INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*

4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

*For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.

UMI Dissertation Information Service
University Microfilms International
A Bell & Howell Information Company
300 N. Zeeb Road, Ann Arbor, Michigan  48106
Ramirez-Fernandez, Luis C.

THE EVALUATION OF CHILEAN MEDICAL EDUCATORS' PERCEPTIONS ABOUT ESTABLISHING A NATIONAL MEDICAL EXAMINATION IN CHILE

The Ohio State University

PH.D. 1986

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
THE EVALUATION OF CHILEAN MEDICAL EDUCATORS'
PERCEPTIONS ABOUT ESTABLISHING A NATIONAL
MEDICAL EXAMINATION IN CHILE

Dissertation

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

By
Luis Ramirez-Fernandez, Lic. Ed., M.A.

* * * * *

The Ohio State University
1986

Dissertation Committee:
Gregory Trzebiatowski, Ph.D.
Robert Bargar, Ph.D.
C. Benjamin Meleca, Ph.D.

Approved by:

Gregory Trzebiatowski
Adviser
Department of Educational
Policy & Leadership
Copyright by
Luis Ramirez-Fernandez
1986
ACKNOWLEDGMENTS

Thanks to:

Gregory Trzebiatowski, for being the architect of my growth and development in the field of medical education. You gave unreservedly as a teacher and friend;

Robert Bargar, for your guidance and support. You helped me to discover the field of naturalistic inquiry;

Ben Meleca, for sharing your experience in research. You gave unrelentless encouragement to finish;

Elso Schiappacasse, for the inspiration and permanent guidance in the medical perspective needed for the study;

Larry Sachs, for your generosity in sharing your knowledge and expertise;

Dorothy Brickman, for your care, support, which gave me the courage to finish;

Kevin Harty, for your friendship, care and permanent concern for my family's well being;

Juan Manso, for your guidance, expertise and enthusiasm which gave strength to the development of the study;

Sallie Mitchel, for your friendship, good will and competence which made my work more expeditious and enjoyable;

Patricia and Daniela, my wife and daughter, for your love, care and support needed for the accomplishment of my goals.
VITA

April 3, 1946

Born, Angol, Chile

1971

Lic. Ed. (English Teaching major), College of Education, University of Concepcion, Concepcion, Chile.

1977

M.A. (Educational Communications major), College of Education, The Ohio State University, Columbus, Ohio.

1973

Director of the Learning Resources Center, College of Medicine, University of Concepcion, Concepcion, Chile.

1973-1984

Professor of the Medical Education Program, College of Medicine, University of Concepcion, Concepcion, Chile.

1984

Professor of Instructional Design, Graduate Program, College of Education, University of Concepcion, Concepcion, Chile.

1985

Associate Director of the Medical Education Program, College of Medicine, University of Concepcion, Concepcion, Chile.

1985

Director of the Continuing Education Program of the Chilean Medical Association, Concepcion, Chile.

PUBLICATIONS

VITA (CONTINUED)


Ramirez, L. 1981. "Development of an Independent Study Program in a School of Medicine." Proceedings of a Conference on University Pedagogy, Catholic University, Santiago, Chile.

Ramirez, L. 1974. "Educational Technology in the Medical Curriculum." Seminar on Determining Medical School Goals, College of Medicine, University of Concepcion, Concepcion, Chile.


Schiappacasse, E., Ramirez, L. 1985. "Diagnosis and improvement of clinical teaching in internal medicine clerkship using television." Medical Society of Concepcion. (Rodolfo Armas-Cruz Award by the Chilean Medical Association.


Schiappacasse, E., Ramirez, L. 1983. "Evaluation of professional competence of medical graduates from the College of Medicine, University of Concepcion, 1976-1980." College of Medicine, University of Concepcion, Concepcion, Chile.


VITA (CONTINUED)


FIELDS OF STUDY

Major Field: Educational Development. Field of Specialization: Medical Education
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>ii</td>
</tr>
<tr>
<td>Vita</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>I. Overview of the Study</td>
<td>1</td>
</tr>
<tr>
<td>Identification of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Protecting the Public Interest: A Conceptual Framework</td>
<td>13</td>
</tr>
<tr>
<td>Medical Licensure</td>
<td>17</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>23</td>
</tr>
<tr>
<td>Research Questions</td>
<td>24</td>
</tr>
<tr>
<td>Overview of Methods</td>
<td>25</td>
</tr>
<tr>
<td>Instrument Development</td>
<td>26</td>
</tr>
<tr>
<td>Scope of the Study</td>
<td>26</td>
</tr>
<tr>
<td>Generalizability of the Study</td>
<td>26</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>26</td>
</tr>
<tr>
<td>Impact of the Study</td>
<td>27</td>
</tr>
<tr>
<td>Definitions of Important Terms</td>
<td>29</td>
</tr>
<tr>
<td>Endnotes</td>
<td>30</td>
</tr>
<tr>
<td>II. Review of Related Literature</td>
<td>33</td>
</tr>
<tr>
<td>Historical Perspective in USA</td>
<td>39</td>
</tr>
<tr>
<td>Testing Development</td>
<td>43</td>
</tr>
<tr>
<td>Continuum of Evaluation in the Medical Education Process</td>
<td>44</td>
</tr>
<tr>
<td>Determining Competences in Medical Education</td>
<td>48</td>
</tr>
<tr>
<td>Methods of Assessing Competence in Medical Education</td>
<td>56</td>
</tr>
<tr>
<td>External Evaluation in Medical Education</td>
<td>67</td>
</tr>
<tr>
<td>Testing Programs</td>
<td>72</td>
</tr>
<tr>
<td>Conclusions of Review of Related Literature</td>
<td>75</td>
</tr>
<tr>
<td>Endnotes</td>
<td>77</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (CONTINUED)

III. Methodology ......................................................................................... 82
   Research Design and Procedures .......................................................... 82
   Quantitative Approach ............................................................................ 83
      Dependent Variable ............................................................................. 83
      Independent Variable .......................................................................... 84
   Instrument Development and Data Collection ....................................... 84
   Population and Sample ........................................................................... 87
   Data Analysis .......................................................................................... 90
   Limitations of the Study ......................................................................... 91
   Qualitative Approach .............................................................................. 91
      Purpose of the Qualitative Approach .................................................. 91
      Methodology ...................................................................................... 93
      The Setting ......................................................................................... 95
      Data Analysis ...................................................................................... 96
   Endnotes .................................................................................................. 98

IV. Results ................................................................................................... 99
   Quantitative Approach ........................................................................... 99
      Examination of the Return Rate .......................................................... 99
      Descriptive Response Data and Tables ............................................... 104
   Qualitative Approach ............................................................................ 123
      Data Analysis for the Methodological Issues ...................................... 127
      Content Analysis of Open-ended Comments ...................................... 131

V. Summary, Conclusions and Recommendations .................................... 133
   Introduction ............................................................................................ 133
   Quantitative Approach .......................................................................... 134
      Discussion of Research Question 1 ....................................................... 136
      Discussion of Research Question 2 ....................................................... 138
      Discussion of Research Question 3 ....................................................... 140
      Discussion of Research Question 4: Methods ..................................... 142
      Discussion of Item 14 ......................................................................... 143
      Summary of Findings .......................................................................... 144
   Qualitative Approach ............................................................................ 145
      Conclusion Drawing and Verification .................................................. 145
      Summary of Conclusions ..................................................................... 148
      Recommendations ............................................................................... 149
      Implications for Future Studies ............................................................ 150
   Endnotes .................................................................................................. 152
| TABLE OF CONTENTS (CONTINUED) |
|-----------------------------|---|
| Bibliography                | 153 |
| Appendix A                  | 165 |
| Appendix B                  | 176 |
| Appendix C                  | 179 |
| Appendix D                  | 183 |
| Appendix E                  | 185 |
| Appendix F                  | 187 |
| Appendix G                  | 189 |
| Appendix H                  | 191 |
| Appendix I                  | 193 |

viii
# LIST OF TABLES

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sampling plan</td>
<td>89</td>
</tr>
<tr>
<td>2</td>
<td>Sampling plan for key people</td>
<td>90</td>
</tr>
<tr>
<td>3</td>
<td>Return rate information</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>Factor analysis summary table results</td>
<td>102</td>
</tr>
<tr>
<td>5</td>
<td>Percentages of agreement and disagreement by university respect to the need for a NME</td>
<td>105</td>
</tr>
<tr>
<td>6</td>
<td>Means and standard deviations by university of items that constitute the need factor</td>
<td>106</td>
</tr>
<tr>
<td>7</td>
<td>Means of key people corresponding to the items that constitute the need factor</td>
<td>107</td>
</tr>
<tr>
<td>8</td>
<td>Percentages of agreement and disagreement by university respect to the entity for the management of a NME</td>
<td>108</td>
</tr>
<tr>
<td>9</td>
<td>Means and standard deviations by university of items that constitute the entity factor</td>
<td>109</td>
</tr>
<tr>
<td>10</td>
<td>Means of key people corresponding to the items that constitute the entity factor</td>
<td>110</td>
</tr>
<tr>
<td>11</td>
<td>Percentages of agreement and disagreement by university respect to the items constitute areas factor</td>
<td>111</td>
</tr>
<tr>
<td>12</td>
<td>Means and standard deviations per item by university that constitute areas factor</td>
<td>111</td>
</tr>
<tr>
<td>13</td>
<td>Means standard deviations of key people corresponding to the items that constitute the areas factor</td>
<td>112</td>
</tr>
<tr>
<td>14</td>
<td>Percentages of agreement and disagreement by university respect to the adequacy of current evaluation methods</td>
<td>113</td>
</tr>
<tr>
<td>15</td>
<td>Means and standard deviations of items by university that constitute method factor</td>
<td>113</td>
</tr>
</tbody>
</table>
LIST OF TABLES (CONTINUED)

