The homework/exams were good measures of the material.	1 2 3 4 5	1 2 3 4 5
3) The workload was appropriate for the course.	1 2 3 4 5	1 2 3 4 5
Grading		
1) The instructor graded fairly and consistently.	1 2 3 4 5	1 2 3 4 5
Other		
1) The textbook was helpful in this class.	1 2 3 4 5	1 2 3 4 5
2) The instructor was knowledgeble in the subject matter.	1 2 3 4 5	1 2 3 4 5
3) The class room was acceptable.	1 2 3 4 5	1 2 3 4 5

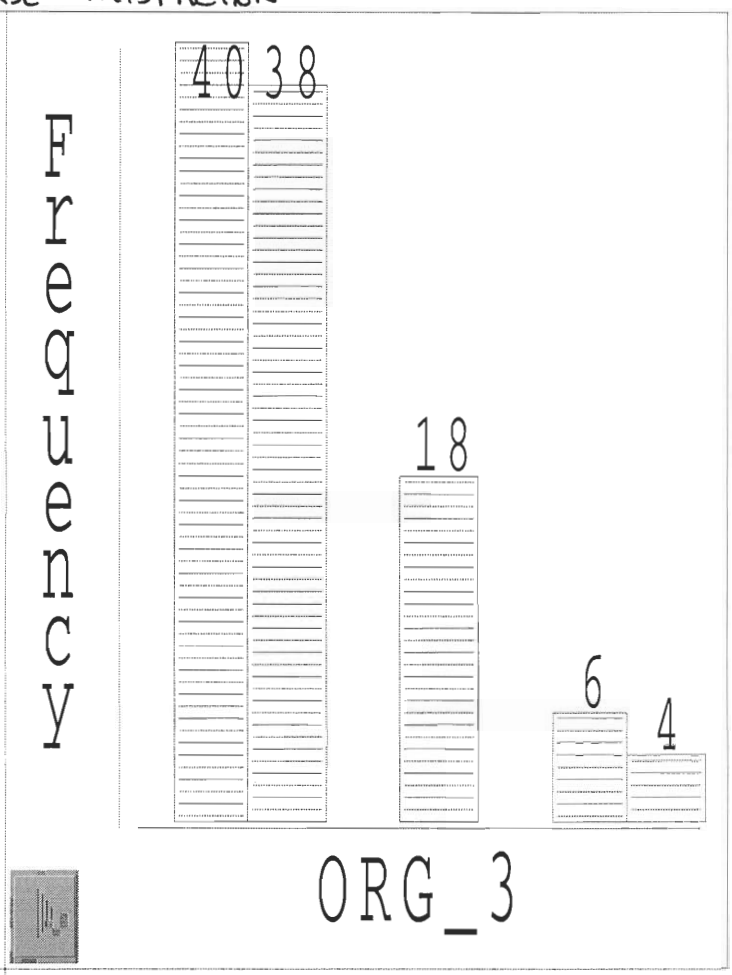
Please rank order the following categories according to their importance to your learning (1-5, 5 being most important)

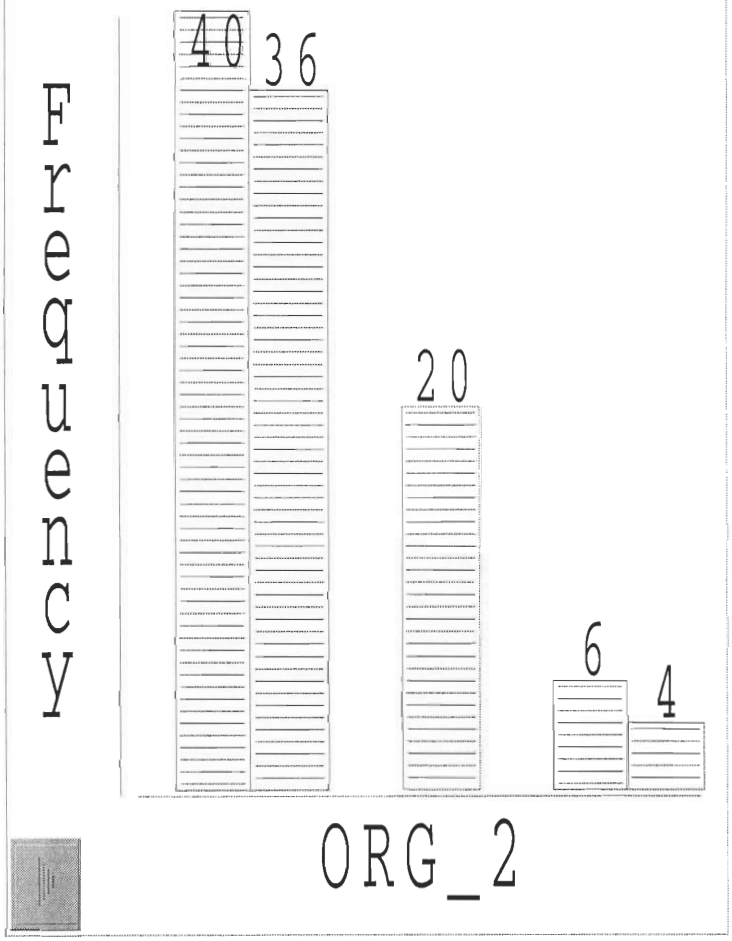
___ Organization
___ Communication
___ Rapport
___ Difficulty/Workload
___ Grading

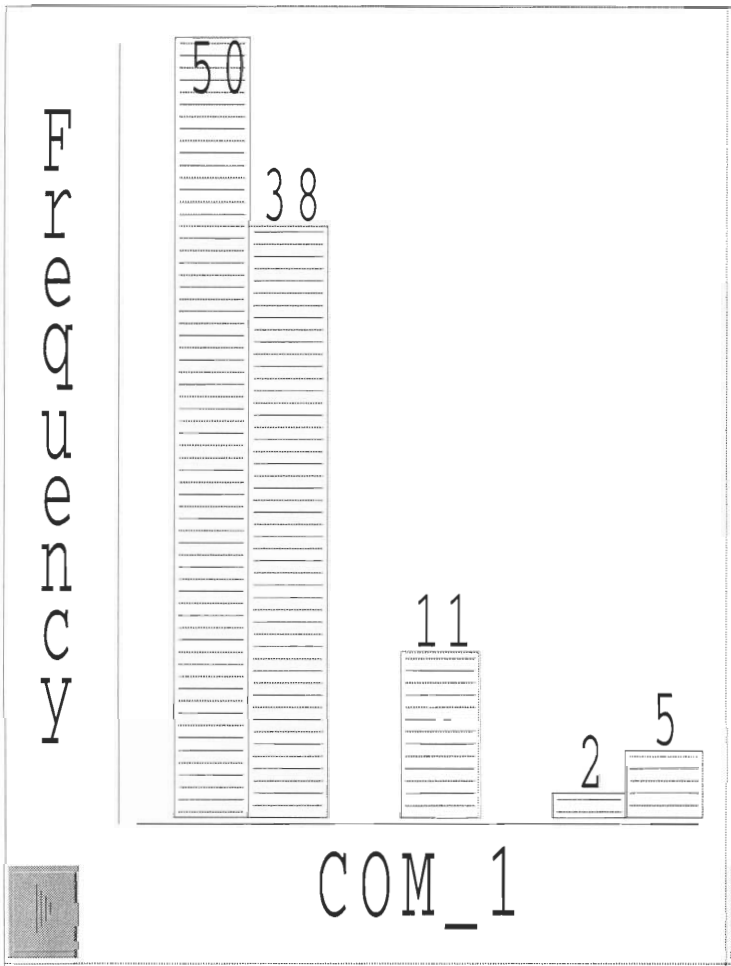
Appendix B: Frequency Histograms of individual questions from "Your Learning" and "Overall Course Satisfaction"

