

1993

# Calibration of Clinical Instruction in a Dental Hygiene Program

Barbara L. Bennett

This research is a product of the graduate program in [Technology](#) at Eastern Illinois University. [Find out more](#) about the program.

---

## Recommended Citation

Bennett, Barbara L., "Calibration of Clinical Instruction in a Dental Hygiene Program" (1993). *Masters Theses*. 2135.  
<https://thekeep.eiu.edu/theses/2135>

This is brought to you for free and open access by the Student Theses & Publications at The Keep. It has been accepted for inclusion in Masters Theses by an authorized administrator of The Keep. For more information, please contact [tabruns@eiu.edu](mailto:tabruns@eiu.edu).

THESIS REPRODUCTION CERTIFICATE

TO: Graduate Degree Candidates who have written formal theses.

SUBJECT: Permission to reproduce theses.

The University Library is receiving a number of requests from other institutions asking permission to reproduce dissertations for inclusion in their library holdings. Although no copyright laws are involved, we feel that professional courtesy demands that permission be obtained from the author before we allow theses to be copied.

Please sign one of the following statements:

Booth Library of Eastern Illinois University has my permission to lend my thesis to a reputable college or university for the purpose of copying it for inclusion in that institution's library or research holdings.

7/27/93

Date

I respectfully request Booth Library of Eastern Illinois University not allow my thesis be reproduced because \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Date

\_\_\_\_\_

Author

CALIBRATION OF CLINICAL INSTRUCTION

---

IN A DENTAL HYGIENE PROGRAM

---

(TITLE)

BY

BARBARA L. BENNETT

**THESIS**

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF

MASTER OF SCIENCE

---

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY  
CHARLESTON, ILLINOIS

1993  
YEAR

---

I HEREBY RECOMMEND THIS THESIS BE ACCEPTED AS FULFILLING  
THIS PART OF THE GRADUATE DEGREE CITED ABOVE

7/27/93  
DATE

7/27/93  
DATE

## CONTENTS

	Page
Acknowledgements.....	1
Abstract.....	2
Chapter I	
Background Information.....	3
Research Questions, Statement of Research.....	5
Purpose of Study, Significance.....	6
Definitions.....	6
Assumptions of Study.....	8
Hypothesis.....	8
Summary.....	9
Chapter II	
Review of Related Literature	
Clinical Evaluation in Health Professions Education.....	10
Rater Reliability in Calculus Detection.....	14
Planning Quality Professional Development Programs.....	17
Establishment of Objectives.....	24
Summary.....	27
Chapter III	
Methodology	
Overview.....	29
Design of Study.....	29
Selection of Patients and Examiners.....	30
Methods and Materials.....	30
Intrarater Reliability.....	31
Interrater Reliability.....	33

	Page
Data Analysis.....	37
Summary.....	38

## Chapter IV

### Results of the Study

Overview.....	39
Demographics.....	40
Research Findings for Question One.....	40
Research Findings for Question Two.....	42
Comparison of Rater Errors.....	44
Survey of Faculty Attitudes.....	45
Summary.....	52

## Chapter V

### Summary, Discussion, Conclusions and Recommendations

Summary.....	53
Results and Discussion.....	55
Conclusions.....	63
Recommendations.....	64
References Cited.....	69
Appendix A.....	75
Appendix B.....	77
Appendix C.....	79
Appendix D.....	83
Appendix E.....	84
Appendix F.....	89

Calibration of Clinical Instruction  
in a Dental Hygiene Program  
Barbara L. Bennett  
Eastern Illinois University

This research was conducted while the author was a master's degree candidate at Eastern Illinois University in Charleston, Illinois. I wish to thank the dental hygiene faculty, Mary Emmons and Judith Henthorn, program directors, and Susan Maurer, department chair of the health professions department at Parkland College in Champaign, Illinois, for their cooperation on this project. I especially wish to thank my advisor, Dr. Deborah Woodley for her generous help and support, and my committee, Dr. Gene Strandberg, Dr. Louis Butler, and Dr. Joyce Felstehausen for their time and efforts. Thanks also go to Dr. David Dodd for his help in the statistical analysis of the research, to Rob Bennett for his technical support, and to Bob Bennett for his encouragement and optimism.

RUNNING HEAD: CALIBRATION OF CLINICAL INSTRUCTION

### Abstract

Continuous quality improvement training was used to improve intrarater and interrater reliability in subgingival calculus detection among the faculty of Parkland College Dental Hygiene program. Clinical trials compared the experimental group to the control group in detection of subgingival calculus on dental mannequins and patients before and after training using group standards, consensus, and feedback. Intrarater reliability was improved to a significant level and interrater reliability improved although not at a statistically significant level. Surveys distributed at the onset and end of the training examined faculty attitudes about clinical evaluation, consistency, and instruction.

## CHAPTER I

### Background Information

Clinical evaluation plays an essential role in the education of any health professional. It provides feedback to students, gives the faculty information about the success of instruction, certifies student competence, and assures the quality of health care provided to the patient. While educators agree that students must demonstrate an acceptable standard of clinical competence, agreement on what constitutes that standard varies. A survey of related literature in dentistry, medicine, and nursing points to difficulty among evaluators in setting an objective standard of clinical performance. Partridge and Mast (1978) state that clinical instruction has long been recognized as lacking in scientific and objective evaluation procedures. Eisner (1993) states that no evidence of quality of assessment in health education can be given without some type of standardization between examiners. In a time of consumer demand for accountability in health care, it is especially critical to have procedures for quality assurance in educational programs.

This study was conducted with the faculty of the dental hygiene program of Parkland College to determine if training could standardize and calibrate the evaluation of the clinical procedure of periodontal scaling. Periodontal scaling is the process by which calcified deposits are removed from the surfaces of teeth, and is the primary treatment performed by dental hygienists. Complete subgingival calculus removal is difficult to achieve, yet incomplete calculus



removal is often a factor in treatment failure in the resolution of periodontal diseases. The thoroughness of subgingival root surface debridement must be done by tactile sensation rather than visual inspection, and is evaluated by most clinicians with the use of some type of explorer. Stambaugh, Dragoo, Smith, & Carasali (1981) found that even highly skilled clinicians left substantial amounts of calculus on areas that had been designated as smooth by exploration. Research cited in the next section of this study repeatedly confirm that subgingival calculus detection is a very subjective procedure that varies widely in effectiveness. Consistency in grading students for detection of subgingival calculus is an ongoing problem and is a primary concern to the dental hygiene faculty at Parkland College.

Tied closely with difficulty in assessment of periodontal procedures are the intrinsic problems of any type of clinical evaluation. The faculty was concerned that specific performance criteria were not being utilized in a universal manner by all instructors and that interrater and intrarater reliability varied widely. Lack of rater reliability leads to evaluation procedures that are not valid according to Eisner (1993). Standardization of clinical evaluation has many benefits for students, faculty, and patients. With standardization, instructors have guidelines that enable them to be consistent in grading, students are ensured better uniformity in clinical assessment, and clinical effectiveness is improved. The American Dental Association (ADA) Curriculum Standard #5.14 requires that accredited dental hygiene programs have some type of calibration

program for clinical instruction, but provides no guidance as to the type of program to be used. Quality assurance and continuous quality improvement are major considerations in health care today. Quality is dependent on reliability, validity, and accurate analysis of testing criteria (Eisner, 1993). For these reasons, the faculty of dental hygiene instructors at Parkland College in Champaign, Illinois agreed to take part in a research project to test the effectiveness of clinical calibration of calculus detection. To study the effect of training on instructor consistency, the following research questions were addressed.

### Research Questions

1. To what degree does training enable a rater to increase intrarater reliability in calculus detection?
2. To what degree does training enable raters to increase interrater reliability in calculus detection?

### Statement of Research

Standardization of clinical evaluation is a desirable goal in providing quality instruction within a health profession program. Lacking guidelines, clinical instruction becomes subjective, rather than objective, and lacks validity. The faculty of Parkland College dental hygiene program expressed an interest in developing a calibration program for instructors since no program was readily available for use.

### Purpose of the Study

The study will explore existing calibration programs, current instructor evaluation consistency, methods of improving evaluation consistency, and the effectiveness of those methods. It will address the question of whether it is possible to create a program that enables instructors to more consistently agree on evaluation of clinical procedures. The issue of quality assurance in health care education will also be addressed.

### Significance of the Problem

The results of this study should prove significant to:

1. Dental hygiene students.
2. Dental hygiene educators.
3. Certifying agencies of dental hygiene programs.
4. Employers of graduates of the dental hygiene program.
5. Patients who receive treatment from dental hygiene students.

### Definitions

For the purpose of this study, the following working definitions are used:

Standardization - the training process that attempts to assure that more than one person, in more than one place, does the task in the same way as all other people engaged in that task.

Calibration - the process of training by which instructors are able to assess clinical procedures in an objective, reliable, and valid manner.

Interrater agreement - ability of two independent raters to duplicate evaluation results (reliability).

Intrarater agreement - ability of a rater to duplicate evaluation results (reliability).

Clinical competence - the ability of a clinician to perform a task within the parameters of accepted standards of that health profession.

Periodontal disease - inflammatory disease of the supporting structures of the teeth; those conditions that are treated by dental hygienists, dentists, and periodontists.

Scaling - the mechanical removal of calcified deposits (calculus) on the teeth, usually performed by the manual instrumentation of the clinician known as scaling.

Subgingival calculus - calcified deposits below the surface of the gum that are detected by tactile exploration with a dental instrument.

Supragingival calculus - calcified deposits above the gumline that are able to be detected visually or manually.

11/12 periodontal explorer - a dental instrument consisting of a very fine wire tip that provides tactile feedback when lightly applied to a surface.

Calculus detection - the ability of a clinician to discern the presence or absence of calculus on a tooth surface.

Internal criteria - subjective evaluation by participants.

External criteria - measures that address behavioral change in order to evaluate the effectiveness of the training program.

Typodont - model of human dentition that contains teeth and gingiva (gums) used to conduct intrarater testing.

False negative response - root surfaces of the tooth that were scored to be free of calculus when calculus actually was present.

False positive response - root surfaces of the tooth that were scored to have calculus where none actually existed.

Continuous quality improvement - a method of quality assurance which calls on practitioners and institutions to measure the quality of their services on a regular basis and make improvements where deficiencies exist.

#### Assumption of the Study

The following assumptions underlie this study:

1. Instructors participating will bring varying levels of expertise to the study.
2. Instructors participating will not agree initially on calculus parameters.
3. Clinical evaluation is able to be quantified by prescribed parameters.
4. Some variation in evaluation will continue to occur due to the individual nature of the evaluators.
5. Some increase in instructor reliability will be due to experimenter bias or the "Hawthorne effect".

#### Hypotheses

To further address the two research questions, two hypothesis were identified. All hypothesis in this study were tested in the null form,  $H_0$ . The null form assumes that there is no statistical difference between the means being compared.

Hypothesis 1 - there will be no significant improvement of intrarater reliability after calibration and training in calculus detection.

Hypothesis 2 - there will be no significant improvement of interrater reliability after calibration and training in calculus detection.

### Summary

This study evaluates the ability of experienced clinicians to consistently determine the presence of subgingival calculus, and the implications of poor reliability to clinical evaluation in a health professions program. Eleven dental hygiene faculty participated in a semester long research project that examined both interrater and intrarater reliability in detection of subgingival calculus. Training designed to improve reliability included continuous quality improvement techniques such as group defined standards, feedback, and consensus.

Chapter II presents a review of related literature on the topics of subgingival calculus detection, clinical evaluation in health professions programs, and integration of quality assurance into professional training and development. Chapter III outlines the research methodology used in the study. Chapter IV and V discuss the results of the study and the conclusions and recommendations of the researcher.

## CHAPTER II

## Review of Related Literature

This chapter presents a review of the literature as it relates to: (a) clinical evaluation in health professions educations, (b) the difficulty in detecting and removing subgingival calculus, and (c) implementing quality assurance measures in clinical education programs. These areas of study are integral to the background of the research project.

Clinical Evaluation in Health Professions Education

Clinical teaching is complex, involving many tasks that must be integrated into comprehensive patient care. Evaluation of clinical procedures is a problem throughout all types of health professions programs because it tends to be subjective rather than objective, often without specific criteria for judging success or failure. If an evaluation of a student's performance is reliable, the resulting evaluation score reflects the performance accurately (Marsick & Smedley, 1989). Unreliable measurements may base the student's score on factors that are unrelated to performance, and introduce errors into the evaluation (Pavlish, 1987). Educators agree that objectivity, reliability, and validity are important factors, yet these factors may be difficult to implement in clinical situations (Morganstein, 1990). Coates and Chambers (1992) state that information about assessment instruments that were known to be objective, reliable and valid was almost nonexistent in the nursing field, and this appeared to be true in all of the health professions researched.