16. Means of key people corresponding to items that constitute the method factor. ............................... 114
17. Percentages of agreement and disagreement by university respect to the inclusion of practical examination in a NME. .......................... 115
18. Means and standard deviations by university of item 3 that constitute practical examination factor. .................. 115
19. Means of key people corresponding to items 3 that constitute the method factor .................................... 116
20. Percentages of agreement and disagreement respect to the university becoming the organization responsible of the design and administration of a NME ................................ 117
21. Means and standard deviations by university of item 4 that constitute the school-management factor. .................. 117
22. Means of key people corresponding to item 4 that constitute the school-management factor .......................... 118
23. Percentages of agreement and disagreement by university respect to the usefulness of objective tests .................. 119
24. Means and standard deviations by university of item 14 that constitute objective-test factor .......................... 119
25. Means of key people corresponding to item 14 that constitute objective-test factor .................................. 120
26. Summary for the effect of university for each factor/item. ................................................................. 122
27. Factors identified as possible inhibitors of the development of a NME in Chile ........................................ 124
28. Summary table of opinions about licensing. ................. 125
29. Opinions about certification of specialties. .................. 126
30. Opinions about pre-grade accreditation ........................ 127
LIST OF TABLES (CONTINUED)

31. Opinions about students' participation in decision-making process about establishing in Chile a NME. ......... 128

32. Opinions about physicians reluctancy to respond questionnaire.............................................. 129

33. Open-ended comments made by respondents of the questionnaire............................................. 131

34. Opinions about the refusal attitude of the Catholic University to participate in the study ......... 132

35. Summary of university means for each factor. .......... 135

36. Analysis of variance summary table for the need factor total. .................................................. 184

37. Analysis of variance summary table for the entity factor total. .................................................. 186

38. Analysis of variance summary table for the areas factor total. .................................................. 188

39. Analysis of variance summary table for the methods factor total. ............................................. 190

40. Analysis of variance summary table for the practice factor total. ............................................ 192

41. Analysis of variance summary table for the school-management factor total . ............................ 192

42. Analysis of variance summary table for the objective-test factor total ..................................... 194
# LIST OF FIGURES

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Public interest protection theoretical model</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Theoretical medical licensing model taxonomy</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>Continuum of evaluation in the medical education process</td>
<td>47</td>
</tr>
</tbody>
</table>
CHAPTER ONE

OVERVIEW OF THE STUDY

This study was designed to determine Chilean medical educators' support for establishing a National Medical Examination (NME) as an external evaluation mechanism to assure medical graduate's competence in delivering health care. Another significant part of the issues studied concerns who develops and administers a National Medical Examination which would then authorize graduates to practice medicine or issue licenses to practice. Currently this function is carried out by the various medical schools of Chile in an uncoordinated fashion.

The study was developed in response to criticisms about the quality of Chilean medical education and its evaluation procedures specifically made at national and international seminars and conferences on medical education held in Chile during the last twenty years. At the same time this study explored medical educators' attitudes toward the type of entity that should manage the examination, the areas to be assessed by an examination if there is one; the attitudes toward the evaluation methods being currently employed in medical education. The study was also designed to determine the possible factors that may have inhibited the development of a NME in Chile which was originally proposed in 1977.

The methodology employed in this research was a combination of a descriptive quantitative study using a survey approach and a qualitative study based on a naturalistic or anthropological approach.
The expected outcomes were the following: (1) Chilean medical educators from universities in the provinces would be more inclined toward the establishment of a NME than medical educators from the universities in Santiago; (2) Chilean medical educators in general would be in favor of developing a NME in Chile as an evaluation mechanism to assure medical graduates' competencies in practicing medicine; (3) Chilean medical educators would prefer the universities being the entity in charge of the design and management of a NME and would be against the involvement of the government in any evaluation process of medical school graduates.

IDENTIFICATION OF THE PROBLEM

The issue of physicians' competence and its evaluation methods have been present throughout the history of medical education. In a study sponsored by WHO, Charvat and McGuire (1968) reported: "those responsible for the health services of a country are concerned above all with quantity and quality of young physicians who graduate from medical schools." The quantity and quality aspects was brought to world's attention during the Thirteen World Health Assembly in 1977 when it was decided (Resolution WHA 30.43) that the main social target of the member Governments and the World Health Organization (WHO) in the decades ahead should be "the attainment by all citizens of the world by the year 2000 a level of health that will permit them to lead a socially and economically productive life." Later, the declaration adopted by the International Conference on Primary Health Care (1981) convened by the WHO, UNICEF in Alma Ata, USSR, clearly identifies primary health care as the future priority in the development of health care systems of all the
countries as the key to the attainment of "Health for all by the year 2000." Chile subscribed to this declaration and therefore, universities, and governmental health agencies have been delineating the strategies that best fit Alma Ata's goal. The plan of action of the Chilean government for the attainment of such a goal has posed an interesting challenge to the Chilean medical education process, especially to the evaluation procedures employed in the schools of medicine since it is necessary to assess constantly if the graduates of the medical schools are competent in fulfilling the government expectations in quantity and what is most important in quality. This last aspect is closely related to the issue of assuring the public's health through sound evaluation methods, which are still under study and development. Since no single assessment method, according to Neufeld (1985), can adequately measure physicians' competence.

In Latin America, according to Andrade (1978) there has been an increasing concern for modifying the traditional systems of producing physicians in order to provide a better quality of health care. He further states that the governments, universities, professional associations, and several international institutions have dealt with the issue of assuring quality of health care by the convening of large number of meetings, conferences with the purpose of sensitizing medical educators to the importance of protecting the public's interest. Similarly many publications currently in circulation indicate the importance given to the issue. For example, Pan American Health Organization publishes Educacion Medica Y Salud, the AAMC publishes the Journal of Medical Education, and most medical journals publish articles
related to medical education issues.

Chile through its socio-economic program for 1981 through 1989, indicates that primary objective in the health care sector's "a recognition of the right of health for all citizens of the country. It is the government's responsibility to maximize the physical and mental well being of all Chileans during all their life from the moment of conception until death" (1981). This objective gives the universities an additional task since their educational output should meet to a reasonable extent the government's priorities. Two questions may be raised in relation to this point: (1) Who is going to certify that graduates have acquired the necessary competencies to practice medicine and the kind needed by the government NHS? and (2) What kind of mechanisms are going to be used in assessing medical graduates' competencies?

In Chile the right to health care is stated in the 1980 constitution of the Nation and was made public during the XXI Pan American Conference 1978-1982 celebrated in 1982: "The government protects the free and equal access to health care."

Medical schools, and health services worldwide have been increasingly concerned about evaluation mechanisms that may assure competence of those that will deliver health care. A good example of this was the Conference on Evaluation in Medical Education held in 1971 in the University of Alberta, Canada where the importance of evaluation was emphasized by the following statement: "Proper evaluation in medical education is as essential from the standpoint of the individual medical student, intern, or residents as it is for the medical school, teaching
hospitals, and community in general." It was also noted at the same conference that currently used methods of evaluation did not satisfy the needs of the individual being evaluated, nor the needs of the schools and the community.

More recent evidence of the actuality of evaluation as a mechanism to assure competence has been provided by The Report of the Panel on the General Professional Education of the Physician and College Preparation for Medicine: "Physicians for the Twenty-First Century" (AAMC, 1984) which states:

The lack of a sound system of evaluating students limits the efficacy of the educational experience in clinical clerkship. Faculties often do not recognize their dual roles as evaluators; in the development of competency, which requires periodical evaluation with feedback to students; and in the assessment of competency which requires the application of specific standards of evaluation for acceptable performance.

Chilean medical schools have also been increasingly concerned about the deficiencies of their evaluation methods. This concern is important because it rises from the fact that the currently employed evaluation methods constitute the only means to assure graduates' competence in delivering health care to patients. This pressing problem and the possibility of establishing a National Medical Examination (NME) in Chile have been the subjects of discussion during several national and international seminars and conferences on medical education. However, there is no evidence of a systematic effort to deal with the issue of assuring the public's health through improved evaluation systems or through a National Medical Examination.

The most critical issue at present in Chilean medical education is the absence of evaluation policies and mechanisms to measure students'
progress and to assure competence in delivering health care to patients.