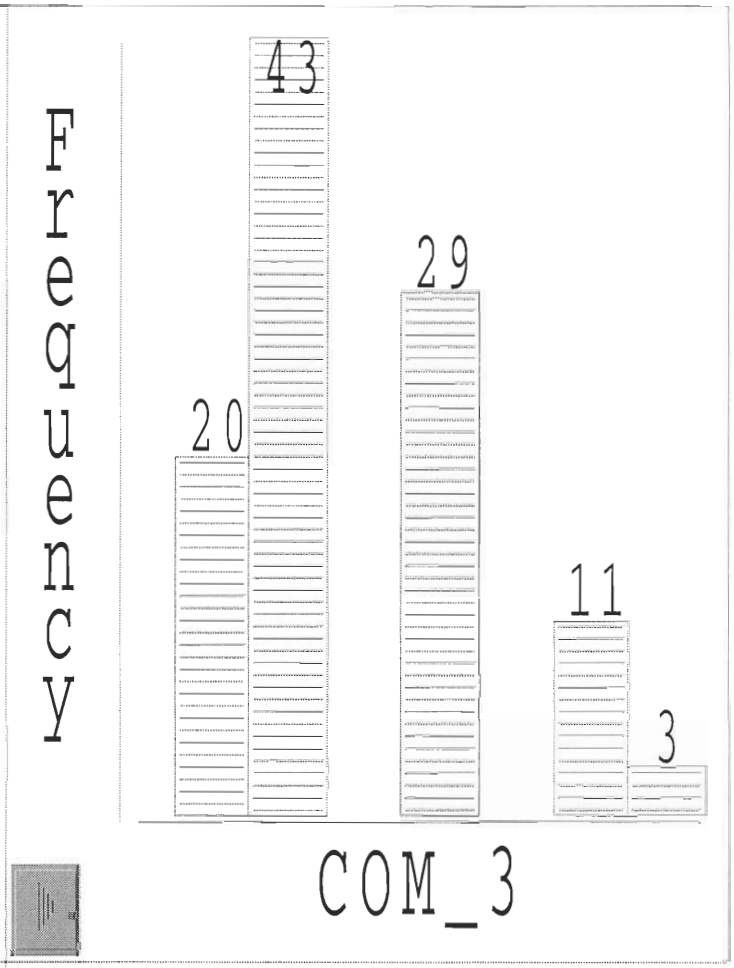
Overall Course Satisfaction

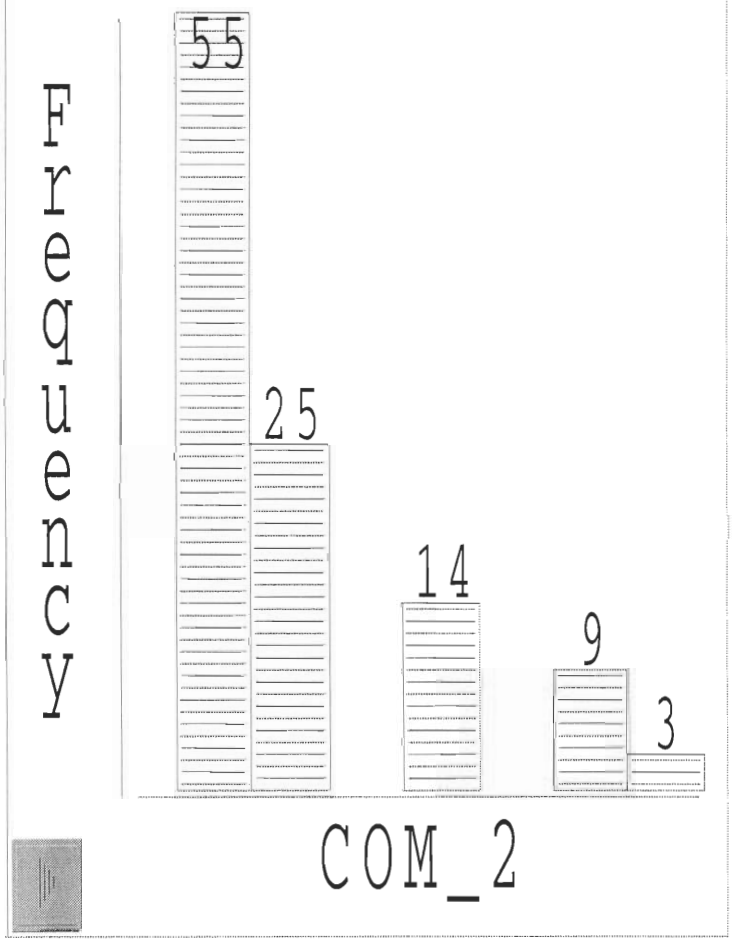


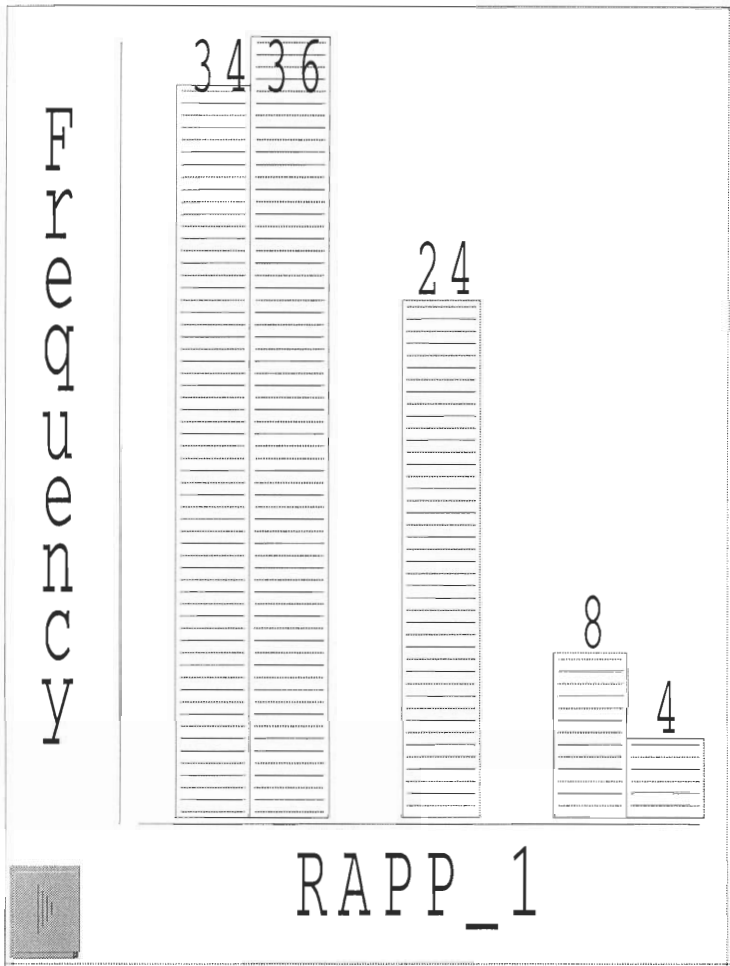


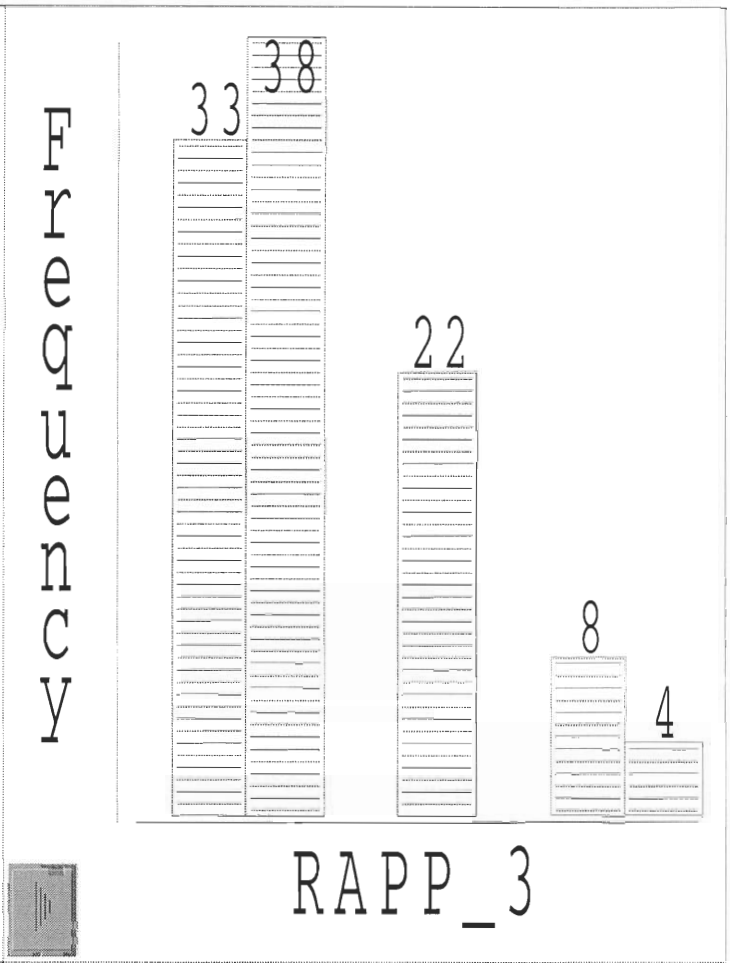


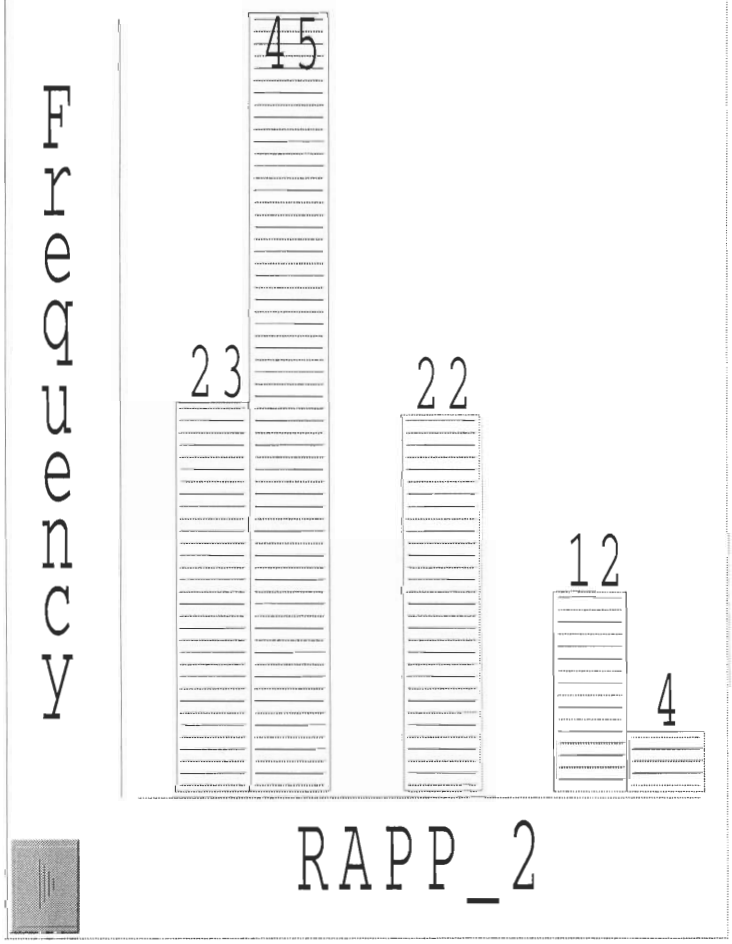

Frequency
50
38
11
2
5
COM_1

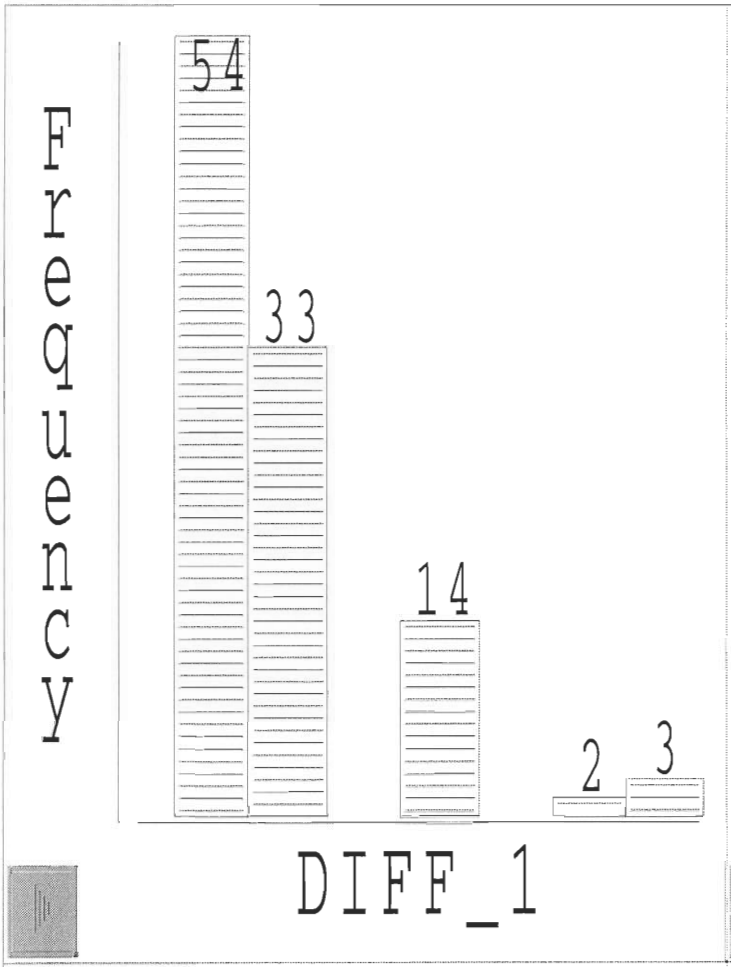

Frequency
43
20
29
11
3
COM_3

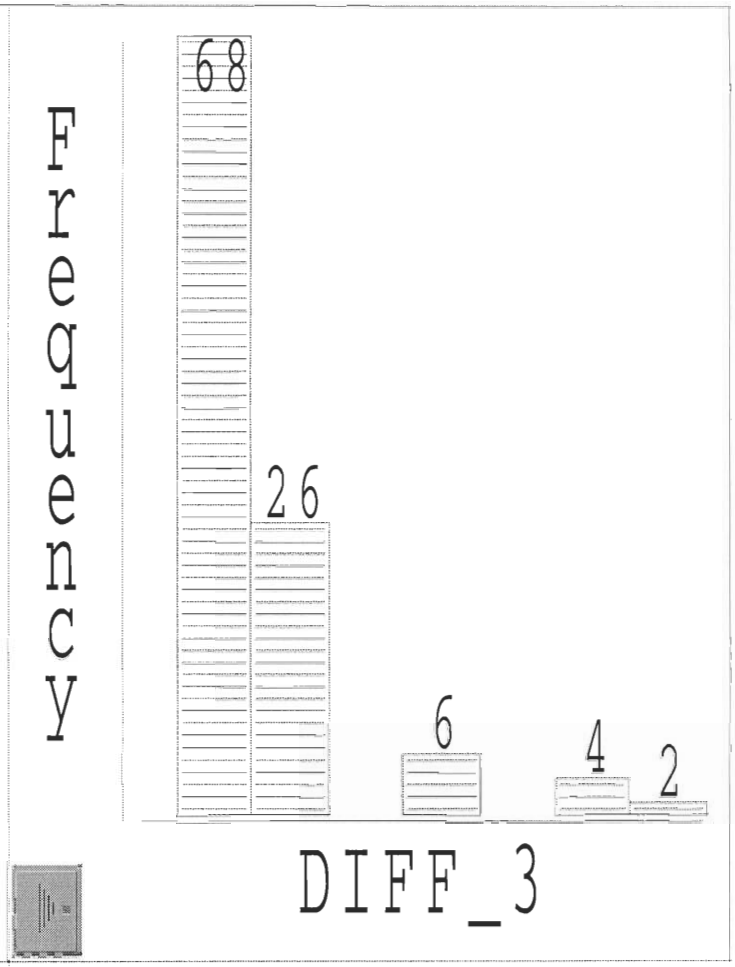

Frequency
55
25
14
9
3
COM_2

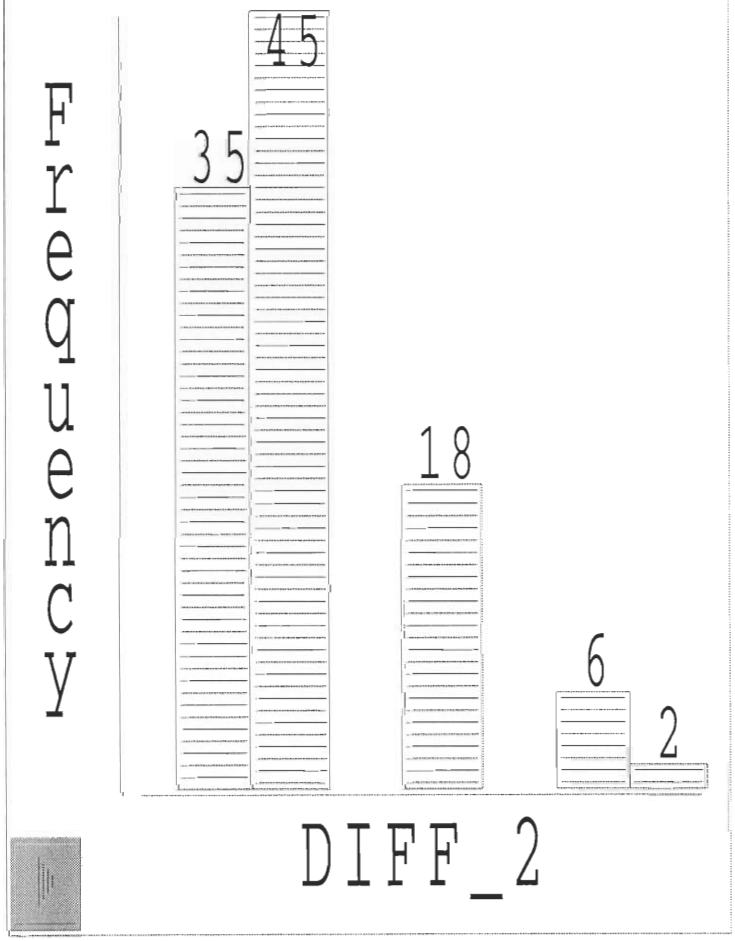

Frequency
34
36
24
8
4
RAPP_1

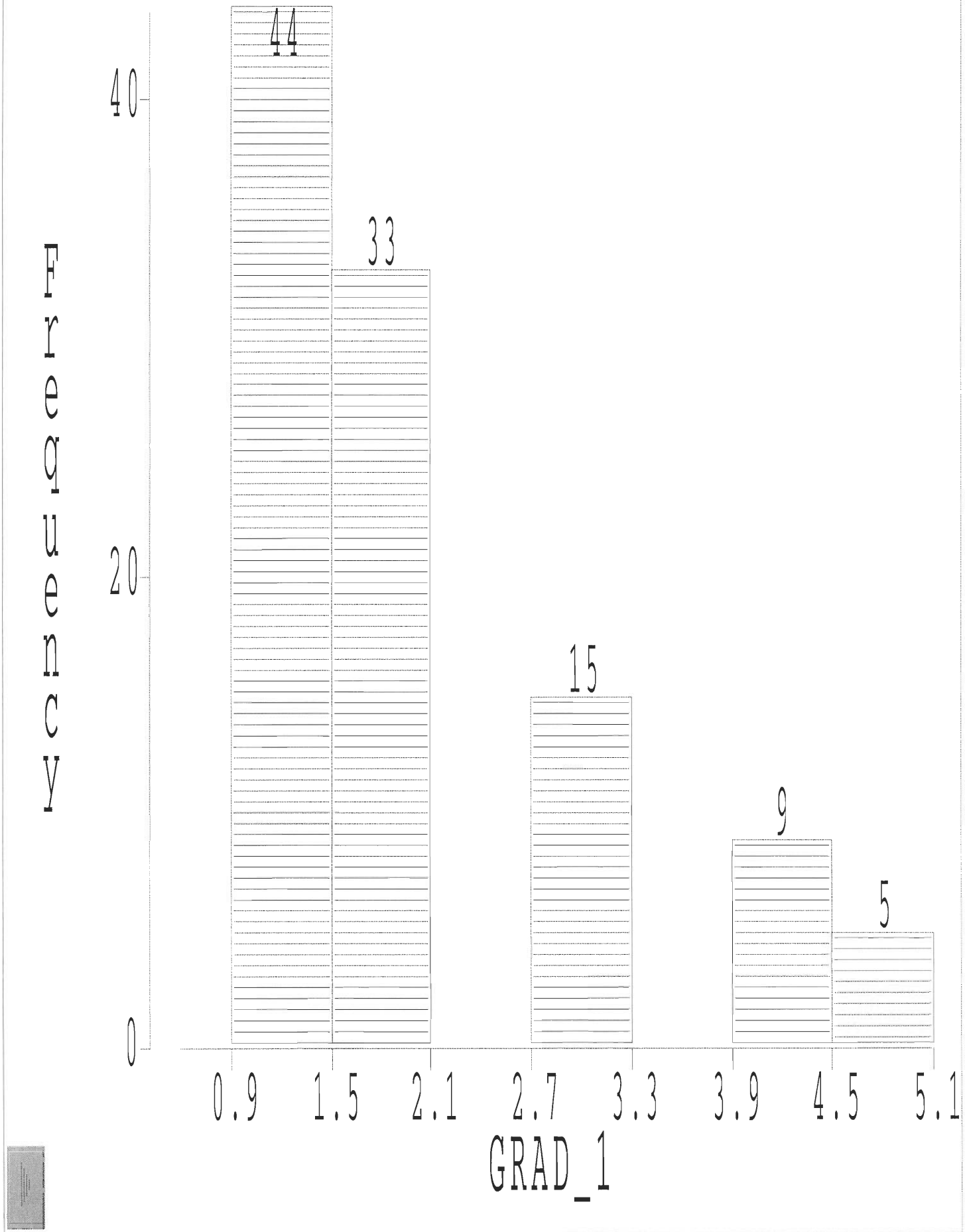

Frequency
33
38
22
8
4
RAPP_3

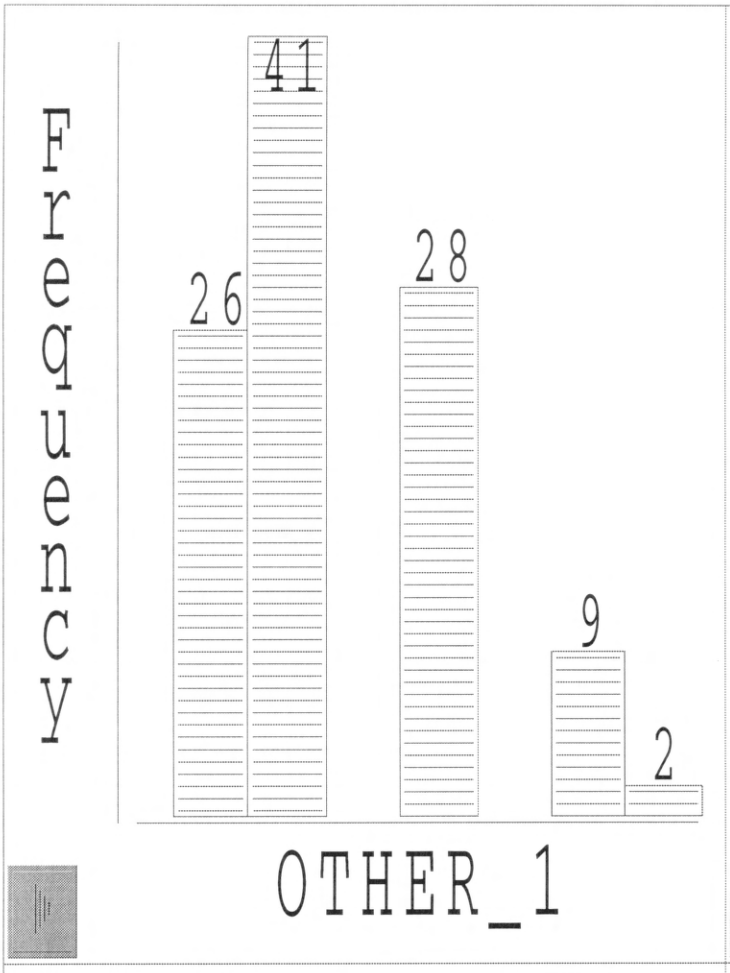

Frequency
45
23
22
12
4
RAPP_2

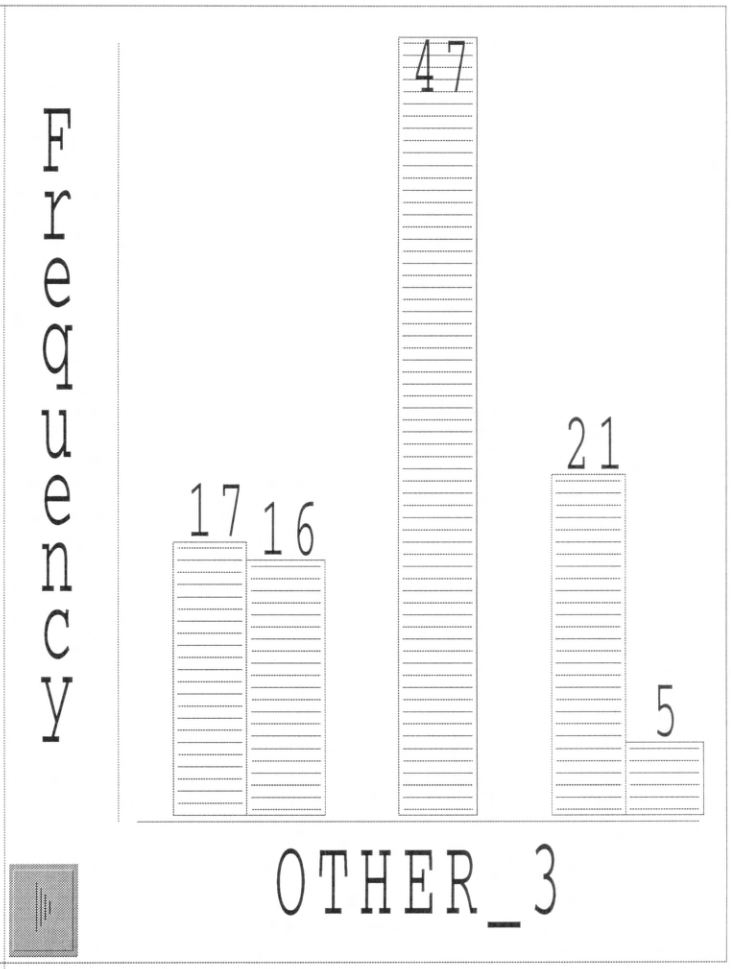

Frequency
54
33
14
2
3
DIFF_1

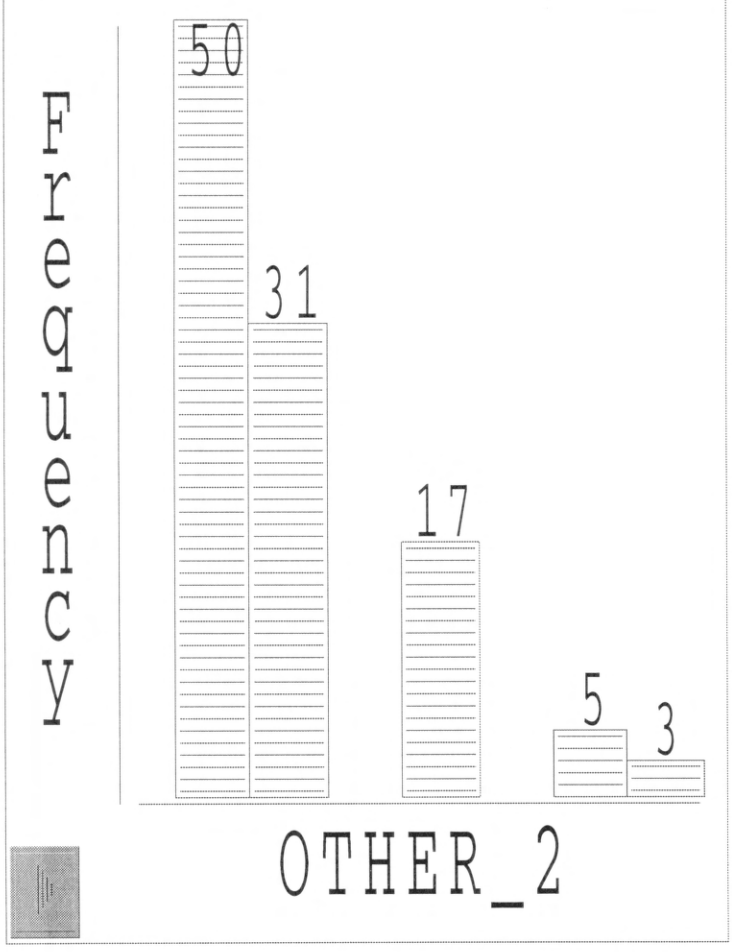

Frequency
68
26
6
4
2
DIFF_3


Frequency
45
35
18
6
2
DIFF_2

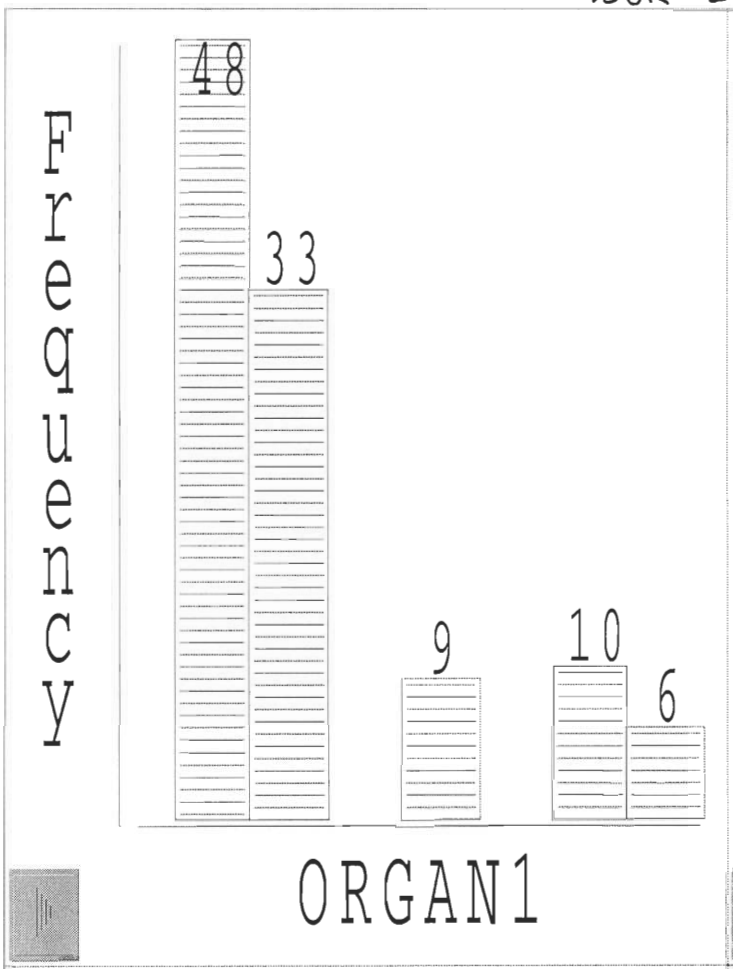

Frequency
40
20
0
44
33
15
9
5
0.9
1.5
2.1
2.7
3.3
3.9
4.5
5.1
GRAD_1

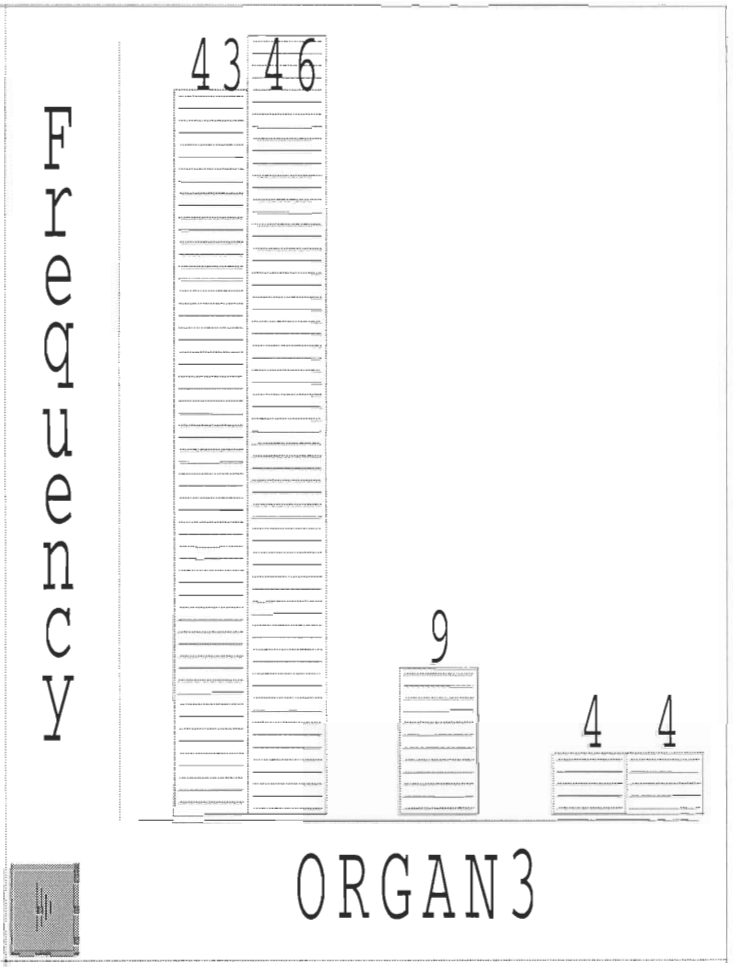

Frequency
41
26
28
9
2
OTHER_1

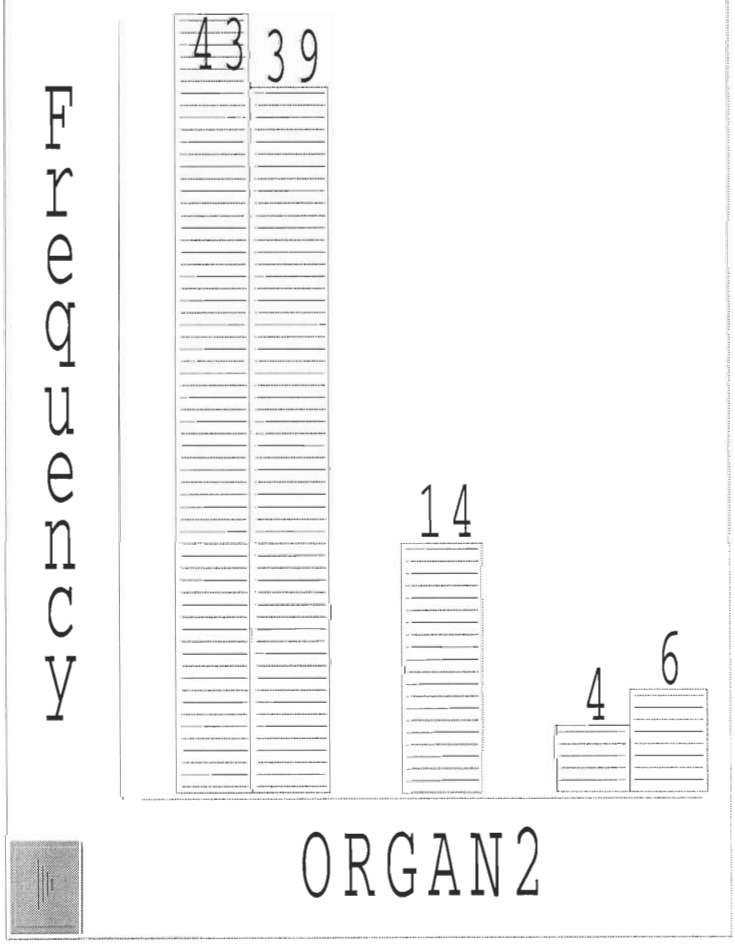

Frequency
47
17
16
21
5
OTHER_3


Frequency
50
31
17
5
3
OTHER_2