In health professions such as dental hygiene, the students' clinical performance constitutes a major portion of the course work. Low reliability characterizes grading procedures and is a source of concern among educators and students. Feil (1982) maintains that the magnitude of measurement error is based on the amount of subjectivity in instructor evaluation of student performance and only components that limit the subjectivity will improve evaluator reliability. The issue of poor performance criteria recurs throughout the research. Partridge and Mast (1978) find that lack of well-defined performance criteria and poor rating or scoring systems lead to lower reliability, which in turn limit the validity of the evaluation. Stemmler (1986) confirmed that mechanisms to ensure that defined criteria of evaluation were applied consistently and equitably were necessary in order to assure valid and reliable results. Emmons (1979) agreed that measurement must be based on specific criteria for evaluation to be consistent from one rater to the next. Bazan and Seale (1982) had inconsistent results in developing performance criteria and rating scales during training sessions to increase instructor reliability in clinical dental instruction, although some increase in reliability did occur.

Clinical instructors tend to be "experts" in their clinical field, with definite ideas of what constitutes professional competence. Evaluation based on these ideas is not always applicable to good instructional methodology and may not be able to be assessed in measurable performance outcomes. Meetz, Bebeau, & Thoma (1988) found



that raters tend to have a narrow conception of adequate clinical performance, often ranking students to conform to their personal ideals of clinical competency. Meetz, et al. (1988) also found that clinicians had a preformed perception of the student's abilities that predicted the outcome of performance. Knowing the student well seemed to be a key to reliable rating, rather than specificity of rating criteria. Raters who were frequent observers were able to achieve a reliability rating of .44 to .61, compared to a rating of .36 to .42 for infrequent observers.

Tonesk (1986) states that personal judgement and educated guesses of the evaluator often determine the grade of a student, rather than measurable objectives. Barrows, Reed, & Moy (1987) found that most clinical teachers or examiners lack the confidence in clinical performance criteria to make critical judgements about clinical competence. A team approach to training faculty helps them understand their role as reliable evaluators (Stemmler, 1986).

Martin and Carey (1991) attributed evaluation inconsistencies to the complexity of the clinical situation, which included factors such as patient assessment and technical expertise. Mackenzie, Antonson, Weldy, Welsch, and Simpson (1982) listed sixteen factors that reduced rater agreement, among them unclear rules, faculty member's memory, unstandardized aids to judgement, inconsistent observational methodology, differences in ability, and differing tendencies toward leniency. One of the key factors was "checkpoint ambiguity", where faculty were unclear about which item to check on evaluations. Two

evaluators might check errors on different checkpoints for the same observed error. Tied to this type of rater error is faulty memory, when the rater does not remember where the item error should be noted and checks the wrong item. If a second examiner checks the right item, the student is penalized for two errors, when only one exists. Inadequate training in observation assessment and criteria also lead to problems in rater reliability.

Mackenzie et al. (1982) stated that important variables for effective training include clearly defined, unambiguous descriptions of grading criteria. Definitions need to be agreed upon by the examiners and each examiner should be recording and grading all dimensions of the same item in a set order each time. If observations are guided by using operational definitions (specific guidelines of what constitutes an acceptable procedure), rater errors can further be reduced. If a checklist can be made that is both reliable and valid, clinical evaluation is greatly facilitated.

Tonesk (1986) found that faculty become more competent in evaluating those narrow components of clinical skills that were most easily quantified. Clinical skills tend to be difficult to quantify, with characteristics such as professional judgement being hard to assess. Raters used their experience with other interns as standard for comparison (norm-referenced criterion), even when instructed to use criterion-referenced standards according to Stemmler (1986). Stemmler (1986) also found that faculty evaluations of clinical experiences were vague and indecisive, with general unwillingness to record negative

findings. This researcher also noted that there was a lack of understanding by faculty members of what was expected of them as clinical evaluators.

Meetz et al. (1988) reported evaluations were often skewed to the positive, with two-thirds to three-quarters of graduates rated as being above average compared to their peers. Faculty members felt they had not been adequately trained as evaluators, and were concerned with their inconsistencies. Faculty meetings of the dental hygiene program at Parkland have echoed these concerns. Abrahamowicz, Tamlyn, Ramsay, Klass, & Murray (1990) noted rater bias in the estimation of individual student ability. These researchers found that raters tended to be more lenient with poor students and demanding for good ones. The reviewed literature suggests that emphasis on clearly defined performance outcomes and expanded evaluator training would improve rater reliability in clinical instruction.

#### Rater Reliability in Calculus Detection

Attempts to estimate and improve interrater and intrarater reliability in calculus detection evaluation in dental hygiene treatment have not been particularly successful. Eisner (1993) states in an ideal clinical department, each faculty member strives to keep reliability above an .80 level of intra- and interrater reliability. Few studies exist that directly address the problem of standardizing grading of periodontal procedures, but those cited consistently show a problem with interrater reliability. Pippin and Feil (1992) report that examined consistency among raters in the detection of subgingival

calculus had moderate to poor reliability, with average reliability being only about 33 percent. Sherman, Hutchens, Jewson, Moriarty, Greco, & McFall (1990) found interexaminer and intraexaminer reproducibility in clinical examination was low. Biller and Kerber (1980) found evidence that 26 to 94 percent of the clinical grade given was due to instructor differences rather than student performance. They found that even experienced faculty members differed as much as 64 percent in calculus detection. Schoen (1992) found that even experienced clinicians were able to duplicate their findings only about 80 percent of the time. O'Hehir (1993) found interrater reliability to be as low as .25, even with experienced clinicians.

There appeared to be two distinct sets of factors that influenced these findings. The nature of the procedure itself lends itself to error. Most dental procedures can be evaluated both visually and by tactile sensation, but periodontal scaling procedures rely primarily on tactile sensitivity since the root surface is usually covered by gingiva. The thoroughness of subgingival root debridement is usually evaluated by use of an explorer, probe, or curette to provide tactile feedback on the smoothness of the surface. Sherman, Hutchen, Jewson, Moriarty, Greco, & McFall (1990) found that clinicians were accurate 72 percent of the time when they determined calculus was present, but 50 percent of the sites they deemed clinically acceptable displayed residual calculus. This means that if they determined a root surface was free of calculus, they were incorrect on half of the surfaces. Sherman et al. (1990) also found that the ability of the clinicians to

reproduce their own findings was very low, perhaps because microscopically the explorer tip used to evaluate root surfaces was larger than the remaining calculus. Pippin and Feil (1992) found that of the surfaces deemed acceptably treated, more than 57 percent of the surfaces had residual calculus. Stambaugh et al. (1981) observed that even clinicians who were deemed as being exceptionally skilled at calculus removal left areas with residual calculus in areas that they had designated as smooth.

Schoen (1992) found that no definitive instrument for evaluating root surfaces existed. Textbooks on periodontal instrumentation do not agree on the instrument that is most suited for exploring root surfaces. While most clinicians use some type of explorer, the types used vary widely from setting to setting. No studies could be found that substantiate recommending any one instrument. Educational settings tend to be dependent on the subjective opinions of the educators and clinicians and studies by Wilkins (1989) show that clinicians continue using whatever instrument was used during their training. Schoen (1992) attempted to determine if evaluation was affected by the use of different instruments. Slight differences were noted, but the primary factor in increasing tactile ability was attributed to clinical experience.

Sherman et al. (1990) found that perceptions of unsatisfactory clinical results increased with the severity of the periodontal pocketing. Microscopic examination of the root surfaces did not bear out these perceptions, as the amount of residual calculus was

relatively equally distributed among all sites. Sherman et al. (1990) questioned the usefulness of this type of traditional evaluation of root surfaces, and recommended that other clinical parameters be used to aid the detection of subgingival calculus.

Instructor preconceptions of the location where calculus was likely to be present influence evaluation. Pippin and Feil (1992) found differences in evaluator agreement according to the specific tooth surface and area of the mouth. Agreement among raters was highest in the interproximal molar areas at 25.3 percent and lowest for anterior teeth at 13.3 percent. The reviewed literature emphasized the difficulty in evaluation of a procedure that is dependent on tactile sensation.

#### Planning Quality Professional Development Programs

The importance of continuing professional development and methods of creating quality training programs have been the subject of several studies. Program planners for continuing professional education in health care education need to consider the whole climate of health care in today's culture. Cervero (1989) feels that the work of professionals is important not only because of their technical expertise, but also their power to define the problems with which they work. Symbolic leadership sets the context in which society sees the problem the professional can solve. This is significant because it gives the professional the power to dictate the conditions of service. Professionals are those who have a complete understanding of the complexity of issues relating to their field of expertise and the

ability to use it wisely. Health professions educators should be able to make accurate assessments of the correct use of power that their professional status allows. Future health care providers should be trained to do what is in the best interest of the patient, while providing the highest possible standard of care.

In considering the control that professionals exercise over the lives of others in society, Cervero (1989) discussed four viewpoints that should be considered when continuing professional education is planned: the functionalist viewpoint, the conflict viewpoint, the critical viewpoint, and the consumer-driven viewpoint. The functionalist viewpoint paints the professions as service- or community-oriented occupations that apply research based knowledge to relevant problems of society. Expertise is stressed, and the role of continuing education is to improve professional service by increasing knowledge, competence, or performance. It is assumed that there is general consensus on what is considered "good practice." The role of the educator is to help the professional stay current with their field.

The conflict viewpoint asserts that professionals are in competition with other societal groups for power, status, and money. Professionals use their power not to promote their expertise, but to dictate what people need. This is a condition that is currently exemplified by the struggle between physicians and nurses, dentists and dental hygienists, and other hierarchies or power within the health profession. The purpose of continuing professional education in this

perspective is to reduce the power of professionals to a more equal standing with other professionals or the client or patient.

The critical viewpoint evolved as a response to problems with the functionalist and conflict viewpoints. According to the critical viewpoint, professionals construct the problem from the situation. This approach stresses the need for professionals to be aware of conflicting values and the choices they impose. In this view, the ethical, political and technical aspects of the situation set professional standards.

Cervero's ideal philosophy integrates the functionalist, conflict, and the critical viewpoints into a fourth consumer-driven viewpoint based upon competence and expertise, rendered in a manner that allows the health professional to critically analyze the technical and ethical choices that must be made. The program planner for this type of professional development program provides data and moral and ethical guidance in the way it is to be utilized (Cervero, 1989).

The growth of the consumerism movement may dictate that health care providers and educators use the functional-conflict model in future practice. Kock and Fairley (1993) state that the 1990s mark a new era in focus in health care quality and that health care providers have been slow to respond. Marsick & Smedley (1989) report that the public will no longer allow health care professional autonomy and respect unless the needs of society are being met. They state that continuing professional education is a tool that allows a partnership of professionals, paraprofessionals, and consumers to address the



complexity of health care today. Only an educational model that does not allow one group of professionals to exercise undue power will satisfy an increasingly savvy public.

Marsick & Smedley (1989) reported that informed consumers will insist that the education of health professionals play an ever increasing role in raising the standard of practice that is expected. Providing quality programs that focus on the life-long needs of the professional is also vital.

Quality assurance is also determined by the group within a profession that will be responsible for providing continuing professional development. Cervero (1989) states that leaders of most professions believe that continuing education must be directed by its own members and that each of the professions feel themselves to be unique. In the task of educating their members they tend to use the same types of techniques and processes. It is possible for a program planner to look for guidance within the literature base to assure the quality of future programs. This concept allows the experiences of other fields of health care to be used in the development of programs. Professionals that are highly regarded within their fields tend to be technically competent, but may have little or no educational experience according to Berwick (1989). This is a common problem in health fields where the professional has prolific clinical credentials with minimal formal educational training. Instructors may have difficulty presenting their material in a way that is conducive to a high level of learning to the beginning clinician. To insure a quality educational

experience the program administrators must provide the instructor with the resources to facilitate learning. The quality of the program must be the primary consideration. Berwick (1989) suggest that although guiding a renowned clinician may be difficult, it is critical if the goals of the program are to be met. This is probably the area that will prove the most difficult for most health care educators although it is imperative to the quality of the program (Emmons, 1993).

Another factor to consider is the monetary cost to sponsoring institutions to develop quality educational program. Eisner (1993) states that quality programs will emphasize high levels of expertise among faculty, in spite of the cost in time and money for training. The challenge for the program planner is to produce training programs that achieve the desired outcomes without unacceptable costs for the sponsoring institutions. Failing to address the issue of quality of instruction threatens the credibility of the institution and could affect future enrollment. Quality educational programs will benefit the public by improving the standard of health care available (Berwick, 1989).

Another challenge to learning in training programs is the varying degrees of experience and competency participants bring into a program. According to related research, the program planner can do several things to increase the quality of the learning experience for this varied type of audience. Kemerer (1991) suggests that six factors inhibit learning transfer: (a) lack of clarity, (b) lack of focus on knowledge, (c) behavior, (d) poor timing, (e) unrealistic expectations

or changing expectations, and (f) not taking ownership of the material. Additional factors that can contribute to the success of the learning experience include clarifying learning objectives, providing realistic expectations of what can be learned, focusing on the behavior or skill to be attained, and acknowledging that the learner must be ultimately responsible for his/her own learning (Kemerer, 1991).