Repeto and Zalazar (1984) studied the state of the art of evaluation systems in medicine in Chile. They concluded that there was no philosophy or theory guiding evaluation and that there were no criteria set for evaluator's behavior. They also reported that while evaluation norms exist for grading the cognitive domain, that no norms exist for the affective and psychomotor domains. Finally, they reported that most medical educators in Chile understand evaluation mainly as measurement and that they have not clearly defined the type of professional they are helping to produce.

Their findings seem to agree very well with the data provided by medical educators during the National Seminar on Medical Education held in Chile in June 25-27, 1984, which emphasized, that despite the advances achieved in the field of evaluation -- patient management problems, objective tests, use of computers, patient simulations, use of television, observation techniques and a host of other techniques -- the process of evaluating the education of Chilean medical students can still be characterized by its lack of planning, lack of appropriate instruments, absence of objectivity, validity and reliability, especially at a clinical level, and a lack of understanding of evaluation techniques. Also expressed at the 1984 National Conference that current evaluation methods stress cognitive knowledge ignoring psychomotor skills and problem solving, which are also very important in providing efficient health care. The conferences also noted that this last fact should concern medical educators since the National Health Service, the main employer of physicians, has been hiring graduates
basically according to grades obtained during their medical studies. The National Health Service in 1984 assigned a weight of 85% to grades and 15% to other educational activities such as assistantships, policlinical services, etc. (1984). A similar situation occurs with applicants to residencies, wherein grades are also considered more important (80% weight) than skills and problem solving abilities.

Schiappacasse & Pomeroy (1985) reported that in Chile due to the new governmental health policies, new health care systems, new social security legislation, new laws governing universities, new professional association regulations: "evaluation systems must be revised and new evaluation methods be explored." They added that new evaluation methods should be explored because schools of medicine will have to revise the curricular objectives for undergraduate and graduate medical education. They said: "Chilean medical education lacks specific educational objectives which makes the selection of content, design of instructional strategies and evaluation of students' competencies very difficult. Medical educators should be informed and trained in evaluation methods."

Another source which supplies valid data relative to the medical evaluation process are the Offices of Medical Education in various Schools of Medicine of the country. These units are in charge of medical pedagogy whose functions include faculty development in medical pedagogy, consultation to the dean in matters of medical education, research on medical education, etc. The Offices of Medical Education are organized into a national association and they periodically meet to analyze the medical education process of Chilean Schools of Medicine.

During many national and international conferences on medical
education well-known spokespersons from the Offices of Medical Education have made public their concern about the inadequacy of evaluation process in medicine. They have frequently stated that the problem has been aggravated by the following factors: first, evaluation has been practiced in medical schools with little educational value. This has become obvious after observing medical educators' lack of interest in acquiring or improving evaluation techniques and procedures. Second, medical educators do not have a clear profile of the type of professional they want to produce. Third, there is a lack of understanding of the evaluation process as an essential part of the medical education process which is designed to provide feedback needed to improve teaching and learning. Fourth, there is lack of tested mechanisms, procedures and standards to assure competency in delivering health care to patients.

Cristoffanini (1981) in his article on evaluation stated that at present the majority of medical schools worldwide consider evaluation one of the most important processes in the continuum of medical education. He identified the following three closely interconnected main aspects of evaluation: process, program and product. He described process evaluation as an effort that deals with the progress and attainment of instructional goals by students. Program evaluation tends to measure the quality of the educational program itself and its different elements. Finally, he described product evaluation as actions that measure the competency of physicians already graduated in delivering health care to patients.

Chilean medical schools have dealt mainly with process evaluation,
very little with program evaluation, and almost nothing with product evaluation. The lack of related literature is a testimony of the veracity of the above statements. Wilford Aiken (1972), referring to American education, stated: "It is an amazing fact that our schools and colleges know little of the results of their work." His words seem to be very appropriate to illustrate the present situation of Chilean medical schools regarding product evaluation.

It seems likely that there is no other field in which individuals continue to work with so little knowledge of their impact. The present status of evaluation in Chilean medical schools can be further described through the words of George Miller (1961). He implies that it is not lack of interest in the outcome that has blocked the development of meaningful evaluation, nor the lack of technical improvement in devices, tools and responsibilities. The fact that we know little of the results of our work "must then be attributed chiefly to a general ignorance among teachers of the complexities, the importance and the nature of evaluation (Miller, 1961).

There have been other factors inhibiting the development of product evaluation efforts. They are of a varied type: financial, a dearth of skilled specialists, lack of appropriate technology and knowledge.

The intent to deal formally with the issue of assurance of competencies of graduated physicians was high on the agenda of the Third National Seminar on Medical Education held in Chile in 1977 and was summarized in the following recommendation:

The evaluation in medical schools should be homogenized through a National Examination administered to all students from the various colleges and schools of medicine of the
country. This examination should be designed with the participation of representatives of all universities that have medical schools. It is believed that this examination would be a means for evaluating the teaching programs of colleges and schools of medicine (1977).

These recommendations are important because they are the first to be made by university based medical educators.

At a governmental level, President Ibanez in 1955 passed Law N11.861 which established requirements for obtaining the Doctor of Medicine degree (1955). The evaluation component of this legislative action stated that the University of Chile, a government supported institution, was in charge of determining the type of examination that medical school graduates were required to pass in order to obtain the M.D. degree and a license to practice medicine.

However, this law did not suggest the type of examination nor the characteristics of the examination.

The University of Chile School of Medicine immediately established the Final Examination or Graduating Examination, which had to be passed by the interns of medicine in order to graduate and obtain a license to practice medicine.

The Graduating Examination consisted of two parts. Part one called Pre-Grado, was a practical evaluation of students performance in Santiago hospitals. Students were required to spend, after completing their internship, seven days on a medicine service and seven days on a surgery service. In addition, they were required to take written tests on anatomopathology and pathophysiology. Once Part one had been successfully achieved, candidates were required to complete Part two, which included a theoretical evaluation of medicine surgery and either
pediatrics or obstetrics.

Since the examination was given in Santiago, students from schools of medicine other than the University of Chile had to travel to Santiago to fulfill this requirement. Over the years some modifications were introduced and the time was reduced to one morning and special committees were nominated to administer the examination at the various universities. This procedure was followed until January, 1981, date of abolition of Law N 11.861. Authorization was then given to each individual university to license their own medical graduates (1981).

Many medical educators in Chile testify that through this examination forced all medical schools in the country to align their curricula with the University of Chile, and served as a selective barrier, it could not be considered as a national medical examination for several reasons. First, the examination was not the same for all students, the content of each examination was different. Second, the evaluators rotated, thus were different for each student; especially during the stay in medicine and surgery services. Third, patients and pathologies were different for each student. Fourth, the examinations' conditions, i.e., place, time, anxiety, were different for those living in Santiago than for those coming from the provinces to the Capitol. Finally, the examination tested an extremely limited range of intellectual abilities. Therefore, many types of learning remained untested.

The only national examination officially established in Chile is the Prueba de Aptitud Academica (Academic Aptitude Test), which is a College Entrance Examination that has been in use since January, 1967
(Pacheco, 1979). This test was designed on USA College Entrance Examination models.

The abolition of Law N11.861 permitted the universities to obtain academic independence, to design, develop and implement their own curricula, instruction and evaluation systems to graduate and confer M.D. degrees and, in effect, licenses to practice medicine.

This independence, though desired by the universities, places additional responsibilities on the medical schools, since external academic control diminishes. Therefore, the schools of medicine themselves are responsible for assuring the quality of the delivery of health care that their graduate physicians will provide. This fact is seen by many medical educators as another compelling factor for correcting existing deficiencies in current evaluation processes.

The recommendations about homogenizing evaluation in medical schools by means of a national medical examination made during the Third Seminar on Medical Education, have not produced any positive impact on evaluating the quality of physicians graduates in delivering health care to patients. Moreover, only few isolated studies have attempted to generate some information on national medical examination and quality of the physicians graduates during the past eight years.

For example, in 1977 Venturini et. al. (1977) reported interesting data relative to the evaluation of general practitioners' perceptions about the training received at their schools of medicine as an indirect way of measuring the extent to which the medical training enables the physicians to initiate his professional work.

Venturini et. a. conducted a second study in 1984 as a result of
the changes that had occurred in Chile's National Health Service in the area of primary care. The purpose of the study was to obtain information about the adequacy of physicians' training received at the various medical schools and the capacity to practice general medicine at the primary level. In general the study reports that the physicians surveyed considered that the training received in internal medicine and pediatrics to be good, but the training received in areas that require psychomotor skills was less adequate (Lavados, 1985).

The subject of a NME was again seriously discussed during the June 25-27, 1984, Seminar on Medical Education held in Chile. Medical educators representing the Chilean Association of Colleges of Medicine, and many other medical institutions throughout the nation, again made public their intention and desire to establish a National Medical Examination. This intent was to address the deficiencies of the present medical education evaluation process, the lack of valid data about the quality of their graduates, the information provided during the Seminar about the state-of-the-art of evaluation in medical education, (Lauados, 1985) the abolition of Law N11.861, and the excessive weight given by the National Health Service to the ability to recall memorized information in selecting candidates for positions as general practitioners.