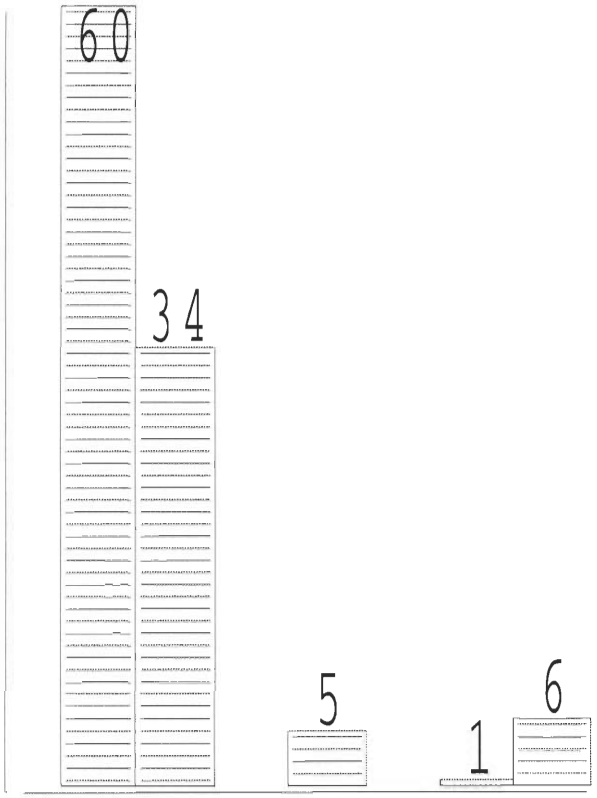
YOUR LEARNING





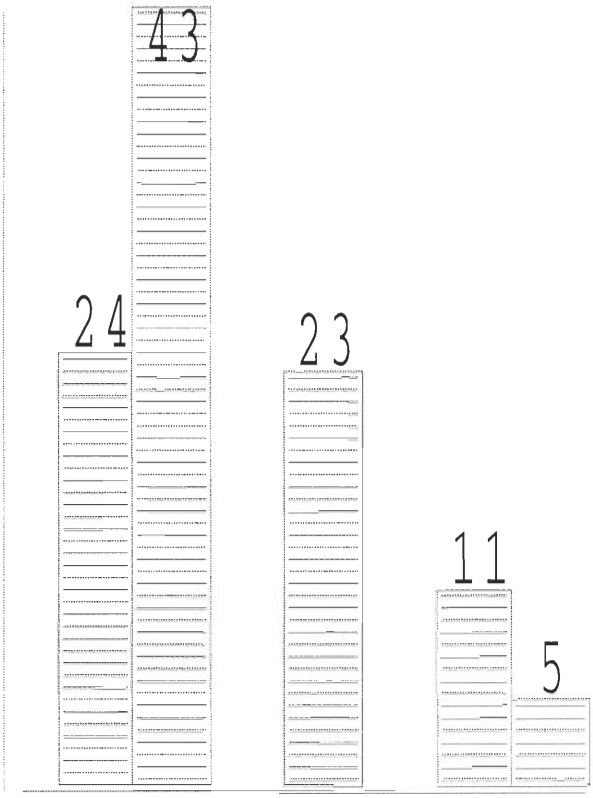


Frequency



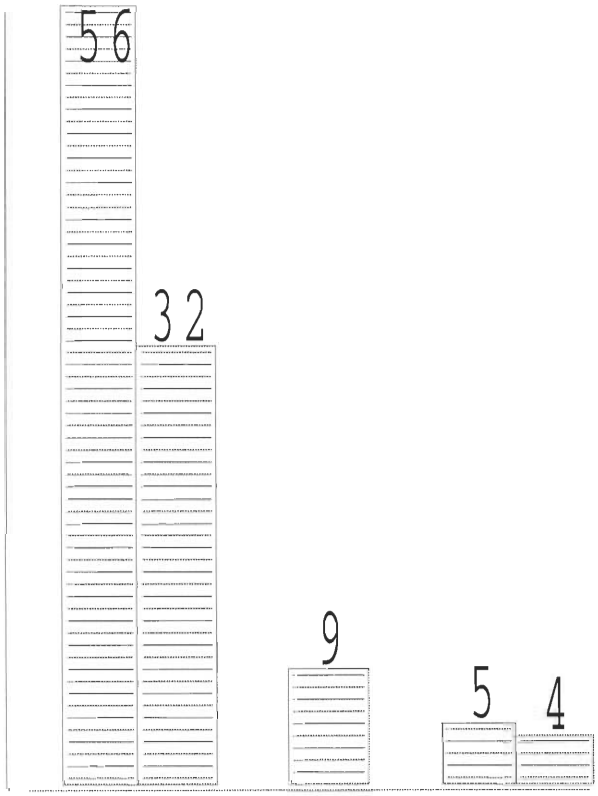
COMM1

Frequency

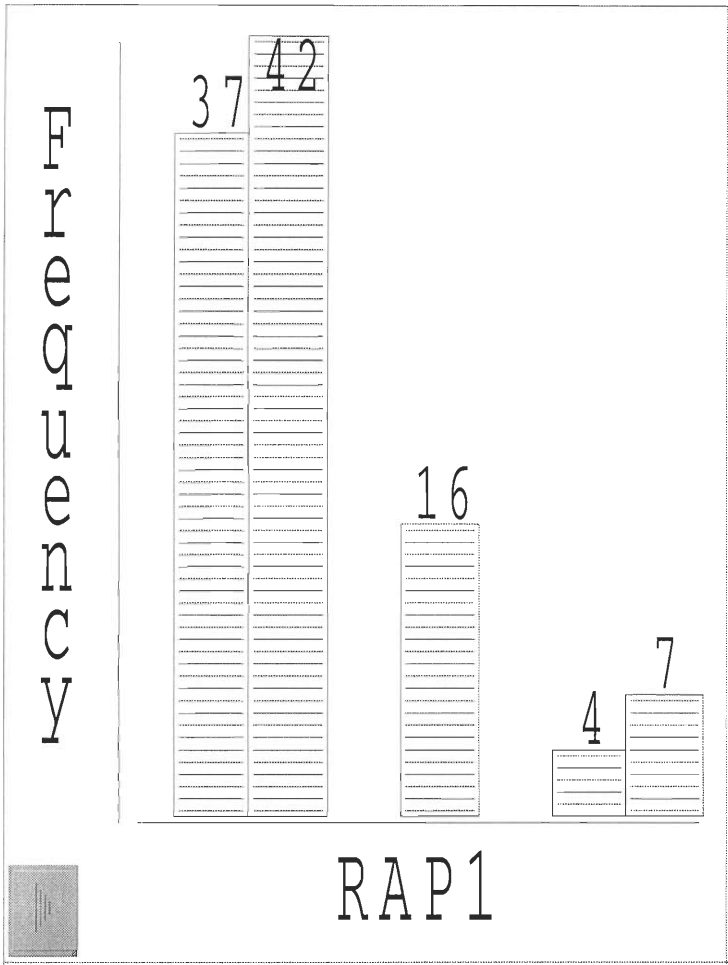


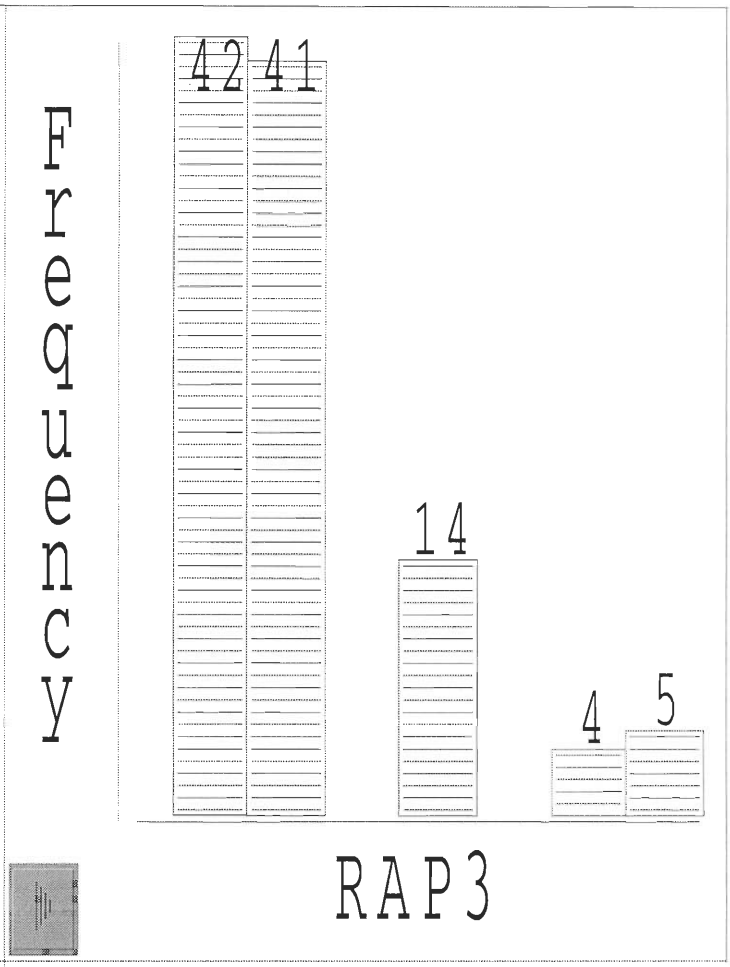
COMM3

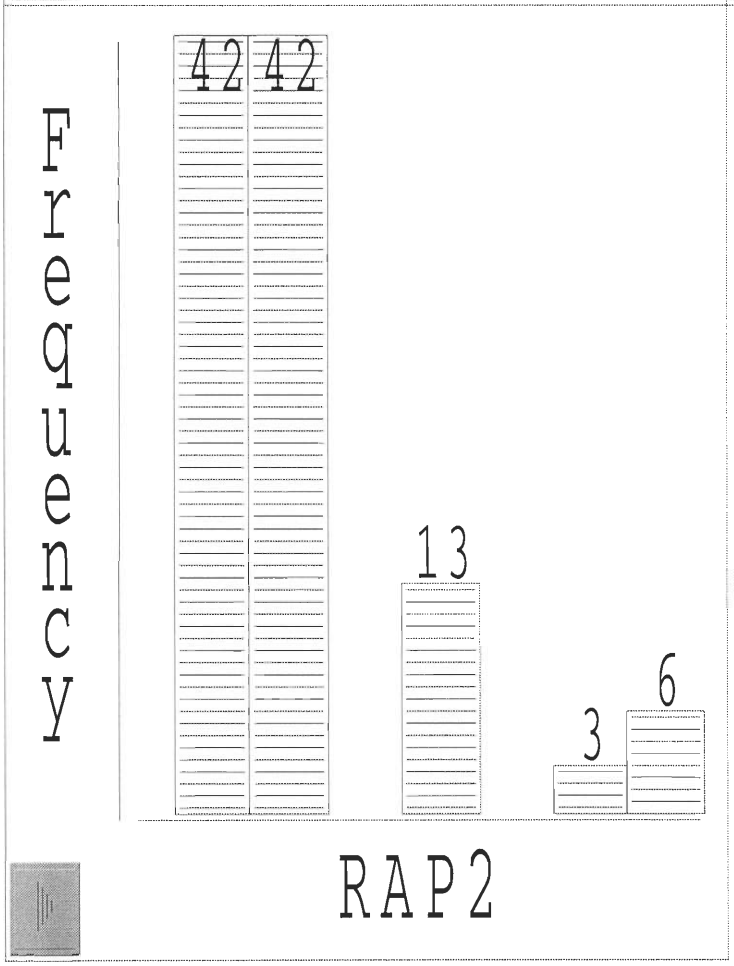
Frequency

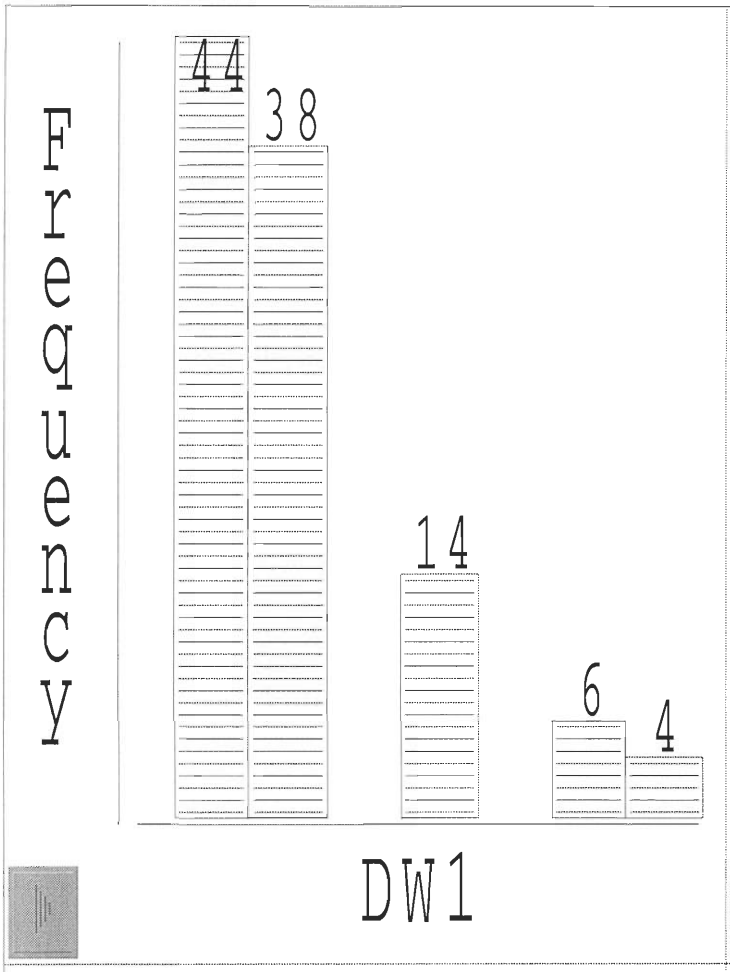


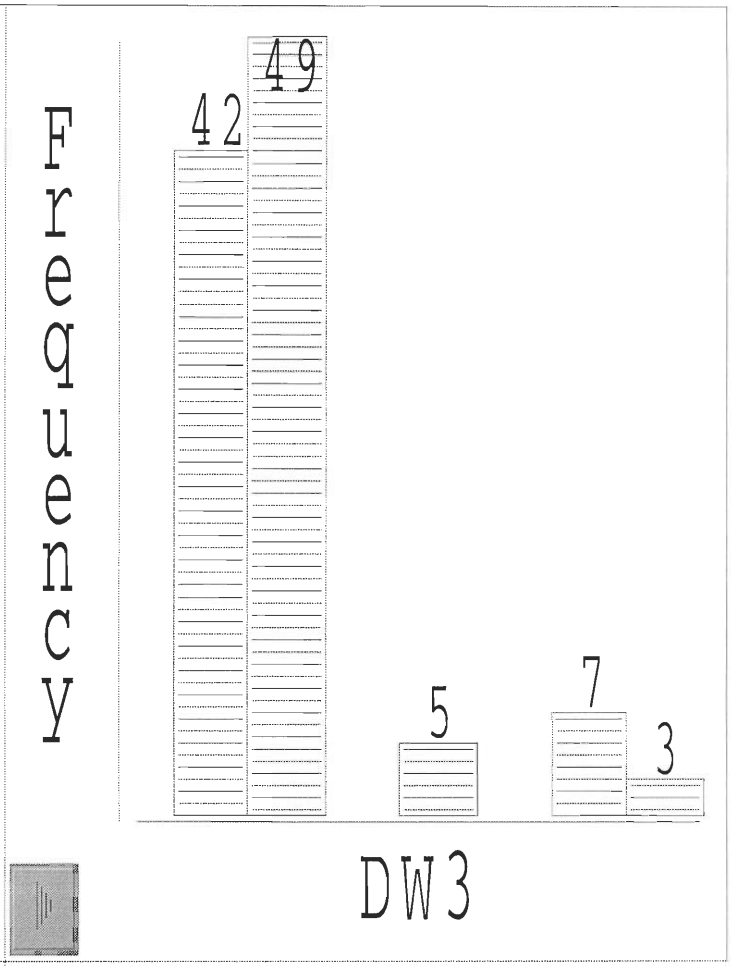
COMM2

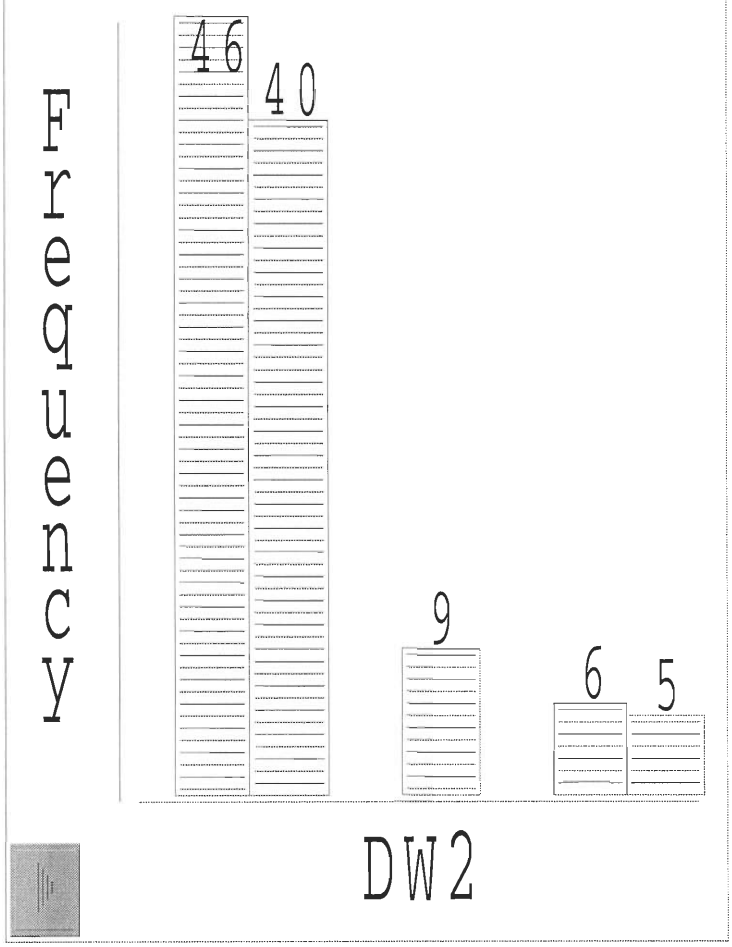

Frequency
37
42
16
4
7
RAP1

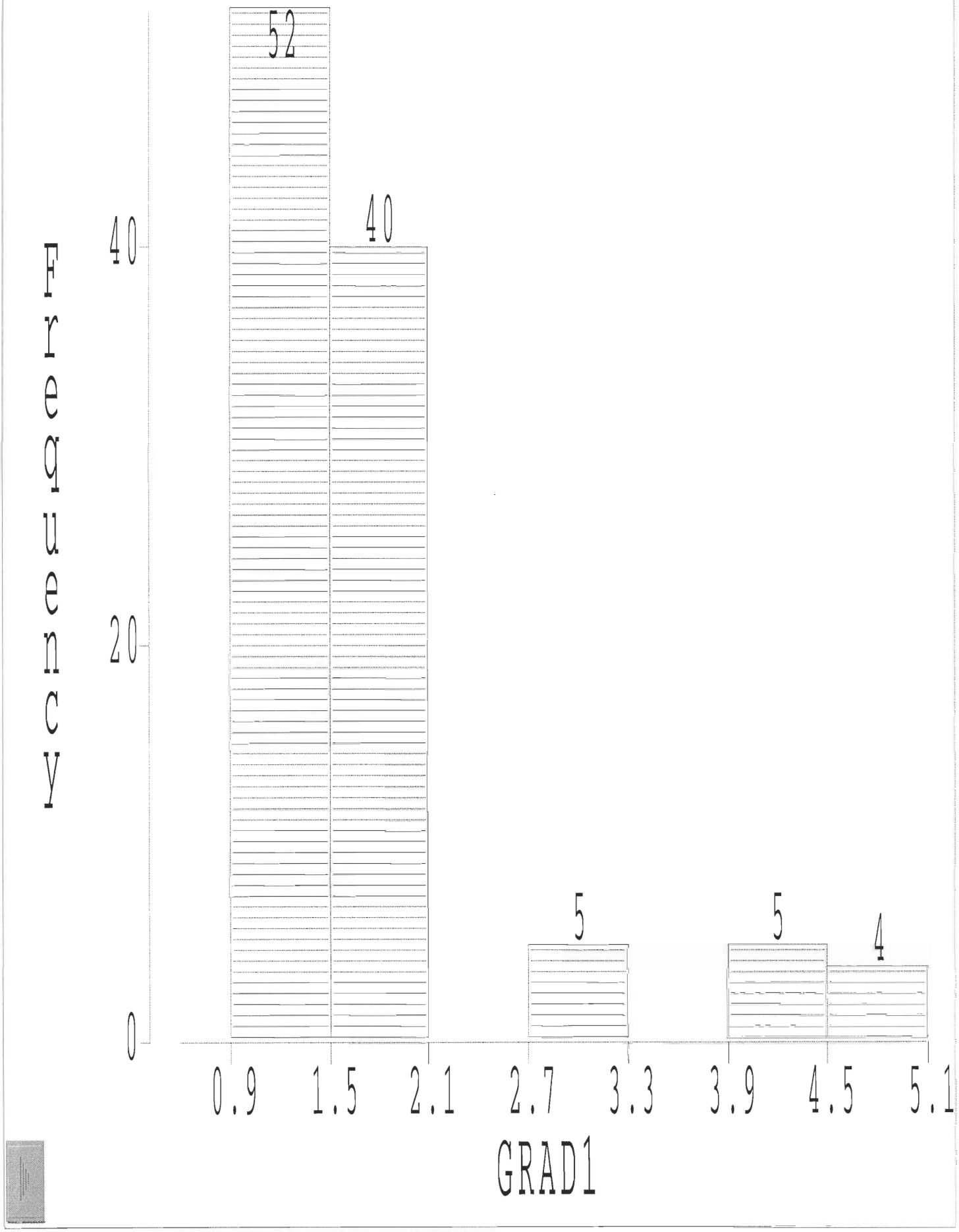

Frequency
42
41
14
4
5
RAP3

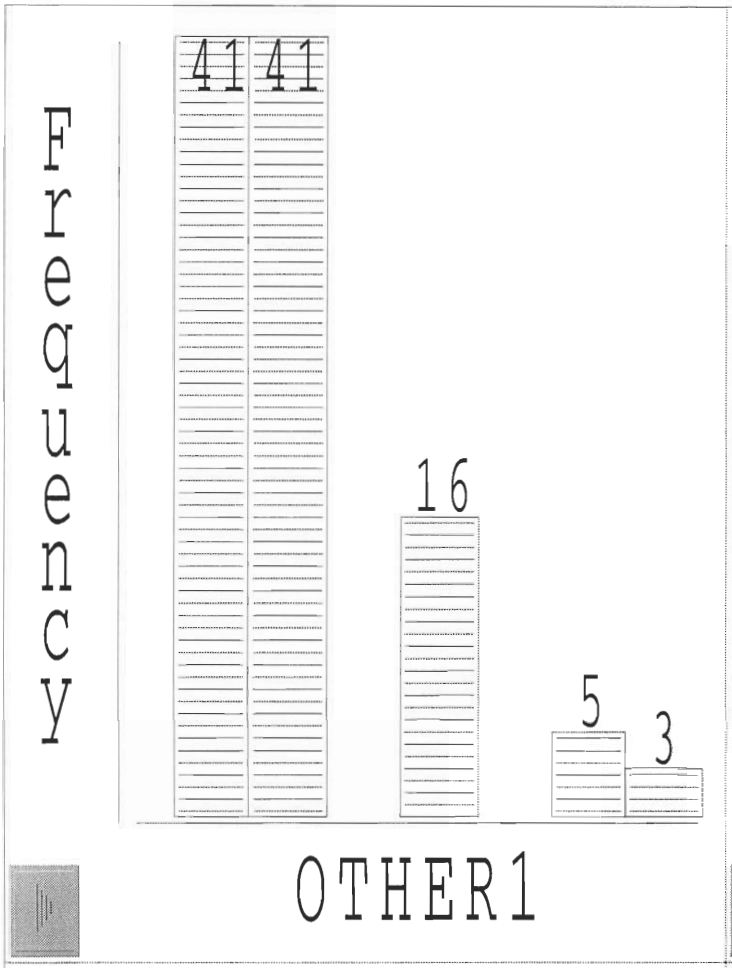

Frequency
42
42
13
3
6
RAP2

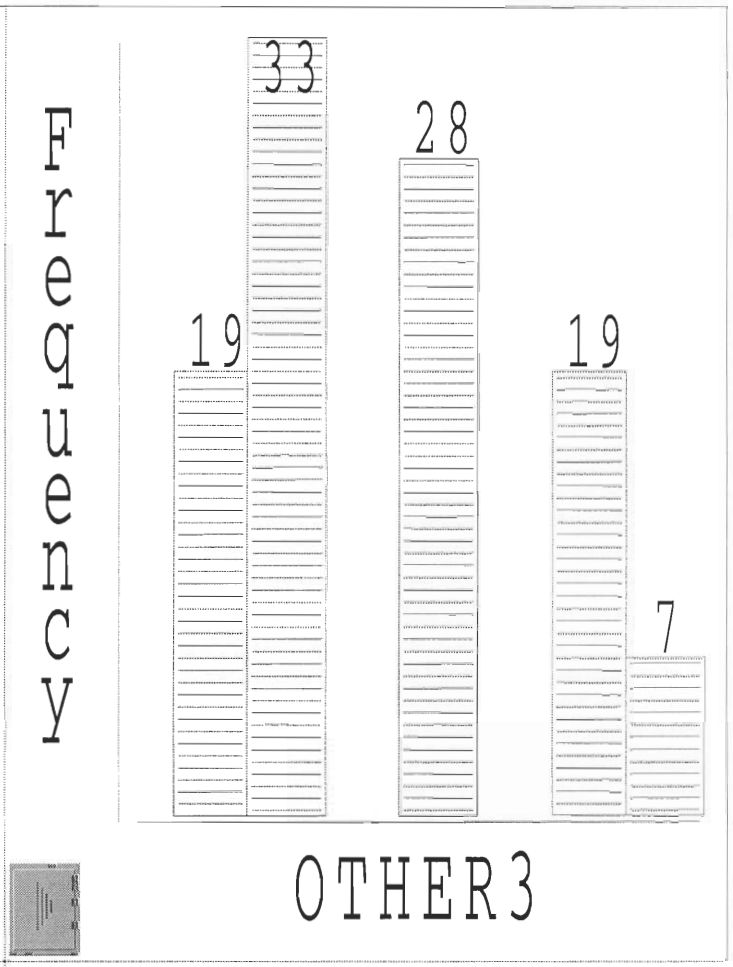

Frequency
44
38
14
6
4
DW1

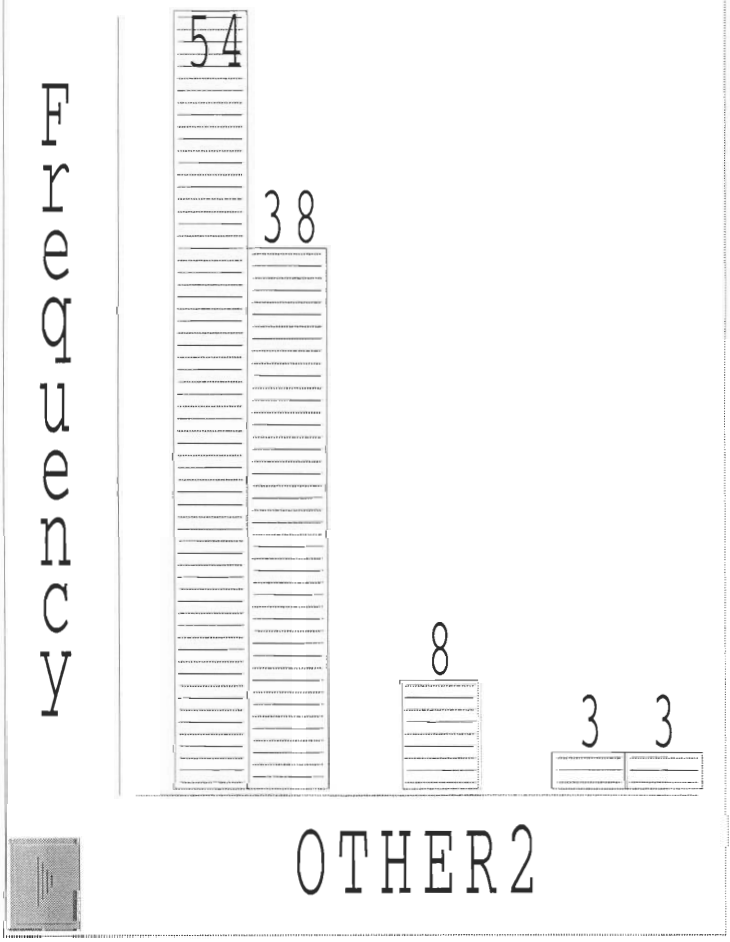

Frequency
42
49
5
7
3
DW3


Frequency
46
40
9
6
5
DW2


Frequency
40
20
0
52
40
5
5
4
0.9
1.5
2.1
2.7
3.3
3.9
4.5
5.1
GRAD1


Frequency
41
41
16
5
3
OTHER1


Frequency
33
28
19
19
7
OTHER3


Frequency
54
38
8
3
3
OTHER2

Appendix C: Correlation Matrices of questions from "Your Learning" and "Overall Course Satisfaction"

VARIOUS DATA INCLUDING CORRELATION COEFFICIENTS.

ORG_1 ORG_2 ORG_3 ORGAN1 ORGAN2 ORGAN3