Other quality assurance measures include providing an environment in which the special needs of adult learners are considered. Apps (1981) states that adult learners tend to be self-directed and wish the instructor to guide rather than command. Learning best occurs when adults are allowed to draw on life experiences in the context of the subject content, and new knowledge and skills should be applicable to their everyday work experiences (Johnson, 1991). In an earlier study, Apps (1979) reports that adults who are participating in voluntary programs expect amenities such as good audio-visual aids and quality handouts. Apps (1979) also states that professional interaction should always be allowed in the schedule since sharing of professional information is an integral component in the development of quality programs. All of the factors discussed above are critical for the assurance of quality within a program.

The professional is an adult learner with characteristics including life experience, enthusiasm, and a commitment to learning (Apps, 1981). Learning is often shaped by past experiences and this may add to the resistance to new ideas. Bennett and LeGrand (1990) report that differing life experiences due to age may actually enhance

the quality of a program by providing interesting diversity in viewpoints during discussion. Participant involvement can be a valuable addition to the educational experience for the professional. Since attendance is usually voluntary, the participant may have expectations of what needs to be learned and is willing to make the effort to master the material. Active involvement increases learning effectiveness for the adult and methods such as group discussion, role-playing, and question and answer sessions with the opportunity for feedback should be employed. Professionals will need a clear image of the intended change and why it is beneficial. Since one of the goals of adult learning is to help people make changes that will update skills and knowledge, it is critical that the change needed is clearly identified (Bennett and LeGrand, 1990).

An additional factor to consider is climate setting. Gilley and Eggland (1989) list four learning climates that are common in adult education and training. These include: a) friendly learner-centered, where learners set their own goals in a supportive environment, b) friendly traditional course, where the climate is supportive and the instructor sets the goals, c) nasty traditional, where the instructor sets the goals in an unfriendly environment, d) and sour T-group, learners set the goal, but unfriendly feelings prevail. These researchers stress the importance of building the relationship process between learners and instructors. Sisco (1991) feels that creating a positive learning environment is especially important for the adult

learner. Examples of this type of environment would include clinical sessions where the participants are able to see the technique being discussed, then receive actual "hands-on" instruction when new techniques are being learned. The integration of the special needs of the adult learner, and the unique considerations of continuing professional education and training can result in a quality training program.

#### Establishment of Objectives

As Bennett and LeGrand (1989) noted, professionals need clear objectives in order to be motivated to change behavior. The establishment of objectives as the outset of the training program can guide learning, instruction, and evaluation (Johnson, 1991). Kozoll (1992) reports that there are two levels of objectives in any planned program. The implicit objectives are unstated, but contribute to the success or status of the sponsoring organization. The explicit objectives are those that are specifically outlined to the learner. In a 1978 study, Apps states that behavioral objectives are often used because they serve as both a guide for organizing learning activities and then as a basis of measurement for the results of those activities. Boyle (1981) breaks behavioral objectives into three classifications - the cognitive, affective, and psychomotor. The cognitive domain is concerned with development of intellectual skills and knowledge. The affective domain describes attitudes and interests. The psychomotor domain involves physical and motor skills. While most behavioral objectives will deal with the cognitive dimension, some

affective and psychomotor elements will be incorporated. Although the objectives are different for each training program, they have several characteristics that make them meaningful (Boyle, 1981). The objectives should indicate what the participant can expect to achieve through the program and deal with things that are important in maintaining professional competency. These objectives should be attainable during the program period and focus on what is most crucial to the program. The objectives should be clear and specific enough that it is possible to determine if they have been met. The result is clearly identifiable. The participant should be able to achieve the stated objectives through careful use of learning activities (Houle, 1976).

Gilley and Eggland (1989) describe a specific learning objective as one that precisely describes knowledge or behavior that occurs as a result of the learning activity. The well-written learning objective should a) identify the type of learning the program wishes to accomplish, b) describe an observable behavior that will demonstrate that learning has occurred, c) identify an acceptable level of performance for the learned behavior, and d) describe the condition under which the performance will be measured.

### Evaluation

It is possible to establish how well objectives have been met through the use of evaluation. Laird (1985) feels that only one part of evaluation criteria is met by achievement of learning objectives,

the other two aspects are contributions to organization goals (implicit objectives) and the perceptions of the participants.

The use of the appropriate evaluation instrument is critical to accurate measurement. Miller (1990) defines two types of evaluation: formative and summative. Formative evaluation uses assessment during the instructional process to identify inadequacies and competencies in skills, knowledge, or attitudes. This information is utilized to influence instruction during the instructional sequence. Summative evaluation is the assessment at the conclusion of instruction. The summative evaluation determines the extent of achievement of the established objectives for which the instructional program was designed.

Nowlen (1988) reports that the most common use of evaluation outside the field of continuing education is in: a) diagnostic, formative, and summative judgements of learning, b) assessment of skills and knowledge prior to learning activities, c) providing feedback on learner progress, and d) measuring achievement. Within performance-based continuing education programs, formative evaluation of learning and teaching, and summative evaluation of learning occur in the context of performance improvements the learning activity is expected to make (Nowlen, 1988).

Formal evaluation instruments such as surveys for participants are the most common form of program evaluation and are usually considered the most important. Evaluation instruments must be "user-friendly" enough that participants are willing to complete them, yet

provide useful data. A well written evaluation gives the program planner information to improve future programs. Areas that should be included are program content, speaker effectiveness, facilities, administration, and expectations of the participant (Caffarella, 1988).

### Summary

The review of related literature in this chapter has dealt with the topics of clinical instruction, subgingival calculus detection, quality assurance techniques in training, and instructional methodology as they related to the study. Literature reviewed confirmed the difficulties of establishing interrater reliability in all areas of clinical instruction and in subgingival calculus detection specifically. Research reviewed stressed the low levels of reliability that exist in all types of clinical evaluation, with subgingival calculus detection exemplifying the types of rater error that occurs.

Low reliability affects the ability of clinical instructors to provide quality instruction. Reviewed literature stressed the need for the development of accurate, reliable and valid measurement instruments. The effect on continuous quality improved training on clinical evaluation was examined with special emphasis on the special educational needs of the adult professional.

Numerous studies cite the differences that must be considered in development of learning settings, using specific examples of training objectives and evaluation procedures.

Providers of health education are under increasing pressure to produce evidence of performance outcomes. The reviewed literature



examined the various components that should be addressed to reduce rater errors and to increase the level of quality within clinical instruction.

### CHAPTER III

#### Methodology

##### Overview

This chapter describes the methods and procedures that were used in this study. Descriptions of the design of the study, selection of subjects and patients, and methods and materials will be provided. The sample size and data analysis, and selection and development of research instruments will be discussed.

##### Design of the Study

The research design for this project was experimental. Data was collected for participants for the following trials: a) intrarater reliability on dental mannequins, b) intrarater reliability in clinical trials, c) interrater reliability in clinical trials, d) control reliability in clinical trials, e) pre-training attitudes, and f) post-training attitudes.

A t-test with a dependent variable was used for comparison to the experimental and control group, with the dependent variable being the level of skill each examiner brought to the research project. A Pearson's coefficient correlation for team scores was calculated before and after training. A survey of participant satisfaction compared attitudes before and after the training exercise using percentage as the basis for comparison.

Selection of Patients and Examiners

Ten patients, ages sixteen to fifty-four, with varying degrees of periodontal health and amounts of calculus, were selected for the study. Prior to any treatment procedures, each patient completed a comprehensive medical history form and signed a consent form (Appendix E) approved by the Parkland Dental Hygiene program. Patients with systemic disorders were not included in the study.

The examiners who volunteered to take part in the training exercise were clinical instructors in the dental hygiene program at Parkland College. All of the examiners were registered dental hygienists who varied in clinical experience from six to more than 20 years in practice. All had been selected for their instructor positions because of a high level of clinical expertise, so all could be considered experienced clinicians. Seven of the 10 faculty members had worked together for over eight years, two of the faculty had three years experience with the group, and one member was a new faculty member. In addition to their clinical background, all had expressed a strong commitment to improving the quality of instruction within the dental hygiene program.

Methods and Materials

A faculty meeting of the dental hygiene instructors at Parkland College in Champaign, Illinois was held at the initiation of the research project. Instructors were asked for their ideas on program design and for desired product outcomes. The number of training sessions and times scheduled was determined by group consensus. Since

participation would be voluntary, it was decided that all results would be coded by random number so that each individual's identity would be protected. This was important to insure randomness of the control versus the experimental group and to reduce competitiveness within the group. Standards were set by asking faculty for their input on desired goals for performance outcomes. This group agreed that the focus of the training would be improvement of the consistency of instructor evaluation of calculus detection. Implicit objectives included the assurance of continuous quality improvement within the dental hygiene program, and explicit objectives included increasing the level of intrarater and interrater reliability. Suggestions made were incorporated into the program, if it was at all possible. Suggestions were also noted for future use in developing this type of project. Pretest and post-test surveys assessed the attitudes of the group and their satisfaction with clinical evaluation and training. Two processes were used to increase the validity of these surveys. A review of the literature suggested recurring concerns in clinical evaluation. The second process included gathering information about evaluation procedures from the Parkland faculty. The two sources of information were used to construct the pre-training and post-training survey instruments.

#### Intrarater Reliability

Intrarater reliability was tested on six faculty members of the dental hygiene program of Parkland College in Champaign, Illinois. A dependent t-test was used to analyze the pre-training and post-training

results, with the level of examiner skill being the variable. Two "Columbia dentiform" models of human dentition were prepared by removing the teeth and placing various mixtures which simulated calculus on the root surfaces (Appendix D). Mixtures used included epoxy resin and salt, poppy seed and cyanoacrylate, and commercial finger nail polish and pumice. This variety of materials was used to simulate the different textures of subgingival calculus that occur on human teeth.

Once the mixtures had dried, the teeth were replaced in the models, with the calculus now only detectable by tactile exploration. Examiners chose one of the models, were given a dental chart (see Appendix D) and an American Eagle XCP11-12 periodontal explorer, and instructed to mark any area that they felt was rough on the chart. This was designated as the pretest and was scored by comparing the number of surfaces marked to the actual number of surfaces where calculus had been placed. False negative scores (areas marked as being free of calculus where it existed), and false positive scores (areas marked as having calculus where none existed), were also scored.

Participants repeated the procedure using the same model one week later, after participating in the first training exercise, and the score was compared to determine intrarater reliability. Once reliability had been scored, the participants were allowed to see the master answer sheet, remove the teeth, and go over any area that was missed. Any instructor who wished to repeat the test could work until

they achieved the reliability score they personally wanted to achieve, and four of the instructors worked to increase their reliability to their desired goals.

Intrarater reliability was also tested during the clinical exercises. Statistics of the ten individuals were kept as they worked within their teams. Individual reliability was analyzed using a dependent t-test to compare accuracy of calculus detection from the first trial (pre-training), and the post-training trial.

#### Interrater Reliability

Examiners were given a survey (Appendix A) to complete on their perception of the level of interrater reliability that existed before the calibration program, and to examine their feelings, concerns and instructional goals concerning clinical evaluation. This survey was taken at the beginning of the first clinical exercise. Each survey distributed to the group had a number chosen from a table of random numbers at the top of the page. Four of the numbers were preceded with a "C", and the people who chose these surveys became the control group. The remaining six instructors were to use the first two numbers of the sequence as their identifying number throughout the clinical trials.

For the first clinical exercise held on April 15, 1993, five patient volunteers from the community were examined by a designated "expert clinician" for the presence of subgingival calculus. The clinician was selected because of intensive training in calculus detection and over 20 years of clinical experience. The presence or absence of subgingival calculus was detected by using an American Eagle

XCP 11/12 periodontal explorer to explore the root surface of every tooth of every patient. The findings were dictated to the clinical dentist as each tooth surface was explored and the dentist marked each area that was determined to be rough on the master scoresheet used to calculate reliability. The standard was any area of roughness that an instructor judged could be made smoother with further scaling or instrumentation. This is the grading criteria that is used when calculating dental hygiene students' final scaling score. Roughness due to the anatomy of the tooth or that would not be counted as a student error was not to be marked. Charting of calculus was recorded as a heavy "dot" on the representative area of a geographic chart of dentition using three concentric circles to designate which clinician was charting (see Appendix F). This form is based on a form used by the Central Regional Examination Testing Agency, a certifying agency for dentists and dental hygienists. The outermost circle was marked by the first team examiner, and the innermost circle represented the "expert" examiner. Use of a three-part form with a carbon between the second and third sheets allowed each clinician to record their findings without being able to see what the other had marked. The circles were divided into fourths, with each fourth representing the mesial, facial, lingual, and distal tooth surfaces. The lines dividing the circles themselves were representative of the line angles of the tooth, for example, mesial-lingual, distal-lingual, mesial-facial, and distal-facial. On each tooth were eight possible sites where calculus could be marked as a dot.