PROTECTING THE PUBLIC INTEREST: A CONCEPTUAL FRAMEWORK

To assure the protection of public health, safety and welfare is one of the most important goals of any society. Educational institutions play a very important part in the attainment of such a goal since they are in charge of the education and training of the
professionals who will serve the public. Society through its organizations and with the cooperation of its members, also plays an important role in providing protection of the public interest. Figure 1 shows a theoretical model of occupational regulation for public interest protection.

FIGURE 1
PUBLIC INTEREST PROTECTION THEORETICAL MODEL
To ensure the protection of the public interest is an issue that is constantly influenced by the interaction of the public/consumer, the government and the professions. They all contribute in order to maintain a balance of individual interests, i.e., public, governmental and professional.

To accomplish the goal of ensuring the protection of public interest, society uses several mechanisms. Educational programs and institutions adequately accredited constitute the first mechanism. E. Stokes "states that accreditation may ensure that the right environment is present for good education to occur" (1980).

Accreditation according to Bratton (1980) is defined as: "the process whereby an agency or association grants public recognition to a school, college, university, or specialized study program that meets certain predetermined qualifications or standards".

Accredited educational institutions have the responsibility of adequately preparing students to practice the profession in a safe and efficient way. The concept that the quality of the graduates depends on the quality of the educational program is widely accepted today. It is an example of voluntary self-regulation of professional preparation by peer professionals.

A second mechanism to ensure the protection of public interest is licensing. Benjamin Shimberg (1982) states that those who meet the training and experience prerequisites, should be required to demonstrate on an examination that they possess the knowledge, skills and other abilities for the level of performance attested by the credentials: A license according to Donald H. Stewart, permits "a business to occur,
serves as an informational tool for the consumer, and to some degree implies a greater degree of knowledge and skill than that available from the unlicensed" (1982).

The U.S. Department of Health Education and Welfare defines licensing as:

A process by which an agency or government grants permission to an individual to engage in a given occupation upon finding that the application has attained the minimal degree of competency required to ensure that the public health, safety, and welfare will be reasonably protected (1971).

Licensing is regulated by government and is delegated, as in the United States and Canada, to state medical boards who are in charge of all aspects of the licensing process.

Certification is the third mechanism to ensure protection of public interest. Donald Stewart (1979) states that specialization should occur and certifying boards are needed to endorse the reliability and truth of the statement that an individual is a specialist. Besides the public in general and employers in particular should be able to identify those practitioners who have met a standard that is usually set well above the minimum level required for licensure.

Certification is defined by the U.S. Department of Health, Education and Welfare as:

The process by which a nongovernmental agency or association grants recognition to an individual who has met certain predetermined qualifications specified by that agency or association. Such qualifications may include: (a) graduation from an accredited or approved program; (b) acceptable performance on a qualifying examination or series of examinations, and/or (c) completion of a given amount of work experience (1971).

This action is another example of voluntary regulation by peer colleagues.
The last mechanism to ensure protection of public interest is reexamination, as a means to assure professional continued competence. The concept of reexamination and/or recertification and relicensure emerges from the belief that professionals may lose expertise and may not keep up with the new techniques and advances in the profession as the length of time since completion of training increases.

Shimberg (1982) states that increased concern about continued competence has given rise to demands that steps be taken to assure the practitioners have kept up-to-date and have maintained their skills (1982). This has prompted an enormous interest in continuing education as a means to ensure that professionals maintain competence.

The final intervening force of the theoretical model of occupational regulation to protect the public interest is the permanent feedback that must be provided in order to correct and improve the model.

MEDICAL LICENSURE

Better health care is clearly a national priority of the highest order throughout the world. The achievement of this goal requires more and better trained physicians who maintain their competence and better mechanisms for delivering health care.

Medical licensing as a procedure to ensure protection of the public's health, fits within the theoretical model represented in the preceding paragraphs of the conceptual framework.

The medical licensing concept emerges from the necessity to protect the public against quacks and charlatans that were so abundant sometime ago and that still exist to a lesser extent today. According to
Seigerist (1935) medical licensure as an institution became general in Europe during the Middle Ages.

The evolution of medical licensing can be summarized as follows: first, preceptors licensed their students to practice. Later medical societies assumed the power of licensure; this function was then partially taken over by the universities. Then licensing boards or other governmental licensing agencies came to regulate medical licensing, assessing practitioners' competence by the administration of an examination of high quality. Finally the highest stage of the evolution comes, reexamination as a response to public concern for assurance of continued competency in the delivery of health care. Figure 2 shows five theoretical medical licensing procedures ordered hierarchically.
LEVEL | MODEL | COUNTRY
---|---|---
V | State/or Government Examination + Accreditation + Reexamination | USA 1986 Canada
IV | State/or Government Examination + Accreditation | Europe Japan
III | No State/or Government Examination + Accreditation | Africa Asia Europe
II | No State/or Government Examination No Program Accreditation | Latin America
I | Apprenticeship | 1600's

FIGURE 2
MEDICAL LICENSING TAXONOMY
The apprenticeship model (I) is based primarily on the preceptor ability to evaluate. After the students complete the apprenticeship, the preceptor is the sole authority for admitting him/her to practice. Hubbard (1978) states that:

We are inclined to disparage the apprentice system in the light of today's formal education and legal controls over the practice of medicine, but - with the current large number of students going through the educational system and with the resulting difficulties in providing for close association between the student and the master of the art and science of medicine - we might well give some thought to its values.

The second model (II) is the one based on the completion of the educational program with no final competency examination. The assurance of physicians' competence rely on students transcripts. This model does not include educational program accreditation as a procedure for self-regulation of the educational institutions.

The third model, (III) though it does not use any final qualifying examination, the educational institutions and educational programs undergo accreditation procedures, which assure to some extent quality of training.

The fourth model (IV) is based on a state or government qualifying examination as a licensing procedure and based on a complete accreditation program. The fifth model (V), state/or government examination + accreditation + reexamination is at the top of the taxonomy and is based on the belief that competence should be continued so reexamination procedures are required or highly recommended.

It might be thought that there are as many licensing procedures as there are countries, however, a review of 68 countries' licensing procedures showed that the predominant models are no-examination - no
accreditation (II) and state/government examination with no academic program accreditation (1979).

In Latin American countries the licensing model is non-examination, no-accreditation. This means that the evaluation conducted by the medical school prior to granting the degree of medical doctor is what effectively counts in allowing the physician to practice in the medical profession (NBME, 1976).

In Europe, eight countries use the model state or government examination as a medical licensing procedure. Details about accreditation were not found. Nine countries use the model no state/or government examination. Accreditation is not mentioned either and six countries required a year or more of practice after completion of the educational program before a license is issued.

In reviewing 28 other countries' licensing procedures from Africa and Asia (World Directory of Medical Schools, 1979), the trends can be summarized as follows: eighteen countries use no state/or government examination and no accreditation was mentioned. Two countries required thesis. Six countries required one or two years of practice after completion of education program. Japan and USSR required a state/or government examination. No countries reported using the apprenticeship model.

Medical licensing in the United States seems to have reached the highest level of the taxonomy. At present, the model: state examination accreditation reexamination is currently used. Two types of examination are accepted by state medical boards for the purpose of licensing physicians to practice. The National Board of Medical Examiners (NBME)
and the Federation Licensing Examination (FLEX). The latter resulted in response to the expressed interest of state medical boards in achieving greater uniformity of standards for medical licensure, and to have an objective, standardized and validated examination that each member board could utilize as its qualifying assessment for licensure (1977).

Canada's licensing systems is also at the highest level of the taxonomy and closely resembles the one used in the USA. In Canada, the Medical Council of Canada, established in 1912, under governmental auspices and representing schools and Provincial Councils, designs and administers an examination for licensing purposes (Neufeld, 1985). This examination attempts to assess a candidate's factual knowledge of clinical medicine and his ability to solve and treat clinical problems. The licensing process in the United Kingdom is even more closely tied to the educational area than the systems described above. Acquisition of a medical degree is tantamount to a license to practice medicine, since the central licensing authority (General Medical Council) establishes national standards and imposes general conformity through accreditation of the educational system. Even if it does not conduct an external evaluation of the candidate, this licensing system is also considered to be at the highest level of the taxonomy. However, before obtaining an MD degree candidates must pass a final examination with external examiners from other universities who set a common educational standard (Neufeld, 1985). The General Medical Council first issue a Professional License which permits one year of practice under supervision in a hospital setting. Following this year of approved training the General Medical Council issues a full license to practice medicine.
In Chile the granting of a MD degree relies at present upon the completion of the prescribed program of study and relies upon the evaluation expertise of the medical faculty for assurance of students' competence. There is no accreditation programs for undergraduate medical education and the authorization to practice medicine and deliver health care is automatically given when one is granted an MD degree. It is interesting to note that until 1981 physicians in order to legally engage in professional practice were required to register at the National Medical Association. This registration was equivalent to a legal license to practice medicine in Chile without a required licensing examination. This procedure was abolished in 1981 by a decree of the Ministry of Justice (1981).

PURPOSE OF THE STUDY

The main purpose of the quantitative approach was to determine to what extent medical educators in Chilean medical schools prefer establishing a National Medical Examination (NME) as a means to physicians' licensure. This study also intended to determine Chilean medical educators' perceptions about what type of organization should design and administer a NME, and what areas should be assessed in a NME. Medical educators attitudes toward the current evaluation methods employed in assessing students' competencies were also be determined in this study.