Univariate Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
ORG_1	106	1.7830	1.0142	1.0000	5.0000
ORG_2	106	2.0377	1.0684	1.0000	5.0000
ORG_3	106	2.0189	1.0599	1.0000	5.0000
ORGAN1	106	1.9906	1.1992	1.0000	5.0000
ORGAN2	106	1.9717	1.0994	1.0000	5.0000
ORGAN3	106	1.8679	0.9863	1.0000	5.0000

Correlation Matrix

	ORG_1	ORG_2	ORG_3	ORGAN1	ORGAN2	ORGAN3
ORG_1	1.0000	0.5877	0.6771	0.0296	0.0286	0.0282
ORG_2	0.5877	1.0000	0.5797	-0.0369	-0.0639	-0.0404
ORG_3	0.6771	0.5797	1.0000	0.0451	-0.0731	0.0297
ORGAN1	0.0296	-0.0369	0.0451	1.0000	0.4188	0.4498
ORGAN2	0.0286	-0.0639	-0.0731	0.4188	1.0000	0.4971
ORGAN3	0.0282	-0.0404	0.0297	0.4498	0.4971	1.0000

COM_1 COM_2 COM_3 COMM1 COMM2 COMM3

Univariate Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
COM_1	106	1.8113	1.0246	1.0000	5.0000
COM_2	106	1.8679	1.1133	1.0000	5.0000
COM_3	106	2.3774	0.9996	1.0000	5.0000
COMM1	106	1.6698	1.0302	1.0000	5.0000
COMM2	106	1.7642	1.0470	1.0000	5.0000
COMM3	106	2.3396	1.0857	1.0000	5.0000

Correlation Matrix

	COM_1	COM_2	COM_3	COMM1	COMM2	COMM3
COM_1	1.0000	0.6208	0.6282	0.0306	0.0025	0.1266