The four instructors designated as the control group explored the lower right quadrant of their assigned patient, and marked any area that they felt was rough according to the criteria described above on the chart that is included in the appendix. This score was designated the baseline score for their reliability of the examiners participating in training exercises. The control group was then finished with the first clinical exercise, but were allowed to remain if they wished to help the other clinicians chart their findings. They were asked not to take part in the team comparisons or listen to the interplay that took place between the teams. They were to chart for each examiner, then move on to the next examiner team, before any team comparisons occurred.

The clinicians who were taking part in the research project were paired into three teams of two examiners. The first teams explored the upper right quadrant of dentition on each patient. The first team member (examiner A) marked any area of roughness on the geographic representation chart on the outside, unshaded circle with a dot indicating areas of calculus. Once the first team member was finished, the top sheet of the grading form was removed, and the second examiners marked his/her findings. The team then compared findings, circling areas that they both agreed were rough. These areas were designated as true calculus. If an area was marked by one team member, but not another, the team members worked together until a consensus had been reached. Team reliability was calculated by comparing the number of sites that the team agreed upon with the number of sites in the quadrant. The team members were instructed to help each other by



offering tips and techniques that had been useful to them. Once a consensus had been reached, the teams dissolved, and new teams were formed. The second team explored the upper left quadrant of dentition. Feedback, advice, and group standards were stressed. The reliability of the teams was measured by comparing their scores from the first exercise to the second clinical exercise using a t-test with a dependent variable (the level of skill of each examiner). Pearson's coefficient correlation was also calculated for team scores.

Individual reliability of each team member was calculated by comparing the number of false negatives (surfaces that had been scored as being free of calculus where it actually was present), and false positives (surfaces that were scored as having calculus that were actually free of calculus) to the findings of the expert clinician.

The second clinical exercise was held two weeks later, on April 19, 1993. Before the exercise began, a roundtable meeting was held to discuss the group's feeling about the first exercise. Discussion about definitions of calculus, and the type of criteria to be used occurred. The group defined the standard that would be used in marking calculus in contact areas and along the gingival margin. Concerns and suggestions for improvement were shared with the group. During the second clinical exercise, the same procedure from the first exercise was repeated, using five different patients. The control group again recorded their findings of the lower left quadrant to determine reliability scores. The experimental group used the same team approach as the first trial, but with some differences. Although the original

teams from the first exercise were preserved, a slightly different procedure was followed. Before either examiner in the team explored their assigned quadrant, the team selected one tooth from one of the other quadrants, and worked together to completely agree on the presence or absence of calculus on that tooth. The team then each independently recorded their findings in the assigned quadrant. Once both team members recorded their findings, those teams dissolved, and the second team was formed. The second team repeated the process described above, as did the third team. A second survey (Appendix B) was distributed to determine participants' perceptions of the process.

#### Data Analysis

Intrarater reliability was determined by computing the number of sites where presence or absence of calculus was agreed upon by the model key and the examiner (percentage of agreement). The number of false negative sites marked was compared with the actual number of sites that were free of calculus, as were the number of false positive sites marked compared to the number of sites where calculus had been planted.

Interrater reliability used the same type of computation for each individual team member and then added the team's agreement with each other to the data. This determined the team's reliability rating for each trial. Each team performed one trial before training at the first session, and one trial after training at the second session.

Surveys (Appendix A and B) administered before and after the clinical exercises were used to determine pre-training and post-

training attitudes and satisfaction levels. Questions concerning effectiveness of instruction, faculty consistency in clinical evaluation, and the general level of instructional consistency were addressed. Pre-test and post-test surveys were coded with examiner numbers so the same examiner's answers could be compared.

#### Summary

This chapter described the methods and procedures that were used in the study. The design of the study, selection of patients and subjects, and methods and materials used were discussed.

## Chapter IV

## Results of the Study

Overview

This chapter presents the findings of the study as they relate to the two research questions outlined in Chapter I. Findings have been organized according to study demographics, intrarater and interrater reliability and survey results.

To answer the research questions, examiners were placed in various situations where their calculus detection accuracy was tested and compared to the control group, the expert clinician, their team members, and to their pre-training scores. A Pearson's correlation coefficient of .8245 was calculated for the first team trial, and the post-training coefficient was .8915. The results of the study revealed a significant reliability improvement in percentage of agreement from .8958 to .9182 ( $t$  value -4.04),  $p < .05$  for intrarater scores using the typodont model. Clinical trials on patients showed an improvement in percentage of agreement for intrarater reliability for the experimental group from .8482 to .8819 ( $t$  value -.00), although it was not statistically significant at  $p < .05$ . The control group also improved in percentage of agreement reliability from .8199 to .8824 ( $t$  value -59), which was not statistically significant at the  $p < .05$  level. The interrater teams improved from .8567 to .8631 in their agreement with each other before and after training, which was not significant at the  $p < .05$  level.

Comparison of the survey on attitudes concerning effectiveness of instruction, faculty consistency in clinical evaluation, and the effectiveness of the training showed an improvement in instructor perception of consistency after the training sessions.

### Demographics of the Sample

All eleven of the dental hygiene faculty of Parkland College participated in the study. All instructors were female with eight of the eleven being part-time faculty.

### Research Findings for Research Question One

To what degree does training enable a rater to increase intrarater reliability in calculus detection?

The hypotheses for this research question stated in the null form was: There will be no significant improvement of intrarater reliability after calibration and training in calculus detection.

Hypotheses one was rejected. Improvement of intrarater reliability by percentage of agreement from .8958 pre-training to .9182 post-training was statistically significant level at  $p < .05$ , when the raters were detecting calculus on typodont models. Improvement in agreement from .8482 pre-training to .8819 post-training occurred when raters worked with patients, but the improvement was not significant at the  $p < .05$  level. Table 1 summarizes the findings for intrarater reliability using typodonts. Trial One occurred before training, Trial Two occurred after training.

Table 1

Intrarater reliability using Typodonts

<u>Trial</u>	<u>Number of Raters</u>	<u>Mean</u>	<u>Standard Deviation</u>	<u>Standard Error</u>
1	6	.8958	.016	.006
2	6	.9182	.013	.005

The t value for this trial was -4.04, with the 2-tail correlation probability being .569 and .239. The degrees of freedom were five.

Intrarater Reliability within Teams Compared to Controls

When compared to the control group, the experimental group did not significantly improve intrarater reliability from pre-training clinical trials to post-training clinical trials. Data is provided in Table Two.

Table 2

Intrarater Reliability Within Teams


---

<u>Pre-Training clinical trial</u> (Percentage of Agreement)				
Group	Number	Mean	Standard Deviation	Standard Error
Control	4	.8199	.090	.045
Experimental	18	.8482	.089	.021

---

<u>Post-training clinical trial</u>				
Group	Number	Mean	Standard Deviation	Standard Error
Control	4	.8824	.149	.075
Experimental	18	.8819	.254	.060

---

The t value was -.57 with a 2-tailed probability of .574 and .597. The degrees of freedom were 4.43 .

---

Although improvement in intrarater reliability occurred after training in clinical trials, the improvement shown was not statistically significant.

Findings for Research Question Two

To what degree does training enable raters to increase interrater reliability in calculus detection?

The hypotheses for this research question in the null form was:

There will be no significant improvement of interrater reliability after calibration and training in calculus detection.

Hypotheses two was accepted. The improvement of the experimental group from .8567 pre-training reliability to .8631 post-training in interrater reliability during clinical trials was not significant at the  $p < .05$  level. The control group improved from .8199 during the first session to .8824 for the second session, although this improvement was not significant at the  $p < .05$  level. Table Three presents the data for interrater reliability before and after training in clinical trials.

Table Three

Interrater Reliability for Teams in Clinicals

<u>Trial</u>	<u>Number</u>	<u>Mean</u>	<u>Standard Deviation</u>
Pre-training	9	.8567	.0761
Post-training	9	.8631	.2861

The Pearson Correlation Coefficient for pre-training was .8245 and post-training was .8915.

Although the interrater teams showed improvement after training in the clinical trials, it was not at a level to be statistically significant if  $p < .05$ .



A Comparison of Types of Rater Errors

Errors in calculus detection can be errors of omission or false negative errors when a tooth surface is designated as clean when calculus still remains. The other error in calculus detection is the error of commission or false positive errors. These errors occur when the examiner designates a root surface as having calculus when none exists. The reviewed literature states that false negative errors are far more common than false positive errors. Calculation of the mean of false positive and false negative scores verified this trend.

During the intrarater pre-test on the typodonts, the mean score for false negative site errors was 26.3 percent, compared to only three percent site errors for false positives. This meant that out of 38 possible sites with calculus, the mean number of sites missed due to omission was 7.6. The number of possible sites without calculus was 102, and the mean of sites missed due to omission was four. The mean number of false negative errors decreased after training to 17 percent and false positive errors dropped to two percent. False negative errors predominated the team errors during interrater training with patients. Pre-training errors had a mean false negative score of 31 percent, with false positive errors at 7.2 percent. The mean of post-training false positive errors was 45 and of false positives, nine.

A Survey of Faculty Attitudes Before and After Training

A survey (Appendix A) was distributed at the onset of the project to the 11 instructors who participated and a similar survey (Appendix B) was distributed at the end of the project. The survey asked questions concerning perceptions on instructor consistency, quality of evaluation, grading criteria, and confidence in evaluation skills. Questions were to be rated on a scale of one to 10, with a one response representing strongly disagree and a 10 response representing strongly agree. Responses were grouped in increments of three (e.g., a response of one, two or three, was designated as disagree) for ease of calculation. Many of the questions were repeated on the post-training survey so that attitude changes could be compared. An item-by-item analysis of each question for the pre-training and post-training survey is provided in Appendix C.

Questions one, two, four, and seven compared participant perceptions of consistency and confidence in evaluation before and after training. Question one, on both surveys, asked instructors to rate their consistency compared to other instructors in the program. This question showed an increase in the perception of instructor consistency from 18 percent agreement to 73 percent agreement after training. Question two asked the instructors to rate consistency of calculus detection. Responses after training showed that agreement in the perception of consistency increased from 27 percent to 45 percent.

Question four indicated a reduction in confidence that instructors were accurately assessing the student's clinical skills occurred after training. When asked before training, 82 percent of the faculty agreed that they were accurately assessing clinical skills. After training, the confidence level dropped to 18 percent. Question seven, on both surveys, asked instructors to rate their abilities as a clinical evaluator. Again, individual confidence fell from 100 percent to 91 percent after training.

Table Four

Participant Perceptions of Consistency and Confidence

Item	Responses			
	Pre-Training		Post-Training	
	Neutral	Agree	Neutral	Agree
N = 11				
1. Personal consistency compared to other program instructors	82%	18%	27%	73%
2. Calculus detection consistent between faculty members (pre-training only)	73%	27%		
2. Calculus detection consistency improved due to training participation (post-training only)			55%	45%
4. Personal confidence in assessment accuracy	18%	82%	82%	18%
7. Confidence in ability as a clinical evaluator	0%	100%	9%	91%

Questions three and six discussed factors that influenced faculty grading patterns. Question three on the pre-training survey asked instructors to agree or disagree with the statement that extraneous factors such as patient management, skill level or personality were factored into the grade given for calculus removal. Forty-five percent of the faculty disagreed, 36 percent were neutral, and only 18 percent agreed that purely objective criteria were used to calculate the student's scores. The instructors were then asked after the training exercises if participation in the calibration project would influence their grading patterns. Twenty-seven percent disagreed that they would change their grading patterns, 55 percent were neutral, and 18 percent agreed that participating in the training would change their grading.

Question six, on both surveys, was concerned with instructor perception of how well prepared they were to evaluate clinical skills. Before training, 18 percent of the faculty disagreed with the statement, 27 percent were neutral, and 55 percent agreed with the statement. The instructors were asked after training if the training exercises had made them better prepared to evaluate clinical skills. Almost half (45 percent) agreed that training had improved their preparedness with the remaining 55 percent being neutral.

Table Five

Factors Influencing Grading Patterns

Item	Responses	
	Pre-Training	Post-Training
N = 11		
3. Extraneous factors reflected in calculus grade (pre-training only)		
Disagree	45%	
Neutral	36%	
Agree	18%	
3. This training exercise will influence grading (post-training only)		
Disagree		27%
Neutral		55%
Agree		18%
6. Instructor well prepared to evaluate clinical skills		
Disagree	18%	0%
Neutral	27%	55%
Agree	55%	45%

When the faculty was questioned on the consistency of clinical evaluation criteria within the program, responses indicated that agreement on universal criteria were being utilized increased from 46 to 55 percent after training. Linked to this was the perception that training exercises had increased understanding of other faculty members evaluation criteria. After training, 91 percent of the faculty indicated that mutual understanding had increased as a result of the exercises.