The main purpose of the qualitative study was to find answers to the questions related to what may have been the factors that have inhibited the development of a NME in Chile as it was proposed several years ago. This qualitative approach also intends to determine
attitudes toward establishing, in Chile, a NME by means of interviewing some key people related to medical education and health care systems. Also, through this approach two methodological issues are explained: first, related to whether students should participate in the decision-making process about establishing a NME in Chilean; and second whether physicians are reluctant to respond to survey instruments.

RESEARCH QUESTIONS

Four main questions were asked in the quantitative approach and two research questions and two methodological issues were posed in the qualitative study. The questions for the quantitative approach are: (1) Do Chilean medical educators perceive a need for a National Medical Examination (NME) as an evaluation mechanism to assure medical graduates' competencies to practice medicine? (2) What type of organization or entity do they believe should design and manage a NME? (3) What areas of competence do they think should be assessed in a NME? and (4) What are their attitudes towards the current evaluation methods employed in Chilean medical education? The questions for the qualitative study are: (1) What factors may have inhibited the development of a NME in Chile? (2) What are the physicians' perceptions about the concepts of licensing, certification and accreditation? The two methodological issues posed for the study were: (1) Should students participate in the decision-making process about establishing in Chile a NME? (2) Are physicians reluctant to respond to questionnaires or any other types of survey instruments?
OVERVIEW OF METHODS

The methodology employed in this investigation was a combination of a concurrent quantitative and qualitative design. First, the quantitative design through a survey research study is designed to evaluate Chilean medical educators' perceptions about establishing in Chile a National Medical Examination (NME) as a model of licensure. The study was designed to determine: (1) the perceptions of medical educators about which entity or organization they prefer for the design and administration of a NME; (2) the areas that they believe should be included for assessment in a NME; (3) their attitudes toward the adequacy of evaluation methods currently employed in medical education in Chile; (4) the potential of an objective test; (5) their perceptions about the university being the organization or entity that should design and administer a NME if there is one; (6) their perceptions about employing practical examination with patients in a NME. Second, the qualitative study is designed to determine the factors that may have inhibited the development in Chile of a National Medical Examination and to evaluate understanding of the concepts of licensing, certification and accreditation. This qualitative study was also designed to find answers to two methodological issues related to the study: (1) should Chilean medical students participate in the decision-making process about establishing a NME in Chile; (2) are physicians reluctant to respond to questionnaires or any other type of survey? In addition to this the qualitative study was designed to determine key medical educators' perceptions about the topic under study by analysis of
open-ended comments made by some respondents to the questionnaires in the qualitative study.

**INSTRUMENT DEVELOPMENT**

For the collection of measures of Chilean medical educators' perceptions in the quantitative study a Likert-Type questionnaire was constructed comprising 25 items and four demographic-related items. For the collection data in the qualitative study, interviews, observations and analysis of open-ended comments were carried out.

**SCOPE OF THE STUDY**

This study was limited to Chilean medical educators of five out of six universities that have medical schools. The study could have enlarged its scope if the target population had included all practicing physicians of Chilean drawing from there the corresponding sampling. However, due to the magnitude of such an enterprise, plus the cost involved it was impossible to do it.

**GENERALIZABILITY OF THE STUDY**

The results of this investigation can be generalized to the majority of Chilean medical educators, perhaps with the sole exception of those academics from the Chile Catholic University whose Dean prohibited them from responding to the questionnaires.

**LIMITATIONS OF THE STUDY**

Three main limitations found in this investigation are: (1) the refusal of the Catholic University College of medicine to participate in the study; (2) the lack of students opinion about such an important
issue that is their direct concern and the scarcity of adequate studies in Chilean bibliographical resources related to the main topic under investigation.

**IMPACT OF THE STUDY**

The importance of this investigation are cited as follows: first, the use for a common, external system for evaluating and licensing in medical education is an increasing concern among Chilean medical educators and answers are needed. Second, medical students from the various schools of medicine throughout the country after several seminars about their medical education process have made public their concern about evaluation deficiencies. This fact is very important since during the last five years tuition and fees have increased significantly. Third, the issue about establishing, in Chile, a National Medical Examination or any other licensure procedure has been discussed formally and informally at many seminars on medical education. However, no formal studies have been undertaken. Fourth, with the abolition of Law N.11861, general public and medical educators are questioning the quality of the product being delivered to society since there is an increasing cost of health care services, and new policies have been delineated. Fifth, the results of this investigation will be used by the Chilean Association of Medical Colleges in their decision-making process about establishing a National Medical Examination or adopting other strategies to cope with the problem of evaluation in medical education. The findings may also be of help for other Latin American countries.
Finally, even though this investigation may not produce important changes in the evaluation and licensing approaches in Chile on the part of the medical schools, it would inevitably produce debate among medical educators, physicians in general and students. The study has already made a great impact among the Chilean medical schools and has brought the issue to the attention of the Catholic University who after discussion refused to participate in the study.
DEFINITIONS OF IMPORTANT TERMS

Competencies defined in Webster's dictionary: "having the capacity to function in a particular way or having requisite or adequate ability."

Clinical competence is understood as the ability of a physician to correctly evaluate and propose appropriate management of an undiagnosed ill patient. This is a minimal acceptance level of competence necessary for all physicians to be certified to practice.

Clinical skills relates to the ability for history-taking, physical examination, use of laboratory procedures, technical procedures, judgment, management and interpersonal techniques.

Key people are defined as a group of experienced and knowledgable Chilean medical educators whose judgment about Chilean medical education is valuable. They represent important institutions related to the health care and medical education field, for example, Ministry of Health, Chilean Association of Medical College, Offices of Medical Education, Schools of Medicine, and Public Health.

National Medical Examination is defined as an external evaluation mechanism, designed to assess Chilean medical graduates' competence to practice medicine and deliver health care.

Evaluation is defined as the process of delineating, obtaining and providing useful information for judging decision alternatives. The decision alternative may be: (1) medical students promotion; (2) licensing to practice in a supervised setting; and (3) certification to practice in an independent setting.

Medical educators are defined as full, associate and assistant professors of Chilean medical schools.
ENDNOTES


CHAPTER TWO
REVIEW OF RELATED LITERATURE

The review of related literature in this study intends to provide knowledge and evidence in the field of assessing medical students and physicians' competence in delivering health care and protecting therefore the public's interest. The specific goals of this section are: (1) to provide evidence of the historical concern of the medical profession, societies in general and the public about the issue of competence in practicing medicine; (2) to provide a brief description of the continuum of evaluation in the medical education process; (3) to discuss the serious efforts made by medical educators in developing and applying strategies to determine the areas of competence to be assessed in medical education; (4) to review some of the most widely used methods of assessing competences in medical education and its relationship to licensing programs; and (5) to discuss the role of external examinations as licensing mechanisms and its possible influence in developing licensing procedures in other parts of the world, such as in Chile.

The review of related literature is organized under five major headings: (1) "Historical Perspective of Evaluation in Medical Education," which points out a few important landmarks in the evolution of evaluation in the field of medical education. This section also presents events that show the constant concern of physician, public in general and specially medical educators for improving the mechanisms of
assurance of professional competence; (2) "Continuum of Evaluation in the Medical Education Process" in which a model is presented in order to facilitate the understanding of the importance of evaluation along the continuum of the medical education process; (3) "Determination of Medical Competences" in which the different methods for determining the competencies that should be measured by the evaluations methods are described; (4) "Methods of Assessing Competences in Medical Education," in which the various methods and their measuring properties are discussed briefly with the competencies; and finally (5) "External Evaluations in Medical Education" which discusses the examinations used for licensing to practice medicine and certification to specialists.

Historical Perspective of Evaluation in Medical Education

"The farther you look back, the farther forward you can see."

Sir. Winston Churchill

This section of the review of related literature [for this study] aims to provide a fruitful perspective about the concern for evaluation of medical students and physicians' competencies capable of contributing to analysis of contemporary problems.

Evaluation is not a new concept. One dictionary definition of evaluation is, "to determine the worth of; to appraise." Given such broad focus for the term, it can be argued that evaluation has been with us always and that everyone is, in his own way on evaluator.

The history of formal evaluation is much longer than is generally
recognized. The concept of evaluating individuals and programs was
evident as early as 2000 B.C. when Chinese Officials were conducting
civil service examinations (Dubois, 1970). Greek teachers, such as
Socrates, used verbally mediated evaluations as part of the learning
process.

Evaluation in medicine is not a new concept either. About 400 B.C.
as it was an actual question, who should learn medicine and who not was
in the mind of scientifics. At the same time another question called
for attention. Who would be, in every respect, the best type of
physician? (Kudlien, 1970)

Galen's treatise on medical examination seems to reveal the
importance of the concept of evaluation as he says, the future physician
must prove that he has "learned" fully not only practice but also
theory. This comment made by Galen (Kudlien, 1970) does not differ much
from the ones heard at the medical education seminars and conferences
today.