COM_2	0.6208	1.0000	0.5673	-0.0550	0.0302	0.1005
COM_3	0.6282	0.5673	1.0000	-0.1276	-0.0598	0.1616
COMM1	0.0306	-0.0550	-0.1276	1.0000	0.7041	0.4162
COMM2	0.0025	0.0302	-0.0598	0.7041	1.0000	0.5068
COMM3	0.1266	0.1005	0.1616	0.4162	0.5068	1.0000

RAPP_1 RAPP_2 RAPP_3 RAP1 RAP2 RAP3

Univariate Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
RAPP_1	105	2.1810	1.0811	1.0000	5.0000
RAPP_2	105	2.3429	1.0546	1.0000	5.0000
RAPP_3	105	2.1619	1.0753	1.0000	5.0000
RAP1	105	2.0762	1.1240	1.0000	5.0000
RAP2	105	1.9524	1.0776	1.0000	5.0000
RAP3	105	1.9524	1.0596	1.0000	5.0000

Correlation Matrix

	RAPP_1	RAPP_2	RAPP_3	RAP1	RAP2	RAP3
RAPP_1	1.0000	0.5776	0.5204	0.0281	-0.0421	0.0412
RAPP_2	0.5776	1.0000	0.6544	-0.0141	0.0484	0.1438
RAPP_3	0.5204	0.6544	1.0000	-0.1296	-0.0597	-0.0185
RAP1	0.0281	-0.0141	-0.1296	1.0000	0.6143	0.6167
RAP2	-0.0421	0.0484	-0.0597	0.6143	1.0000	0.5791
RAP3	0.0412	0.1438	-0.0185	0.6167	0.5791	1.0000

DW1 DW2 DW3 DIFF_1 DIFF_2 DIFF_3

Univariate Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
DW1	106	1.9434	1.0586	1.0000	5.0000
DW2	106	1.9057	1.0826	1.0000	5.0000
DW3	106	1.8679	0.9766	1.0000	5.0000