Table Six

Knowledge of Evaluation Criteria

Item	Responses			
	Pre-Training		Post-Training	
	Neutral	Agree	Neutral	Agree
N = 11				
8. Universal evaluation criteria are being used within the program	55%	45%	45%	55%
11. Training has increased my understanding of other faculty member's evaluation criteria (post-training only)			9%	91%

When asked about the need for more clinical evaluation training, 91 percent of the faculty indicated that they felt the need for more training in clinical evaluation, and 73 percent indicated that they felt the training exercises had been useful. When asked if they would care to participate in similar types of programs in the future, 82 percent indicated that they would be interested in collaborating in this type of project. Ninety-one percent of the faculty agreed that training exercises improved the caliber of instruction within a health profession program.

Table Seven

Participant Perceptions About Training After Calibration

Item	Responses	
	Post-Training Only	
	Neutral	Agree
N = 11		
9. These training exercises were useful to me	27%	73%
10. I would like to participate in similar programs in future	18%	82%
12. Exercises like this improve the caliber of instruction in health professions programs	9%	91%



Previous studies indicated that the fear of "peer review" is an impediment to the initiation of many calibration programs. When the faculty were asked before the training program if they were threatened by the peer review process, nine percent agreed that they were threatened, 55 percent disagreed, and 36 percent were neutral.

The surveys showed an improvement in confidence in evaluation consistency after training for every question but question four. An improvement in the perception of evaluator consistency after training was noted in many of the questions. The value of the training program was indicated by many of the responses.

### Summary

This chapter described the findings of the data analysis for each of the research questions. The research findings revealed an improvement in reliability for both intrarater and interrater reliability, with the results of the intrarater exercises on the typodont showing significant improvement. The control group improved in reliability as well as the experimental group, although not at a significant level.

The improvements that were the most dramatic were in attitudes and perceptions about instructor consistency in calculus detection. Agreement about level of consistency improved on the post-training survey for ten of the eleven questions. A discussion of the implications of these findings will be included in Chapter V.

## Chapter V

## Summary, Discussion, Conclusions, and Recommendations

Summary

Reliable evaluation of clinical procedures within a health professions program is important because it provides feedback to students, gives faculty information on the effectiveness of instruction, certifies student competence to regulatory agencies, and assures the quality of health care provided. The difficulties in maintaining objectivity during evaluation of clinical procedures have been documented by many researchers. Studies examining reliability within clinical evaluation of health professions find it to be low (Partridge & Mast, 1978; Feil, 1982; Mackenzie et al., 1982). Low reliability is often linked to a lack of specific performance criteria and failure to utilize standardized performance objectives. Limiting subjectivity through such techniques as developing unambiguous grading criteria and operational definitions will reduce subjectivity and reduce rater errors (Mackenzie et al., 1982). Eisner (1993) stated that quality in health care education cannot be assured unless evaluation procedures are reliable and valid.

The purpose of this study was to take one component of periodontal treatment, subgingival calculus detection, and to develop specific grading criteria to reduce evaluator subjectivity. Subgingival calculus detection is evaluated during dental hygiene clinical training, and standardizing or calibrating examiner interrater and intrarater reliability for this procedure using continuous quality

training methods was the goal of the training exercise. Because subgingival calculus detection is dependent on the tactile sensitivity of the clinician, conventional wisdom in dental hygiene education implies that low reliability is intrinsic due to the individual personalities and skills of raters. Research reviewed confirmed that interrater and intrarater reliability for this procedure are extremely low (Stambaugh et al., 1981; O'Hehir, 1993; Biller & Kerber, 1990). This research took principles of continuous quality improvement and adult training and applied them to the procedure of calibrating subgingival calculus detection among the faculty of the Parkland College dental hygiene program.

The purpose of the study was to specifically answer the research questions:

1. To what degree does training enable a rater to increase intrarater reliability in subgingival calculus detection?
2. To what degree does training enable a rater to increase interrater reliability in subgingival calculus detection?

The dependent variable was the level of skill each examiner brought to the research project.

Data from eleven faculty members was used for the final analysis. The experimental group was tested and compared to the control group for several types of reliability: intrarater reliability on models of teeth (typodonts) and on human patients, and interrater reliability on human patients. Surveys distributed pre-training and post-training compiled

faculty attitudes on instructor consistency, evaluation methods, and instructional effectiveness.

### Results and Discussion

The most striking result of the study was the perception of improvement of consistency of the faculty after the training project. When asked to rate their consistency of calculus detection compared to other instructors in the program, training improved the perception of consistency from 18 to 73 percent. The perception of consistency in calculus detection between all faculty improved from 27 to 45 percent after training. Evaluation and grading were examined in the third question on both surveys. Almost half of the faculty stated that factors other than absence or presence of calculus were considered during the grading procedure. This introduced subjective factors into the evaluation procedure and lowers the reliability among raters. This is consistent with researchers' findings that factors such as personality are often introduced into the evaluation of clinical procedures (Meetz et al., 1988; Feil, 1982). Awareness of the importance of standardization of the evaluation process emphasized during training resulted in an 18 percent agreement that grading criteria would be changed as a result of the training exercise.

Reduction in confidence that instructors were accurately assessing clinical skills occurred after the training. Before training, 82 percent of the faculty agreed that they were accurately assessing clinical skills of the student. After training, the confidence level dropped to 18 percent. Perhaps increased awareness of

the importance of specific performance outcomes and reduction in emphasis on factors such as patient management and personality could account for this drop in confidence. Confidence that interrater reliability was improved was indicated by a nine percent increase in positive responses after training. Instructors were made aware of each other's criteria in defining subgingival calculus during the training session and this could account for the increased confidence in interrater reliability.

Before training, 55 percent of the faculty felt well prepared to evaluate clinical skills, 27 percent felt neutral, and 18 percent did not feel adequately prepared for clinical skills evaluation. With only slightly more than half of the faculty stating that they felt well prepared for clinical evaluation, training would seem to be indicated. After participating in the training exercises, 45 percent indicated that they felt better prepared to perform clinical evaluation. This validates the importance of training for clinical evaluation.

Although only about half of the faculty felt well prepared to perform clinical evaluation, 100 percent stated that they felt confident in their abilities as a clinical evaluator before training occurred. There was a nine percent decrease in confidence after training. Perhaps this could be attributed to the increased awareness of the complexity of providing a reliable and valid evaluation.

Slightly less than half (45 percent) of the faculty felt that the same set of criteria was utilized by everyone in clinical evaluation. After training, the perception of consistency was

increased to 55 percent. Emphasis on standardizing evaluation procedures could account for this increase.

When asked before the training sessions if they would like more training in clinical evaluation, nine percent disagreed, 45 percent were neutral, and 45 percent wished to have more training. After training had occurred, 73 percent stated that the training had been helpful, 27 percent were neutral, and no one stated that it was not helpful. This verifies the initial resistance that can occur in adult training programs, especially in programs where the workers are experienced and consider themselves good at their jobs. The resistance needed to be overcome in order for the training to be perceived as helpful. It is often difficult for adult learners to put themselves in a position where they are vulnerable to the judgements of others (Bennett & LeGrand, 1990). This appeared to be confirmed by the responses of participants in this research project. After the project was completed, many of the faculty commented to the researcher that they had "dreaded" participating in the project because of judgement by peers, but that their fears were not realized due to the spirit of collaboration and teamwork.

These fears are linked to the process of "peer review" which is the process by which professionals are judged by other professionals, usually in a negative way. Health professionals tend to be confident of their own skills and may be unwilling to expose themselves to criticism from others. When asked if they felt threatened by the process of peer review, 55 percent said they were not threatened, 36

percent were neutral, and nine percent said the process was threatening. In discussion with individual faculty members, however, concerns were expressed about being teamed with other faculty who were feared to be judgmental. Due to the randomness of the selection process, teams were formed that included team members who had expressed concerns, yet the process was declared a positive experience by everyone involved. The emphasis on cooperation and information sharing may have prevented anticipated problems with professional egos. At the beginning of each session, instructors were asked to share techniques and tips with their teammates and to collaborate on effective detection strategies. Learning was structured to occur in a "friendly learn-centered environment" (Gilley & Egglund, 1989). Instead of judging a team member right or wrong, consensus had to be reached by both team members using the group standards that had been agreed on. This eliminated many of the individual variances that normally occur in clinical evaluation. Clearly defined performance criteria provided a goal for the desired outcome. Since the goal had been established by the group (improved rater reliability), the need for change had been clearly identified, a vital component to professional education (Bennett & LeGrand, 1990).

After training, faculty members were asked if they would like to participate in similar types of training programs in the future. Eighty-two percent said they would like to participate in future programs and 18 percent were neutral. Participants were also asked after training if the exercises had given them a better understanding

of the other faculty member's evaluation criteria. Ninety-one percent said that they had a better understanding of the groups' standards and nine percent were neutral. When questioned after training as to whether training exercises of this type improve the caliber of instruction within a health profession program, 91 percent agreed and nine percent were neutral. Attitudes surveyed confirmed the value of training programs to standardize clinical evaluation within a health profession program. The faculty responded to the training sessions with enthusiasm and professionalism. Verbal comments included expressions of a feeling of unity within the faculty with further commitments to excellence in clinical instruction. These types of improvements in faculty morale are hard to measure, but they are important to the process of continuous quality improvement. The emphasis on improving the caliber of clinical evaluation has ramifications for all aspects of instruction within the program. Training programs and increased levels of reliability are measurable efforts to improve the quality of instruction.

Another initial finding will be discussed in relation to research question one concerning intrarater reliability. This question asked if it was possible to improve intrarater reliability in calculus detection by using continuous quality training methods. The research found improvement in intrarater reliability both on typodonts and real patients for both the control groups and the experimental group. The improvement was statistically significant at  $p < .05$  level for intrarater reliability in the experimental group using typodont models of



dentition. The pre-training reliability level for the experimental group was .8958, and reliability increased after training to a statistically significant level of .9182. These findings reinforce comments made to the researcher at the National Dental Examiner Meeting of 1993 in Chicago, Illinois (personal communications, 1993).

Experienced examiners from Northeast Regional Dental Board, Central Regional Dental Board, and Washington State Dental Board stated that they had been able to increase reliability within their examiners using typodont training, but that reliability was not as high when actual patients were used in clinical settings. The use of typodonts should be examined for effectiveness as a method of increasing reliability.

The use of typodonts does seem to have a role in training examiners to be more reliable in subgingival calculus detection. Instrumentation is more difficult on typodonts, because the artificial gingiva (gums) are usually of hard vinyl and are not as easily manipulated as human tissue. Human patients generally have signs and symptoms indicating the presence of subgingival calculus that are obvious to the experienced examiner. Human gingiva usually has inflammation in areas where subgingival calculus exists, and will generally bleed upon exploration if irritants are present. None of these subtle "clues" are able to be adequately represented on typodont models, so the examiner must rely wholly on the tactile sensation of exploration. Simulated calculus does not feel exactly like calculus that forms on human teeth, and this discrepancy can confuse the examiner. Typodonts have the advantage of being readily available to

the examiner. Typodonts have simulated calculus placed on locations that are known to the person making the key, so there is an absolute answer. Human calculus detection is subject to the judgement of the examiner, because unless the tooth is extracted and examined, it is impossible to know absolutely if calculus is present or absent. This absoluteness of the existence of calculus on a typodont is what makes it useful for calibration training exercises. Examiners are able to come to a conclusive agreement about the presence of calculus because the teeth can be removed and examined. Typodonts are a good starting point for gaining examiner agreement. Once they have been mastered, the more difficult task of gaining reliability on human patients can be attempted.

As the research demonstrated, improving intrarater reliability on human patients to significant levels is more difficult. Reliability among the faculty of Parkland dental hygiene program was already at a high level before training. The experimental group improved their reliability from .8482 to .8819 although it was not statistically significant. There may be several reasons for this. Treatment of human patients contains variables such as difficulty of management, differences in tooth structures and soft tissue, and differing bodily response to periodontal disease. Patients have treatment problems such as bleeding and sensitivity making calculus detection more difficult. Since periodontal treatment is performed on "real" patients, it is important that reliability be at the highest level possible. Eisner (1993) stated that a goal for reliability should be levels of .80 or

higher. The Parkland examiners came into the training with a reliability level of .8482 so improvement to the level of .8819 was a tangible improvement. Varying the period of time between training sessions might have affected the results of the study. The small sample size and large standard deviation made achieving statistical significance very difficult.

In response to research question two, concerning reliability between examiners, interrater reliability also showed an improvement that was not significant. Pre-testing reliability of the experimental group was .8567 and it improved to .8631 compared to an improvement of the control group from .8199 to .8824. Neither of the groups' improvements were significant, although for reasons discussed above, the improvements were noteworthy. When reliability is already at a high level, even small improvements have a positive effect on the quality of the evaluation process. The improvement in the control group could be due to practice effect or experimenter bias. Pre-training levels for both the experimental group and control group were very high, which did not allow a large measure of improvement. As was noted earlier, the majority of the faculty had worked with each other for over three years, so this may have been a factor. The faculty may have already integrated many of the standardization techniques into the daily routine of instruction as they worked together for a prolonged period. The increased awareness of the importance of accuracy during the exercises may have boosted the initial reliability levels, or they could have been increased due to the "Hawthorne effect".