Medical educators and physicians in Islam were concerned about the
quality of medical care, and gave special attention to fighting fraud
and quackery in medicine. Al-Razi, a pioneer in the medical field and a
great medical educator deplored in his writings the ignorant
practitioners who deceived the public. He worked hard to uphold
professional standards (Sami Hamarneh, 1970).

Moreover, in Al-Malaki first treatise by Al-Majusi (Arabic),
Chapter Five is concerned with the testing of the competence of a
physician to beware of the charlatan and honor the competent and
virtuous. Earlier, Al-Mugtadir, the Caliph in 931, issued and enforced
an edict that all physicians in Baghdad should take and pass an examination by Sinan before receiving a license to practice (Sigerist, 1935). This situation does not differ much from the present situation in the USA, Canada and other countries.

Medical education in the Middle Ages did not consist of memorizing textual material or assimilating the words of the teacher. There were discussions between master and pupil themselves, to ensure that the student thoroughly understood what they had learned (Talbot, 1970). This resembles a typical up-to-date formative evaluation. Historically, if we look back at the beginning of the 13th century formal examinations were instituted by the King's orders at the medical school in Salerno:

"in order that the King's subjects should not incur dangers through the inexperience of their physicians" (Hubbard, 1971). These early official expression of the public interest in the examination of physicians touched on what continues to be a common source of anxiety today. How can a patient be sure that his physician is worthy of his total trust? Is his physician really qualified and can he be sure that having once achieved qualification, he continues to deserve it? How can these questions be answered and how were these questions handled in Salerno?

Medical licensure as an institution became general in the European Middle Ages (Sigerist, 1935) when medieval regulations were issued in connection with the first medical faculty of the western world, with the school of Salerno. In 1140 the Norman King Roger issued an order stating:

Who from now on, wishes to practice medicine has to present himself before our officials and examiners, in order to pass their judgment. Should he be bold enough to disregard this,
he will be punished by imprisonment and confiscation of his entire property. In this way we are taking care that our subjects are not endangered by the inexperience of the physicians. Nobody dares practice medicine unless he has been found fit by the convention of the Salernitan masters (Sigerist, 1935).

Today the issue of malpractice brings about a loss of trust which often results in legal suits against the practitioners brought by patients. Newspapers headlines present evidence of the increased public and legislators' concern for malpractice occurrences.

Over the centuries (Hubbard, 1978) scholarly theses, and written examinations in essay form were added to the oral questioning Salerno type of evaluation; and reliance was placed upon various combinations of the essay examination, the oral examination and subjective judgments to evaluate the individual that is the product of the medical education system. By the 14th century in Italy the granting of the Baccalaureate, at the end of fourth year of medical school, and Licentiate were subject to the same conditions, namely oral examinations by masters of the medical faculty in the presence of the masters of the higher faculties, theology, law and arts. The introduction of clinical teaching within the medical school so that the students as a group accompanied their teacher to the bedsides of the ill, both to observe their physicians' teacher's methods and to practice those methods under his observation and direction, was a phenomenon of the 16th century and a product of the Italian medical schools (Belloni, 1970). This development indicated the onset of what has become known as formative clinical evaluation.
Over the centuries medical education and practice has long been of concern to the general public and official authorities in most part of the world. For example, in the 18th century in France. The edict of 1767 condemned "the relaxation which has taken place in some of the faculties of medicine either in regard to the duration and quality of studies, or in regard to the number and nature of the examinations leading to a degree" (Coury, 1970).

In 1789 numerous petitions of complaint revealed the deplorable level at which the medical profession had fallen in the country as a whole. The inhabitants of Mantreuil demanded "that a young man who has hardly any notions of surgery and medicine not be able, for a hundred ecus, to obtain a license to kill with impunity, to cripple men throughout the countryside and take away a father or mother from unhappy children, that he be prevented from exercising these arts without previous examinations attesting his capacity" (Coury, 1970). This is an early version of dealing with malpractice and with public concern for the quality of physicians' competence performance.

In Germany during the 19th century many curricula existed. They differ from university to university even within one state, as documented by Stubler (Simmer, 1970). Not even within a medical faculty was full agreement always reached among professors as one can read in Bischoff's pamphlet of 1848. In a way the students were ordered to learn, since they had to pass examinations: the tentamen philosophicum in basic sciences, and the faculty and state examination at the end of the study. Here we can see an early example of the influence of an examination on curriculum and on medical education process. It is an
early intent of achieving homogeneity through an examination in an educational systems based on academic freedom and diversity of curricula.

Aragon, under Alfonso II, was the earliest Spanish realm to make compulsory the examinations of apprentices prior to practice. The appointing examiners for medical and surgical apprentices by the Spanish Kings frequently led to conflict of authority. The Protomedicato was established in Spain in 1479, this medical board was empowered to license and control medical matters, including examination of apprentices and medical graduates. The Protomedicato was established in nearly all Spanish colonies in America with the same authority. Latin American countries received the influence of the mother land and used the same procedures.

HISTORICAL PERSPECTIVE IN USA

In colonial America the controversial William Douglas, MD of Boston in the mid 1700's, nearly at the end of his career assumed to speak for all the colonies when he asserted:

". . if we deduct persons who die of old age, of mala stamina vitae or original bad constitutions, or intemperance, and accidents, there are more die of the practitioner than of natural course of the distemper under proper regimen. The practitioner generally without any considerable thought fall into some routine method, and medicines, such as repeated blood-letttings, opiates, emetics, cathatics, mercurials, Peruvian bark" (Norwood, 1970).

He reasoned at that time that in order to prevent malpractice, there should be regulation of the profession.

The existence of native medical schools in the USA made more acute the old problem of standardization practice by licensing procedures
(Shryock, 1966). The local medical societies, and the first provincial society (New Jersey, 1761), all wished to distinguish between the better-grade practitioners and the those with little or no training even by the apprenticeship.

In colonial days and for sometime thereafter in the USA, reliance upon apprentice system was about all that there was to give the patient some sort of confidence in the knowledge and skills of the doctor. Genievier Miller has given a vivid example of the manner in which the student was evaluated and deemed competent to enter the practice of medicine in colonial America (Miller, 1976): "These Presents will Inform All Whom are Concern" that Mr. Cary Henry Hampton of the county of P. William in the colony of Virginia hath compleat his Apprenticeship to my instruction in the Arts & Sciences [sic] of Anatomy, Chirugery, Physics and Midwifery to all of which for the space of four years he hath been Studious & Diligent. He is well grounded in the teachings of Cheseldens Anatomy, Hesiers Surgery, Cullens Materia Medica, Smellies Midwifery, three Works of our Masters Sydenham & Hippocrates which he hath read in the Latin tongue, as well as many other books of our Proffession and in the Instruction I have give to him at the beds of my Patients & elsewhere. So I repose my Confidence in his Knowledge & Recommend him to all those who require his Skill & Services. Given under hand & seal this the first day of Aug. 1775.

Anderson Roberson Doctor in Medicine

(Hubbard, 1978)

American practitioners and educators as early as 1827 were concerned with medical education defects and delegates from medical
societies in New York and the New England States met in Northampton and passed a remarkable set of resolutions designed to elevate the standards of medical education, but no school had the courage to embark on the recommended program. There is a similarity of this situation with the Chilean situation about establishing a National Medical Examination. It has been proposed several times and no action has been taken. The issue had been dealt with in several seminars and conferences over the years starting in 1960 but no one has ever taken leadership in doing something about it.

Back to colonial America, agitation for reform gradually grew and expanded, and became a major factor in the founding of the American Medical Association in 1847. This is an example of our modern control mechanism that professions use to regulate themselves.

In Chile until 1973 the medical profession through its National Medical Association used to be a peer control mechanism and at the same time was the institution that gave legal authorization to physicians to practice medicine and deliver health care. This institution to some extent controlled indirectly the peers' competence. A decree issued by the Chilean Ministry of Justice in 1981 took away such prerogative from the Medical Association. Presently the control of physicians' performance and the ethical issues are handled by the law (1981).

In USA in 1877, a group of medical school delegates met in Chicago and formed the American Medical College Association, later Association of American Medical Colleges (1890).

As it has been stated before, the concept of evaluation is not new, and evaluation in medical education can be considered as old as
evaluation itself. It can be stated with certain confidence then that the issue of protecting the people's health and assuring physicians' competences has been ever present in the medical education process and in the hearts of the institutions associated with educating physicians and with the ones that provide health care. Also it can be stated that the public, the professions and governments have been always concerned about the medical profession and about setting the standards required in order to assure competence of individual practicing physicians.

By 1910 Flexner report had an impact upon both medical education and evaluation. His concern for the quality of evaluation can be summarized by his own words: "no applicant for instruction who could pay his fees or sign his note was turned down. State board were not yet in existence, the school diploma was itself a license to practice. The examination, brief oral and secret, plucked almost none at all" (Flexner, 1910).