DIFF_1	106	1.7453	0.9568	1.0000	5.0000
DIFF_2	106	2.0094	0.9511	1.0000	5.0000
DIFF_3	106	1.5472	0.9064	1.0000	5.0000

Correlation Matrix

	DW1	DW2	DW3	DIFF_1	DIFF_2	DIFF_3
DW1	1.0000	0.5687	0.5454	0.0609	0.1330	0.0227
DW2	0.5687	1.0000	0.5916	0.0685	0.0194	0.0628
DW3	0.5454	0.5916	1.0000	0.0248	0.0424	0.0932
DIFF_1	0.0609	0.0685	0.0248	1.0000	0.6306	0.7003
DIFF_2	0.1330	0.0194	0.0424	0.6306	1.0000	0.6899
DIFF_3	0.0227	0.0628	0.0932	0.7003	0.6899	1.0000

GRAD_1 GRAD1

Univariate Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
GRAD_1	106	2.0377	1.1541	1.0000	5.0000
GRAD1	106	1.7642	1.0100	1.0000	5.0000

Correlation Matrix

	GRAD_1	GRAD1
GRAD_1	1.0000	0.0077
GRAD1	0.0077	1.0000

OTHER_1 OTHER_2 OTHER_3 OTHER1 OTHER2 OTHER3

Univariate Statistics

Variable	N	Mean	Std Dev	Minimum	Maximum
OTHER_1	106	2.2453	0.9838	1.0000	5.0000
OTHER_2	106	1.8679	1.0335	1.0000	5.0000
OTHER_3	106	2.8208	1.0761	1.0000	5.0000
OTHER1	106	1.9434	0.9936	1.0000	5.0000
OTHER2	106	1.7075	0.9357	1.0000	5.0000
OTHER3	106	2.6415	1.1645	1.0000	5.0000

Correlation Matrix

	OTHER_1	OTHER_2	OTHER_3	OTHER1	OTHER2	OTHER3