The final area of discussion concerning research was the type and amount of errors made in calculus detection. The majority of detection errors were errors of omission, or false negative errors. Examiners were missing areas of subgingival calculus much more frequently than they were finding areas of nonexistent calculus. This could have severe consequence in the success of treatment of periodontal disease. Subgingival calculus that remains is likely to contribute to the continuing loss of support of the periodontal structures and the continuation of the disease process. The other major implication is in clinical evaluation. If an instructor finds remaining calculus when checking a student's progress, it is extremely likely that the calculus exists, for errors of commission are rare.

### Conclusions

Based on the findings of this study, the following conclusions may be drawn:

1. It is possible to improve intrarater reliability for subgingival calculus detection using feedback, consensus and group standard setting techniques.
2. Improving intrarater reliability using models is less difficult than improving intrarater reliability with human patients, but both are possible and have important implications.

3. It is possible to improve interrater reliability for subgingival calculus detection using feedback, consensus, and group standard setting techniques. The reliability between examiners in clinical evaluation should be improved to the highest level possible to assure quality instruction and patient care.

4. Errors in subgingival calculus detection are usually errors of omission where existing calculus is missed by the examiner. Examiners should be aware of the likelihood of committing this type of error and the ramifications to the success of treatment.

5. It is possible to increase the perception of consistency within a group of clinical examiners through training exercise that stress continuous quality improvement such as feedback, consensus, and group standard setting techniques.

6. The level of confidence was improved between members of the group after training.

#### Recommendations and Related Issues

As this research progressed, several recommendations and related issues became apparent. The following recommendations are provided for: a) further research, and b) integration of calibration training in clinical instruction.

Recommendations for Further Research

1. Interrater and intrarater reliability testing should be further developed to find the maximum reliability level that can be achieved and maintained. Rater testing increases awareness of the importance of accuracy for both the faculty and the students. A follow-up study for the experimental and control groups could track reliability levels after cessation of training to see if the higher levels could be maintained.

2. Further testing in rater reliability should be done with students using the same techniques and principles to see if larger sample sizes could improve reliability to a significant level with inexperienced clinicians.

3. The same research project should be tested in other dental hygiene programs to compare results with other groups. It would be interesting to compare the amount of time the various groups had worked together to see how it affected reliability levels.

4. The principles of continuous quality improvement training could be used to standardize the evaluation process in other health professions programs such as nursing. The principles of continuous quality improvement are applicable in many training situations.

Recommendations for Integrating Calibration into Instruction

1. Calibration should be integrated into the routine of regular instruction. Emphasis on quality is obvious when faculty are striving to improve clinical skills in view of students. Faculty can set the example of striving for excellence at all stages of a career and influence the formation of professional ethics that are being shaped during clinical instruction.

Making calibration the exception creates added expense for faculty reimbursement and training sessions outside regularly scheduled clinical periods. Informal calibration exercises should become a part of the evaluation process by having two instructors come to a consensus on a student's final grade. More formal sessions should be a part of faculty orientation at the beginning of each semester. Because such a large proportion of the faculty is part-time, the first few hours of each day of the first week of clinicals should be used for calibration so that everyone could participate. Research strongly suggest that the effects of calibration on reliability are short-lived and that it must be an ongoing process to be effective.

2. Funding for training should be a priority. Accountability for excellence through continuous quality improvement should start with a commitment from administration to training for part-time as well as full-time faculty. Parkland currently has no funding for training for part-time faculty which is perceived as a lack of commitment from administration to these faculty.

3. Training techniques that are geared to the special learning needs of the professional should be used. Egos are very much an issue in this type of training, so every effort should be made to keep training non-threatening. Continuous quality improvement training is ideal for it creates a spirit of teamwork and allows the individual to take ownership of learning.

4. Use of the techniques such as setting group standards, consensus, and feedback should be incorporated into regular clinical activities. If these techniques help improve reliability for experienced instructors, it is likely they would be useful to the beginning clinician. Allowing the student to have greater control over the learning situation should increase cooperation and the quality of instruction.

5. Allow students to take part in improving their reliability by participating in the same types of exercises as the instructors. A process evaluation could be constructed to allow students to work together to improve calculus detection. Students should be evaluated on their ability to work with fellow clinicians as they improve interrater reliability in subgingival calculus detection. Other options would include having the students "grade" their patients at the completion of treatment to compare their results with other students or instructors.

6. Include use of typodont models for calculus detection exercises. Allow students to remove the teeth from the model so that they can reconcile tactile sensation with actual anatomy. Students



should do these exercises as part of clinical requirements or as an adjunct activity. Several models should be available so that a variety of situations are present.

References Cited

- American Dental Association Council on Dental Education. (July, 1992). Guidelines for valid and reliable dental licensure clinical examinations. ADA Curriculum Standard #5.14. Chicago, Illinois.
- Apps, J. (1981). The adult learner on campus. Chicago: Follett Publishing Co.
- Apps, J. (1979). Problems in continuing education. New York: McGraw-Hill Book Company.
- Abrahamowicz, M., Tamblyn, R.M., Ramsey, J.O., Klass, D.K., & Kopelow, M.L. (1990). Detecting and correcting for rater-induced differences in standardized patient tests of clinical competence. Academic Medicine, 65(9), S25-S26.
- Barrows, H.A., Williams, G.G., & Moy, R.H. (1987, October). A comprehensive performance-based assessment of fourth-year students' clinical skills. Journal of Medical Education, 62, 805-809.
- Bazan, M.T., & Seale, N.S. (1982). A technique for immediate evaluation of preclinical exercises. Journal of Dental Education, 46(12), 726-728.
- Bennett, N. & LeGrand, B. (1990). Developing continuing professional education programs. Urbana-Champaign: The Guide Series.
- Berwick, D.M. (1989). Continuous improvement as an ideal in health care. The New England Journal of Medicine, 320, 53-56.

- Boyle, P. (1981). Planning better programs. New York: McGraw-Hill.
- Caffarella, R.S., (1988). Program development and evaluation resource book for trainers. New York: John Wiley & Sons.
- Cervero, R.M. (1989). Continuing Education for the Professions. In S.B. Merriam & P.M. Cunningham (eds.) Handbook of adult and continuing education. (pp. 513-522). San Francisco: Jossey-Bass Publishers.
- Coates, V.E., & Chambers, M. (1992). Evaluation of tools to assess clinical competence. Nurse Education Today (UK), 12, 122-129.
- Eisner, J. (1993). Standardization and calibration in clinical evaluation. Paper presented to the annual meeting of the American Academy of Dental Examiners annual meeting, Chicago, Illinois.
- Emmons, M.R. (1979). The development of a procedure for the evaluation of dental hygiene clinical practice. Unpublished master's thesis, Illinois State University, Normal.
- Emmons, M.R. (1993). Interview with program director of Parkland College Dental Hygiene Program. Champaign, Illinois.
- Feil, P.H. (1982). An analysis of the reliability of a laboratory evaluation system. Journal of Dental Education, 46(8), 489-494.
- Gilley, J.W., & Eggland, S.A. (1992). Principles of human resource development. New York: Addison Wesley Publishing Company.

- Gronlund, N.E. (1976). Measurement and evaluation in teaching. New York: Macmillan Publishing Company, Inc.
- Houle, C.O. (1972). The design of education. San Francisco: Jossey-Bass Publishers.
- Johnson, S.D. (Ed.) (1991). Teaching techniques for part-time community college instructors. Urbana-Champaign: Department of Vocational and Technical Education, University of Illinois.
- Koch, M.W., & Fairly, T.M. (1993). Integrated quality management - The key to improving nursing care quality. St. Louis, Missouri: Mosby-Year Book.
- Knowles, M.S. (1980). The modern practice of adult education. (Rev. ed.) Chicago: Association Press/Follett.
- Kozoll, C. (1992). AHCE 448 class notes. University of Illinois, Urbana, Illinois.
- Laird, D. (1985). Approaches to training and development (Rev. 2nd ed.). Reading, Massachusetts: Addison-Wesley Publishing Company.
- Lauffer, A. (1978). Doing continuing education and staff development. New York: McGraw-Hill Book Company.
- Mackenzie, R.S., Antonson, D.E., Welsey, P.L., Welsch, B.B., & Simpson, W.J. (1982). Analysis of disagreement in the evaluation of clinical products. Journal of Dental Education, 46 (5), 284-288.

- Marsick, V.J., & Smedley, R.R. (1989). Health education. In S. B. Merriam & P.M. Cunningham (Eds.) Handbook of adult and continuing education. pp 502-510. San Francisco: Jossey-Bass Publishers.
- Martin, J. A., & Carey, R. M. (1991). The student-initiated, faculty-assisted system of evaluation of clinical teaching at the University of Virginia. Academic Medicine, 66(12), 773-775.
- Meetz, H. K., Bebeau, M. J., & Thoma, S.J. (1988). The validity and reliability of a clinical performance rating scale. Journal of Dental Education, 52(6), 290-297.
- Miller, W. R. (1990). Instructors and their jobs. Homewood, Illinois: American Technical Publishers, Inc.
- Morganstein, W. M. (1990). Outcome measures in dental education - we've only just begun. Journal of Dental Education, 54(6), 308-310.
- Nowlen, P.M. (1988). A new approach to continuing education for business and the professions - the performance model. New York: Macmillan Publishing Company.
- O'Hehir, T. (1993). Calculus removal should be based on body's reaction, not thoroughness. RDH, 13(3), pp 23.
- Partridge, M. I. & Mast, T.A. (1978). Dental clinical evaluation: a review of the research. Journal of Dental Education, 42(6), 300-305.
- Pavlish, C. (1987). A model for clinical performance evaluation. Journal of Nursing Education, 26(8), 338-339.

- Pippin, D. J., & Feil, P. (1992). Interrater agreement on subgingival calculus detection following scaling. Journal of Dental Education, 56(5), 322-326.
- Schoen, D. H. (1992). Instrument effects on smoothness discrimination. Journal of Dental Education, 56(11), 741-745.
- Sherman, P.R., Hutchens, L. H., Jewson, L. G., Moriarty, J. M., Greco, G. W., & McFall, W. T. (1990). The effectiveness of subgingival scaling and root planing: clinical detection of residual calculus. Journal of Periodontics, 61(1), 3-8.
- Sisco, B. (1991). Setting the climate for effective teaching and learning. In Hiemstra (Ed.) Creating Environments for Effective Adult Learning. San Francisco: Jossey-Bass Publishers.
- Stambaugh, R.V., Dragoo, M., Smith, D. M., & Carasali, L., (1981). The limits of subgingival scaling. International Journal of Periodontics and Restorative Dentistry, 1(5), 30-41.
- Stemmler, E.J. (1986). Promoting improved evaluation of students during clinical education: a complex management task. Journal of Medical Education, 61(9), 75-81.
- Tonesk, X. (1986). AAMC program to promote improved evaluation of students during clinical evaluation. Journal of Medical Education, 61(1), 83-88.
- Tonesk, X., and Buchanan, R. G., (1987). An AAMC pilot study by 10 medical schools of clinical evaluation of students. Journal of Medical Education, 62, 707-718.

Wilkins, E. (1989). The clinical practice of the dental hygienist.  
(Sixth ed.). Philadelphia: Lea & Febiger.

Appendix A

Pre-exercise survey

Circle the answer that best describes your feelings.

Strongly disagree          neutral          strongly agree

1    2    3    4    5    6    7    8    9    10

1. My detection of subgingival calculus is consistent with other instructors in this program.

1    2    3    4    5    6    7    8    9    10

2. Detection of subgingival calculus is usually consistent between most of the instructors in this program.

1    2    3    4    5    6    7    8    9    10

3. The grade I give is entirely dependent upon the calculus remaining or absent. I do not consider other factors such as patient management, skill level, or personality.

1    2    3    4    5    6    7    8    9    10

4. I feel confident that my clinical evaluation of students reflects an accurate assessment of their skill.

1    2    3    4    5    6    7    8    9    10

5. I feel confident that another instructor "checking" my calculus detection will get the same results.

1    2    3    4    5    6    7    8    9    10

6. I feel well prepared in clinical evaluation of student's skills.

1    2    3    4    5    6    7    8    9    10

7. I feel confident in my abilities as a clinical evaluator.

1    2    3    4    5    6    7    8    9    10



## Instructor Calibration

76

8. The faculty of the dental hygiene program consistently use the same set of criteria in clinical evaluation.

1 2 3 4 5 6 7 8 9 10

9. I would like more training in clinical evaluation.

1 2 3 4 5 6 7 8 9 10

10. I feel threatened by the process of "peer review".

1 2 3 4 5 6 7 8 9 10

Comments:

Appendix B

Post-exercise survey

Circle the answer that best describes your feelings.