In 1930's some faculties were worried about medical education being over-regulated by AMA ratings and requirements of state licensing (Shryock, 1966). Some faculties have even voiced a hope that licensing power could be returned once more to schools, in order to permit free experimentation with curricular and education methods. This demand did not become widespread, in part because of memories of the chaos of the days when schools did control licensing (Shryock, 1966). This issue of controlling licensing and the like has relevance to day in Chile since in 1981 the Schools of Medicine were given power to license their own students and this fact has put some pressure in the idea of establishing a NME as an external controlling force.
TESTING DEVELOPMENT

By 1950, Psychometrics, the science of testing had come of age and more reliable methods of examination became available. Multiple-choice examinations began to replace the essay examinations in the medical licensing and qualifying examinations in the USA. What is new in evaluation from a historical perspective is the introduction of accurate, valid, and reliable measurements of medical knowledge and competence. Despite early critics, evidence accumulated to demonstrate the gains that had been achieved in reliability and validity of multiple-choice examinations.

Medical school faculties began to recognize their usefulness in measuring their students class-by-class and subject-by-subject and deriving thereby useful information not only about the individual students but also about the effectiveness of the educational process.

One further point about the introduction of objectivity into evaluation of medical education, is the attempt to evaluate clinical competence through the programmed testing technique, better named patient management problems. This objective examination was first developed by Rimoldi (Hubbard, 1971). Subsequently, adaptations of the method have been developed by George Miller and his associates Christine McGuire and the NBME.

Objectivity has been introduced into evaluation of medical education for students and faculty at the undergraduate and graduate levels, both for the product and process of medical education. At present we have more reliable methods to assure the competence of a physician. Questions asked some centuries ago are still critical at
present: how can people be assured of the competence of their physicians? Is there any justification for a person certified as a competent specialist 35 years ago, to be so listed today? These questions bring about the issues of licensing, certification and continuing medical education. The issue of continuing medical education was already foreseen in 1971 when it was stated that it is no longer a function of the desire of the physician to maintain an understanding of current affairs in medicine (Samph, 1979). Today a physician's involvement in continuing medical education includes going to local hospital, staff meetings, attending local, regional and national medical society meetings and home study books, medical journals and other special materials.

CONTINUUM OF EVALUATION IN THE MEDICAL EDUCATION PROCESS

The protection of public health is a continuous process that begins with the selection of the individual wanting to pursue the profession of medicine, continues through undergraduate and graduate medical education and continues during the entire practice of practitioners through continuing medical education. The medical education process has been defined by Samph (1979) as a process by which physicians are educated to develop the knowledge, skills, and attitudes that are necessary for the delivery of effective health care to patients. This process in USA begins at the moment of admission for college level which takes generally four years then is followed by medical undergraduate school that also takes generally four years, then the internship and residency and continues throughout the continuum of physician's lifetime of practice. A similar system exists in England, Scotland and Canada
(Cristofanini, 1982) where the two or three years of pre-medical school are oriented to the study of basic and biological sciences. In Spain, France, Italy, Singapore and Japan undergraduate medical education takes six years plus one year internship (Cristofanini, 1982).

Undergraduate medical education is designed to prepare the undifferentiated physician; graduate, the differentiated; and continuing medical education to update medical skills and knowledge of practitioners. In USA the undergraduate level training terminates with the awarding of the MD, DO degree, and at the graduate level, which takes three to four years, with the awarding of board certification. For relicensure in continuing medical education, physicians are required to obtain 60 hours of training.

Graduate medical education terminates with the awarding of board certification as it was stated previously, and all physicians are expected to undergo some graduate training.

Chilean medical education process follows approximately the same patterns. Some differences exist as it would be expected. First, the internship is not considered graduate level; it is included within the undergraduate program. Second, only a small number of graduates from the medical schools have the chance to go into residency. Vacancies are very scarce.

Evaluation has played and still plays a very important role in the medical education process. Much attention has been given to it, and its evolution is marked with refinements and improvements.

Due to the rapid growth in new knowledge and medical technology, due to the public expectation of health care and the fact that medical
education process continues throughout physician's life time the medical profession is aware that it is imperative to establish mechanisms to assure the qualifications of individuals to assume clinical responsibilities at the appropriate stages (Samph, 1979).

Evaluation throughout the continuum depends upon a series of repeated observations of the performance of an individual trainee. Evaluation of applicants to medical schools constitute a critical stage in assuring students' competence and in protecting the public health. Rezler (1983) summarizes the characteristics found to be important in the selection of applicants to medical schools. First, measurement of intellectual development are established. Graduate Point Average (GPA), class rank, aptitude and achievement scores are good examples. Second, assessment of intellectual motivation, attitudes, values, personal adjustment and social habits are included in the evaluation. The information for this category is obtained through interviews and letters of recommendations. Third, evaluation of biographical data such as age, race, family status, place of origin, volunteer activities, work experience, and specific career goals is carried out. Fourth, assessment of psychomotor skills is performed. Likewise, the opportunities for evaluating individual performance in medical school provide for a more reasonable estimate of an individual's competence is those areas necessary for affective performance as a physician. The granting of the MD degree according to Samph (1979) is a reflection of the institution's recognition of an individual's competence to move on into graduate specialty training.
Figure 3 provides a graphic display of the continuum of evaluation in the medical education process used by Samph (1979).

**FIGURE 3**

**CONTINUUM OF EVALUATION IN THE MEDICAL EDUCATION PROCESS**
DETERMINING COMPETENCIES IN MEDICAL EDUCATION

One of the basic issues throughout the history of medical education has been the concern for the development of accurate methods for determining competencies. The problem was not scientifically addressed until 1960. Techniques of assessing professional competence were commonly based on the intuitive and usually implicit judgments of senior teachers, supervisors, and leaders about what was important in their respective fields. Clearly, such judgments varied widely among experts and were necessarily biased by the often narrowly focused interests and experience of each, however, more systematic methods to develop more valid definitions of the essential components of professional competence have been proposed (McGuire, 1983).

The task of determining the dimensions of competence, regardless of the method chosen, is described in the literature as the detailed specification of the necessary knowledge or skills of the competent practitioner. These dimensions are generally arranged according to domains of competence, such as knowledge, technical skills, and interpersonal skills. Reasons for defining competence evolve around the improvement of planning and evaluation along the continuum of medical education. Having an organized overview of what the product of a medical education program should be able to do allows those who have responsibility for the program, or for program accreditation, to recognize significant gaps in the individuals with the requisite competence to perform adequately. A definition of competence can make goals of medical education clearer to students, teachers, and the public. A definition of competence also helps to define the programs
that a student is exposed to and for certification if students have met the requirements of the program.

Burg and Lloyd (1983) write that the reasons for defining competence in medicine are several, but, from the perspective of American Specialty Boards, perhaps the most important reason has to do with ensuring the validity of their certification procedures.

Definitions of competence within a medical specialty discipline serve the purpose of providing a first step toward the development of more valid procedures for the certification of specialties. This is because one form of validity, content validity calls for a comprehensive delineation of the skills and abilities the board is attempting to measure. Ideally, measures of competence in a specialty should sample from the components of competencies identified as important by members of that specialty (or by empirical methods). Burg further states that a definition of competence also can serve each specialty board in providing a rationale for its evaluation procedures. Agencies with the responsibility for certification of physicians need to know and agree on just what are the skills and abilities they plan to certify.

Dr. C.F. Schumacher (1969) in his studies on validation of Part III Examination, the final step in National Board certification, developed a definition of clinical competences: "those skills and abilities which are acquired as the physician gains experience with patients and increasing responsibility for their care. Another attempt of definition was done by the task force of Study B in its reports to the research Advisory Committee, as "the ability of a physician to correctly evaluate and propose appropriate management of an undiagnosed ill patient for
whom he was the physician first contacted (Rogers, 1968). This
definition attempts to delineate a minimal acceptance level of
competence necessary for all physicians to be certified to practice.
Senior in 1974 tends to complete this definition by suggesting that more
advanced levels of competence in special areas of medicine might require
specific definitions and that different tests might require specific
definitions and that different tests might have to be devised. It was
suggested also that competence was scalar, rather than all-or-none.
Senior (1974) further states that there is a true barrier in reaching a
generally acceptable definition of clinical competence in medicine, and
which is "the inability to reach consensus on what constitutes a good
physician."

Neufeld (1985) suggests that the methods for determining
physicians' activity range from the writings of individual experts,
consensus of a group of experts to studies done in disciplines such as
psychology, and epidemiology. He also states that the earliest
definition of the profession, was conducted by the reflexive/
philosophical method which involved consideration of quality deemed
important by a single recognized authority, and generally appear in the
context of philosophical writings on textbooks on clinical diagnosis.
For example, Tyrer and Eadie (1976) have described one aspect of
clinical reasoning in the following manner:

"As more and more evidence of various sorts is accumulated,
the physician, by reasoning in relation to his knowledge, is
able to formulate possible diagnosis in his mind and then
narrow down the list of suspect diseases until a definite
diagnosis can be made. Should insufficient or invalid
evidence of facts be obtained . . . a correct diagnosis is
unlikely . . . Should the knowledge of the clinician be
defective . . . it would be difficult for him to reach a correct diagnosis . . or should his reasoning be defective he may be unable to isolate the correct diagnosis from the forest of alternatives presented to him. . ."