OTHER_1	1.0000	0.5661	0.2218	-0.1123	-0.0972	-0.1470
OTHER_2	0.5661	1.0000	0.1583	0.0112	0.0188	-0.1426
OTHER_3	0.2218	0.1583	1.0000	-0.0096	-0.0904	0.0090
OTHER1	-0.1123	0.0112	-0.0096	1.0000	0.4020	0.1881
OTHER2	-0.0972	0.0188	-0.0904	0.4020	1.0000	0.1301
OTHER3	-0.1470	-0.1426	0.0090	0.1881	0.1301	1.0000

Appendix D: New WPI Student Course Evaluation

CLASS_____ TERM_____ INSTRUCTOR_____

SCALE: 1:Strongly Agree 2:Agree 3:Somewhat Agree 4:Nuetral 5:Somewhat Disagree 6:Disagree 7:Strongly Disagree 8:N/A

Grading: Weight=20%

The instructor graded fairly and consistently.	1 2 3 4 5 6 7 8
The instructor used evaluations that were good measures of the material covered.	1 2 3 4 5 6 7 8

Difficulty / Workload: Weight=20%

The workload was appropriate for the course.	1 2 3 4 5 6 7 8
The instructor assigned homework that aided my learning.	1 2 3 4 5 6 7 8
The material to be learned in this course was difficult.	1 2 3 4 5 6 7 8

Communication: Weight=20%

The instructor spoke clearly and comprehensibly.	1 2 3 4 5 6 7 8
The instructor gave good explanations of the material.	1 2 3 4 5 6 7 8