Strongly disagree                  neutral                  strongly agree

1      2      3      4      5      6      7      8      9      10

1. My detection of subgingival calculus is consistent with other instructors in this program.

1    2    3    4    5    6    7    8    9    10

2. Detection of subgingival calculus is likely to be more consistent as a result of participating in this exercise.

1    2    3    4    5    6    7    8    9    10

3. Participating in this exercise will influence the way I grade a student.

1    2    3    4    5    6    7    8    9    10

4. I feel confident that my clinical evaluation of students reflects an accurate assessment of their skill.

1    2    3    4    5    6    7    8    9    10

5. I feel confident that another instructor "checking" my calculus detection will get the same results.

1    2    3    4    5    6    7    8    9    10

6. I feel better prepared to evaluate a student's clinical skills as a result of this exercise.

1    2    3    4    5    6    7    8    9    10

7. I feel confident in my abilities as a clinical evaluator.

1    2    3    4    5    6    7    8    9    10

## Instructor Calibration

78

8. The faculty of the dental hygiene program consistently uses the same set of criteria in clinical evaluation.  
1 2 3 4 5 6 7 8 9 10
9. The training and collaboration that I participated in during these exercises was useful to me.  
1 2 3 4 5 6 7 8 9 10
10. I would like to participate in similar types of training programs in the future.  
1 2 3 4 5 6 7 8 9 10
11. Participating in these exercises has given me a better understanding of other faculty member's evaluation criteria.  
1 2 3 4 5 6 7 8 9 10
12. Exercises such as this improve the caliber of instruction within a health profession program.  
1 2 3 4 5 6 7 8 9 10

Comments:

## Appendix C

Comparison of Attitudes Survey

Question One: (the same on both surveys) Number = 11

My detection of subgingival calculus is consistent with other instructors in this program.

Pre-training                      Neutral 82% N=9    Agree 18% N=2

Post-training                    Neutral 27% N=3    Agree 73% N=8

Question Two: (the same on both surveys) Number = 11

Detection of subgingival calculus is usually consistent between most of the instructors in this program.

Pre-training                      Neutral 73% N=8    Agree 27% N=3

Post-training                    Neutral 55% N=6    Agree 45% N=5

Question Three: (Pre-training survey) Number = 11

The grade I give is entirely dependent upon calculus remaining or present. I do not consider other factors such as patient management, skill level, or personality.

Disagree 45% N=5              Neutral 36% N=4              Agree 18% N=2

Question Three: (Post-training) Number = 11

Participating in this exercise will influence the way I grade a student.

Disagree 27% N=3              Neutral 55% N=6              Agree 18% N=2

Question Four: (the same on both surveys) Number = 11

I feel confident that my clinical evaluation of students reflects an accurate assessment of their skill.

Pre-training                      Neutral 18% N=2    Agree 82% N=9

Post-training                      Neutral 82% N=9    Agree 18% N=2

Question Five: (the same on both surveys) Number = 11

I feel confident that another instructor "checking" my calculus detection will get the same results.

Pre-training                      Neutral 36% N=4    Agree 64% N=7

Post-training                      Neutral 27% N=3    Agree 73% N=8

Question Six: (Pre-training survey) Number = 11

I feel well prepared in clinical evaluation of student's skills.

Disagree 18% N=2              Neutral 27% N=3    Agree 55% N=6

Question Six: (Post-training survey) Number = 11

I feel better prepared to evaluate a student's clinical skills as a result of this exercise.

Disagree 0% N=0              Neutral 55% N=6              Agree 45% N=5

Question Seven: (the same on both surveys) Number = 11

I feel confident in my abilities as a clinical evaluator.

Pre-training                      Neutral 0% N=0    Agree 100% N=11

Post-training                      Neutral 9% N=1    Agree 91% N=10

Question Eight: (the same on both surveys) Number = 11

The faculty of the dental hygiene program consistently uses the same set of criteria in clinical evaluation.

Pre-training                      Neutral 55% N=6      Agree 45% N=5

Post-training                      Neutral 45% N=5      Agree 55% N=6

Question Nine: (Pre-training survey) Number = 11

I would like more training in clinical evaluation.

Disagree 9% N=1                      Neutral 45% N=5                      Agree 45% N=5

Question Nine: (Post-training survey) Number = 11

The training and collaboration that I participated in during these exercises was useful to me.

Disagree 0% N=0                      Neutral 27% N=3                      Agree 73% N=8

Question 10: (Pre-training survey) Number = 11

I feel threatened by the process of "peer review".

Disagree 55% N=6                      Neutral 36% N=4                      Agree 9% N=1

Question 10: (Post-training survey) Number = 11

I would like to participate in similar types of program in the future.

Disagree 0% N=0                      Neutral 18% N=2                      Agree 82% N=9

Question Eleven: (Post-test survey) Number = 11

Participating in these exercises has given me a better understanding of other faculty member's evaluation criteria.

Disagree 0% N=0                      Neutral 9% N=1                      Agree 91% N=10

Question Twelve: (Post-training survey) Number = 11

Exercises such as these improve the caliber of instruction within a health profession.

Disagree 0% N=0

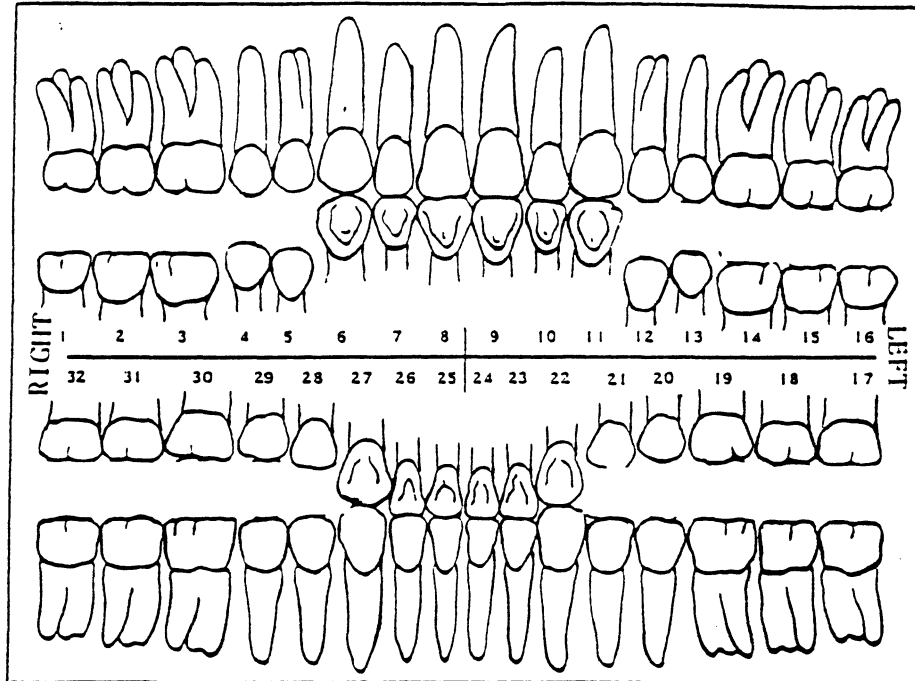
Neutral 9% N=1

Agree 91% N=10

NAME \_\_\_\_\_ Date \_\_\_\_\_ Inst. Init. \_\_\_\_\_ C/NC

CALCULUS DETECTION EVALUATION  
Calculus Charting

DIRECTIONS: Cross through all missing teeth in assigned area. With blue ink, draw lines on the chart which correspond to calculus detected on the patient's teeth. Be careful to note specific placement. Each tooth has a maximum of 12 areas. Instructor will note, by circling in red ink, areas where calculus was (1) not detected, and/or (2) incorrectly detected.



Total number of areas student correctly identified \_\_\_\_\_ = \* \_\_\_\_\_ %  
Total number of areas instructor identified \_\_\_\_\_

Instructor observation of exploring techniques

	Always (2)	Sometimes (1)	Never (0)
Reaches junctional epithelium			
Thoroughly covers interproximal surfaces			
Uses appropriate lateral pressure			
Stroke length & method ensure total tooth coverage			

Student description of calculus present: \_\_\_\_\_

\*Percentage must  $\geq 85\%$  to receive credit  
calcdet.evl





Parkland College  
Dental Hygiene Clinic  
Room L148  
351-2221

## Medical History

See number(s): \_\_\_\_\_

☐ MC

PATIENT'S NAME (last)		(first)	DATE OF BIRTH		HEIGHT	WEIGHT
HOME STREET ADDRESS		CITY	STATE	ZIP	HOME PHONE ( )	
SCHOOL BUSINESS STREET ADDRESS		CITY	STATE	ZIP	SCHOOL BUSINESS PHONE ( )	
OCCUPATION	SEX <input type="checkbox"/> M <input type="checkbox"/> F	RACE	PARENT OR GUARDIAN, IF UNDER 18 (last name first)			
PHYSICIAN'S NAME (last name first)		ADDRESS		ZIP	PHYSICIAN'S PHONE ( )	
LOCAL EMERGENCY FACILITY			PATIENT CLINIC NUMBER (if applicable)		EMERGENCY FACILITY PHONE ( )	
DENTIST'S NAME (last name first)		ADDRESS		ZIP	DENTIST'S PHONE ( )	

	DATE	DATE	DATE	DATE	DATE
1. Has there been any change in your general health within the past year? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
2. When was your last visit to a physician? .....					
What was the reason for the visit? .....					
3. Are you now under the care of a physician? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, what is the condition being treated? .....					
4. Have you had any serious illness or operation? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, describe illness/operation and when it occurred. ....					
5. Have you been hospitalized or had a serious illness within the past five years? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, describe the problem and when it occurred. ....					
6. Have you ever received a blood transfusion? .....	Yes/No	Yes/No	Yes/No	Yes/No	
If so, give date of transfusion .....					
7. Do you have or have you had any of the following conditions or diseases?					
a. High blood pressure .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
b. Stroke .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*c. History of heart murmur .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*d. Rheumatic fever or rheumatic heart disease .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*e. Congenital heart defect .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*f. Cardiovascular disease — heart attack or angina (circle) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(1) Do you have pain in your chest upon exertion? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(2) Are you short of breath after mild exercise? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(3) Do your ankles swell? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(4) Do you get short of breath when you lie down or do you require extra pillows when you sleep? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(5) Do you have a pacemaker? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
g. Low blood pressure .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(1) Have you ever fainted? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, how often? .....					

	DATE	DATE	DATE	DATE	DATE
h. Anemia or other blood disorder .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*i. Abnormal bleeding associated with previous dental extractions, surgery or trauma .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*j. Do you wear an implanted device? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, explain .....					
*k. Organ transplant .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, explain .....					
l. Allergy to plants, animals, dust (circle) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
m. Sinus trouble .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
n. Asthma .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, how frequent are the attacks? .....					
If you have asthma, do you use an inhaler? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
**o. Tuberculosis .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
p. Persistent cough .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
q. Seizures (epileptic or other) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, how often do they occur? .....					
r. Central nervous system disorder (Parkinson's disease or other) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
s. Diabetes .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, type _____ How is it controlled? .....					
(1) Has any blood relative had diabetes? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
t. Arthritis .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
u. Stomach ulcers .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
v. Sudden weight loss or gain .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
w. Hearing, speech, or sight impairment (circle) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
x. Do you wear contact lenses? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*y. Hepatitis, jaundice, or liver disease (circle) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, what type of hepatitis? .....					
(1) Have you had contact with anyone who has had hepatitis in the last 6 months? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, what type of hepatitis? _____ Nature of contact .....					
*z. Kidney disease .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, have you had kidney dialysis? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
*aa. Have you been diagnosed as having AIDS or tested HIV positive? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
(1) Have you had contact with anyone who has AIDS? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, what is the nature of your contact — social, intimate, or are you a healthcare provider? (circle) .....					
bb. Venereal disease — gonorrhea, syphilis, herpes? (circle) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
**cc. Have you had contact with anyone who has had any infectious disease? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, what is the nature of the disease? .....					
**dd. Have you ever had cancer? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
When? .....					
Where was it located? .....					
What treatment was used? .....					
Are you still being treated for this condition? .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
8. Are you taking any nonprescription medications such as cold or sinus medication, aspirin, weight control medication, or other? (circle) .....	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
If so, explain .....					

[illegible]

CHECK THE BOX AND SIGN IF APPLICABLE.

☐ I have answered these questions to the best of my knowledge

for my \_\_\_\_\_  
(relationship to individual) who is non-English speaking.

RESPONDEE'S SIGNATURE \_\_\_\_\_

☐ I give my permission for my child, who is under 18 years of age,  
to be treated in the Parkland College Dental Hygiene Clinic for  
as many appointments as it is necessary to complete treatment.

PARENT OR GUARDIAN'S SIGNATURE \_\_\_\_\_

I SIGNIFY THAT THE HISTORY OF MY HEALTH CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_

I SIGNIFY THAT I HAVE REVIEWED THE PATIENT'S HEALTH HISTORY.