The instructor shared his/her class agenda/schedule.	1 2 3 4 5 6 7 8
The instructor demonstrated a good understanding of the material being taught	1 2 3 4 5 6 7 8
The instructor used the blackboard/visual aids in an effective manner.	1 2 3 4 5 6 7 8
The instructor clearly defined the requirements for preparing lab reports.	1 2 3 4 5 6 7 8

Organization: Weight=20%

The instructor arrives well prepared for lectures.	1 2 3 4 5 6 7 8
The class objectives were defined.	1 2 3 4 5 6 7 8
The instructor used class time effectively.	1 2 3 4 5 6 7 8

Rapport: Weight=10%

The instructor stimulated my interest in the subject matter.	1 2 3 4 5 6 7 8
The instructor seemed really concerned about students.	1 2 3 4 5 6 7 8
The instructor challenged me to extend my capabilities.	1 2 3 4 5 6 7 8

Other: Weight=10%

The textbook(s) helped me to learn the subject matter.	1 2 3 4 5 6 7 8
The room used for the course was acceptable.	1 2 3 4 5 6 7 8
The lab and/or computer equipment was in good operating condition.	1 2 3 4 5 6 7 8
I rate myself in general as an excellent student.	1 2 3 4 5 6 7 8

Student year:_____ Student Major:_____ Sex: M F