Date \_\_\_\_\_ Clinical Dentist \_\_\_\_\_

I SIGNIFY THAT THE UPDATE OF MY HEALTH HISTORY CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_

I SIGNIFY THAT THE UPDATE OF MY HEALTH HISTORY CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_

I SIGNIFY THAT THE UPDATE OF MY HEALTH HISTORY CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_

I SIGNIFY THAT THE UPDATE OF MY HEALTH HISTORY CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_

I SIGNIFY THAT THE UPDATE OF MY HEALTH HISTORY CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_

I SIGNIFY THAT THE UPDATE OF MY HEALTH HISTORY CONTAINED ON THIS FORM IS TRUE AND ACCURATE TO THE BEST OF MY KNOWLEDGE.

Date \_\_\_\_\_ Patient's Signature \_\_\_\_\_

BP \_\_\_\_\_ Student's Signature \_\_\_\_\_

Instructor's Signature \_\_\_\_\_



Parkland College  
2400 West Bradley Avenue  
Champaign, Illinois 61821-1899  
217/351-2200

## Dental Hygiene Clinic Agreement

Parkland College is a teaching institution that offers dental hygiene services to residents of the area. The patients who seek the services of the dental hygiene clinic are accepted on their merits as suitable educational experiences for the students. Based upon an initial appointment, you may or may not qualify as an acceptable patient. If your dental requirements are too extensive or too complicated, your treatment in this clinic will be terminated at the discretion of the clinical dentist and the dental hygiene staff. You will then be advised to seek dental care from your personal dentist or other agency.

The following information provides a basis for an agreement between you and the college if you are accepted as a patient.

1. It is important for you to assist your student dental hygienist by providing a complete history of past and present dental and medical conditions. Certain medical conditions may necessitate medical consultation with your physician prior to receiving dental hygiene services in our clinic. Patients who have active communicable infectious diseases will not be treated in the Dental Hygiene Clinic.
2. In the event your student dental hygienist has an accidental occupational exposure, you will be requested to have blood drawn for necessary lab tests. The cost of the lab tests will be paid for by Parkland College. By signing this document, you are agreeing to the release of your dental and medical information to the Public Health Department and/or Covenant Medical Center should an accidental exposure occur.
3. For the protection of your eyes, you will be asked to wear either the safety glasses provided by the clinic or your own glasses during your treatment.
4. Our goal and responsibility is to provide you with considerate, respectful, and confidential treatment. An initial exam will be performed, which will determine the treatment plan necessary for your existing condition. You will be informed of the recommended treatment and treatment alternatives. You will have the option of refusing treatment. You will also be advised of the risk of no treatment. In addition to cleaning your teeth, the following procedures may be done when indicated: complete mouth series x-rays and/or bitewing x-rays, impressions for study model, pit and fissure sealant placement, polishing of restorations, nutrition and plaque control counseling. Due to student clinic schedules, we may not be able to provide all of the additional procedures that are recommended. After your cleaning, you will be referred to your personal dentist. Dental hygiene is only a part of your total dental care. We will work with your personal dentist to assure you the best and most timely care. We ask that you identify a personal dentist prior to any treatment.
5. The clinic fee of \$10 per semester is to be paid at the first appointment unless you are otherwise exempted from this fee. Exemptions include children under twelve years of age; senior citizen (sixty-two and over); and those on disability, social security, or I.P.A.C.
6. It is important for you to be prompt and keep all appointments. We request twenty-four hours notice in advance if you are unable to keep your appointment. If you fail to keep your appointment or are tardy, students lose valuable experience that may adversely affect their progress in the program.
7. Several appointments, one hour and thirty minutes in length, may be necessary to complete your treatment. We feel confident that our service is of the highest quality and well worth your time. If, after a discussion with your student dental hygienist, you feel that you cannot spend the required amount of time, please let the student know immediately.
8. Your patient records become the property of the Dental Hygiene Clinic. However, your x-ray films will be sent to your private dentist within two weeks after treatment is completed in the clinic.
9. If you have any questions concerning the condition of your mouth or teeth, we encourage you to ask the clinical dentist or dental hygiene staff.
10. From time to time, photographs and television cameras are used in the clinic for educational purposes. Unless you inform us to the contrary, we will consider your signing of this agreement as authorization for you to be included in group photographs and television filming when these occur in the clinic. We will ask you to sign a release and consent form for any individual photography or television filming.

*I have read the preceding information and consent to the dental procedures performed by dental hygiene students as part of my supervised dental treatment. I agree to abide by the rules and regulations as herein stated above.*

Date

Patient's Name and Signature

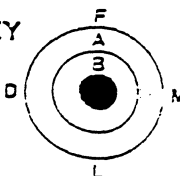
Parent or Guardian (if patient is under 18 years of age)

Parkland College does not discriminate on the basis of age, color, race, national origin, sex, religion, or disability.

# STANDARDIZATION EXERCISE

## Detection of Subgingival Calculus

KEY

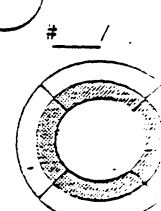
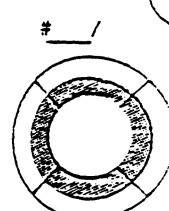
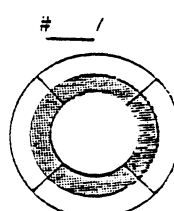
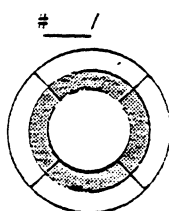
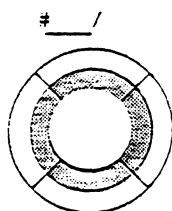
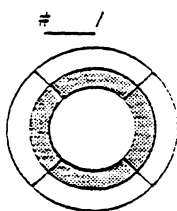
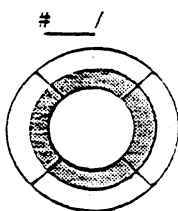


Examiner A \_\_\_\_\_ Exam Site \_\_\_\_\_

Examiner B \_\_\_\_\_ Date \_\_\_\_\_

Trial # 1 2 3 Tooth Sample \_\_\_\_\_

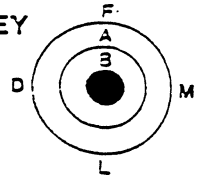
INSTRUCTIONS: Record tooth sample numbers above wheel grids. Record calculus by placing a heavy dot (•) in the appropriate space of the wheel grid.



# STANDARDIZATION EXERCISE

## Detection of Subgingival Calculus

KEY

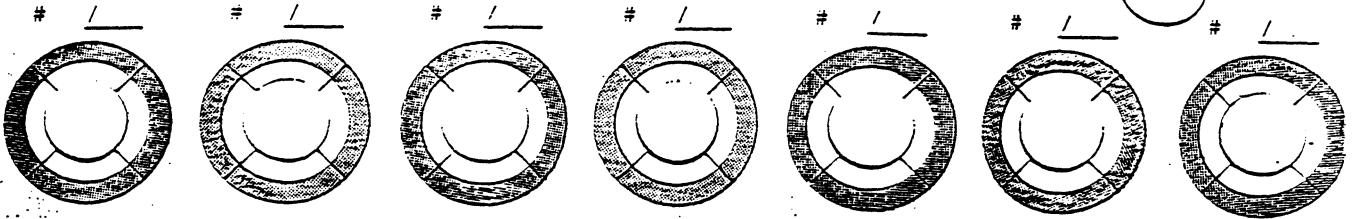


Examiner A \_\_\_\_\_ Exam Site \_\_\_\_\_

Examiner B \_\_\_\_\_ Date \_\_\_\_\_

Trial # 1 2 3 Tooth Sample \_\_\_\_\_

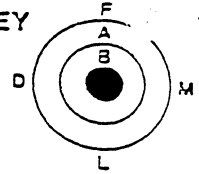
INSTRUCTIONS: Record tooth sample numbers above wheel grids. Record calculus by placing a heavy dot (•) in the appropriate space of the wheel grid.



# STANDARDIZATION EXERCISE

## Detection of Subgingival Calculus

KEY

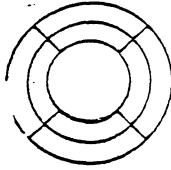
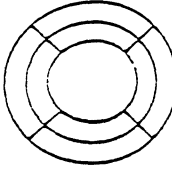
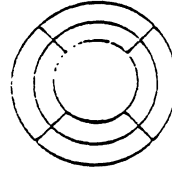
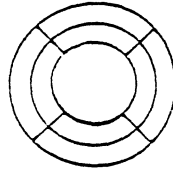
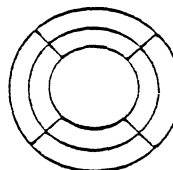
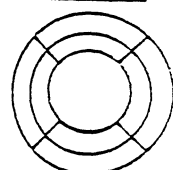
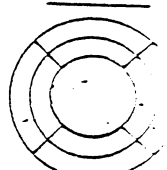


Examiner A \_\_\_\_\_ Exam Site \_\_\_\_\_

Examiner B \_\_\_\_\_ Date \_\_\_\_\_

Trial # 1 2 3 Tooth Sample \_\_\_\_\_

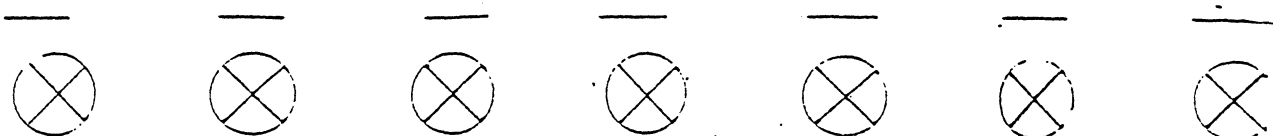
VALIDATION: Confirm teammate's findings by placing an "X" in your space of the wheel grid. Deny teammate's findings by writing "NO" in your space of the grid.

# /	# /	# /	# /	# /	# /	# /
						

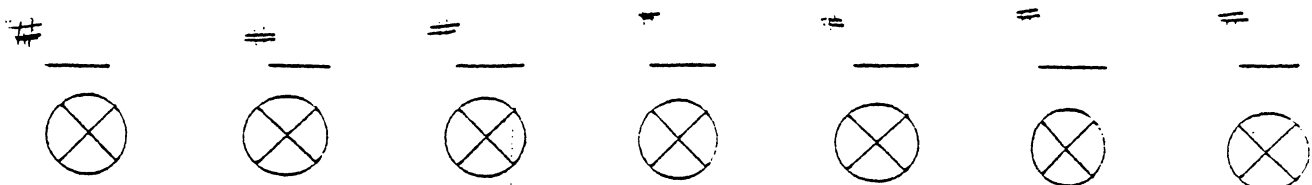


UPPER RIGHT QUADRANT

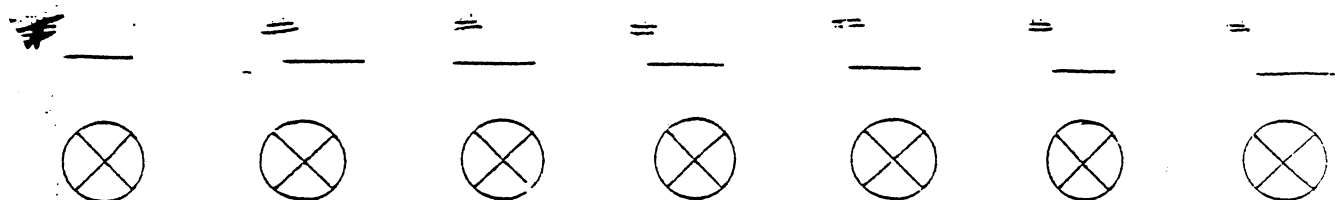
92



UPPER LEFT QUADRANT

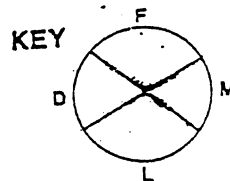
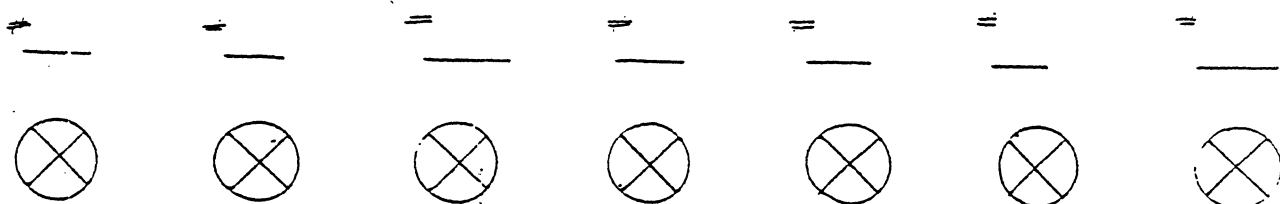


Lower Left Quadrant



Key

Lower right Quadrant



date \_\_\_\_\_

Patient # \_\_\_\_\_