The review of the literature has provided other methods which rely more on consensus and on empirical studies. In contrast to the reflexive approach on the assumption that "professional performance is socially determined by the situation in which occurs," La Duca, Engel, and Risley (1978) developed a model - the Professional Performance Situation Model (PPSM) - for utilizing expert judgment in defining required skills and knowledge, for occupational therapy, clinical dietetics and medical technology. They found that for the three professions the skills and knowledge could be defined in terms of three dimensions: setting, client, and problem.

A task analysis approach to defining competences has been widely used in the medical field. Andrew (1976) states that it is a process whereby activities that physicians engage in . . . are documented and described in such a way as to make explicit the purpose of the activity, the procedures that the physician utilizes in performing the activity, and the outcomes or expected products of the activity. She further identifies three streams in task analysis: (1) the consensus of expert opinions; (2) the critical incident technique approach; and (3) the use of activity logs.

The approach, consensus of experts, in which those experts in a field are frequently used in defining the content of a discipline and the production of explicit statements of competence has been used by a number of Specialties, such as the American Board of Pediatrics (1974), The American Board of Internal Medicine (1979), The College of Family
Wigton (1980) in an attempt to rank order the variables that are "most meaningful in evaluating the performance of a first-year house officer on a Clinical Service, employed a modified Delphi technique to obtain a consensus among full-time faculty, volunteer faculty, and residents. He reported that this technique provides a list of variables that are important and valid for the faculty and residents involved. He found important differences among the three groups, as did Sweeney and Resan (1982), in an analogous attempt to have educators, employees, and new graduates define the essential skills for baccalaureate nurses. Schiappacasse and Ramirez et al. (1984) in an attempt to develop a Professional Profile of the Physician also employed a modified Delphi technique and they found that the competences of the physicians can be organized in five major categories: the personal qualities of the physician, professional ethics, professional competence, professional improvement, and the administrative aspects of the physician's activity.

The critical incident technique has played an historically significant role in defining clinical competence and has several methodological advantages, as the one that addresses the representativeness of the criterion group being consulted and deliberately chosen. Incidents are not deleted in the final summary, but are simply categorized, removing personal bias from this stage of the process. La Duca (1978) reproduces Flanagan definition of the critical incident: "by an incident is meant any observable human activity that is sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act. To be critical, an incident must occur
in a situation where the purpose or intent of the act seems fairly clear to the observer and where its consequences are sufficiently definite to leave little doubt concerning effects." The NBME in 1959 undertook to analyze and define the clinical competence that its Part III was intended to measure, and to devise testing techniques permitting objective, valid and reliable assessment of clinical competence (Hubbard, 1978). The study employed the critical incident technique with the cooperation of Dr. John Flanagan who had successfully developed objective methods for measuring skills of airplane pilots (Hubbard, 1978). The study involved 600 practitioners and a total of 3,300 critical incidents, described as "clinical situations in which they had personally observed interns doing something that impressed them as examples of conspicuously good or poor clinical practice." These incidents were then grouped and classified into nine major categories.

A similar technique was employed to develop a certification examination for the American Board of Orthopedic Surgery (Miller, 1968). This time involving 7,700 critical incidents contributed by 1,100 specialists and grouped into 94 sub-categories. Miller (1968) reporting about the same study stated that:

"the ultimate study goal was that of improving efficiency and effectiveness of the educational process, but the immediate objective was to assure the adequacy of methods by which professional competence in Orthopedics was assessed, an essential requirement if alternative training patterns were ever to be established with any confidence" (p. 602).

He further describes the critical incident technique as a systematic assembly of episodes observed and described by qualified experts, which they have judged to represent outstanding or shabby performance in the specialty.
Sanazaro and Williamson (1970) used the technique to study a number of primary care specialists involving a total of 2,000 physicians and 12,000 incidents. They used a modified form of the critical incident technique. They define critical incident as:

any episode of patient care in which one or more specific actions by a physician had one or more specific beneficial or detrimental effects on a patient. The term critical simply means that very likely, the physician's actions were directly responsible for the effects observed in the patients."

One of the major strengths reported by them is that the classification used is based on episodes of care in which specialists were reasonably certain of the causal relationships between physician's action and patient's effects. However, they stated that caution should be taken when dealing with representativeness as they stated:

the information cannot be construed as representative of medical practice at large. First, in the critical incident technique, only unusually effective or ineffective performance tends to be recorded and the reports are not a representative sample of actual experience. Second, the respondents in our study were not a representative sample of all specialists in their particular field (Sanazaro, Williamson, 1970)

Cowles (1965) modifying the technique, grouped a total of 2,300 critical comments retrospectively derived from written evaluations of clinical clerks and scored each comment in an attempt to develop an objective means of evaluating verbal statements. He reports: "In spite of individual faculty differences in number and adequacy of comments, the pooling of comments from a number of instructors did result in excellent verbal pictures of most of the students with much internal consistency."

An attempt to define clinical competence in a broader scope was done by Hubbard et al. (1965). By interviews and mail questionnaires,
requested physicians to record clinical situation. A total of 3,300 incidents were collected from approximately 6,500 physicians. These incidents were divided about equally between "good" practice and "poor" practice. Nine major categories were reported by the authors of the study: history taking, physical examination, test and procedures, diagnostic acumen, treatment, judgment and skills in implementing care, continuing care, physician-patient relationship responsibilities as physician.

Traditional activity analysis has not been extensively employed in the health professions (McGuire, 1983) for competence definition. However, activities questionnaires and log-diary have been used to some extent.

Mendenhall, Girard, and Abrahamson (1978) developed a highly structured log-diary form to gather data on the activities and practice characteristics of generalists and specialists of different types; responses were analyzed to yield a unique practice profile of 24 specialties and subspecialties. In 1972, the American Board of Pediatrics initiates studies to identify the important component of competency needed in the practice of pediatrics (Burg, 1976), and the study reports:

"An examination of the daily activities of the practicing pediatricians revealed that certain principles were operating. The pediatrician is faced with a variety of clinical situations, the management of which depends upon the performance of certain tasks. These tasks require the exercise of a variety of skills and abilities if the situation is to be managed properly. Thus, the overall measurement of a physician's competency is determined by the manner in which he uses his abilities to perform the tasks necessary to resolve the clinical problem."

In relation to the impact of the technique the study concludes:
The direct usefulness of the competency study is not yet established, but the ultimate impact that the continual dialogue on defining and assessing pediatric competency among knowledgeable pediatrics may have on the education of medical students, pediatric residents and pediatricians could be enormous (Burg, 1976).

Though the reasons for defining competences through the various methods, remains the same, however there are serious problems which have been summarized by Christine McGuire (1983): The critical incident technique does not supply information about the importance of behavior in terms of the frequency with which it will be required. On the other hand log diary studies and task or activity analyses inform us only about what is, not about what ought to be; nor do they tell us which of existing activities really make difference in health outcomes. The PPSM model, as it has been applied, relies only on the judgment of academicians and fails to incorporate more empirical data from practitioners. McGuire's words seem to be suitable to conclude this section:

Though we have not yet developed an adequate theory of the structure of professional competence nor totally satisfactory method of defining its constituents, the increasing systemization of the process for utilizing expert judgment and for incorporating relevant empirical sources in defining its dimensions, which has occurred over the past 20 years, constitutes one of the most important developments leading to its more valid assessment.

METHODS OF ASSESSING COMPETENCE IN MEDICAL EDUCATION

Medical educators during the last decades have been very much concerned with the development and refinements of methods for assessing the areas of competence in medical education. This section of the review of related literature focuses mainly on the identification and description of the different methods most widely employed in medical
The process of evaluating students' clinical performance/competence has evolved over the years to constitute the present practices employed in most medical schools. Abrahamson (1985) in relation to the phenomenon writes: "it is interesting to note that the entire evolution process [evaluation] seems to have occurred in the last two or three decades, with the most significant contributions coming very recently." Students, knowledge and skills in the various disciplines was measured in the past almost exclusively by essay-formats, as recently as the 1950's. Hubbard (1978) reports that the National Board of Medical Examiners (NBME) started to shift from essay to multiple-choice in 1950. Oral examinations and essay examinations were employed to measure clinical problem solving (Abrahamson, 1985), always considered a major goal in medical education.

The now-common patient-management problem (PMP) format - an outgrowth of the Rimoldi-Cowles test for problem-solving - was introduced as recently as 1961 and was not widely used for almost a decade (Hubbard, et al., 1965). For the assessment of clinical skills, one could frequently find something called "a practical examination." Unfortunately it took different forms in different places, and, indeed, it therefore defied description as an assessment of data-collection technique. The NBME used it in Part III, and the candidate was expected to examine a patient (doing both history and physical), write the findings in a booklet provided for that purpose and then be the object of an oral examination conducted by two examiners (Hubbard, et al., 1965). However, problems of validity and reliability forced the
abandonment of this practice.

One final component of clinical performance worthy of mention is that of essential habits of the good clinician. In the past in this area one frequently encountered the accepted use of impressions - impressions gained during repeated or extended encounters between student and teacher. Over time, reasonable clinicians came to understand and accept the need for a more objective means of assessing performance, globally or in components - and moved toward the use of rating scales and chart reviews to establish whether students were building good clinician habits (Abrahamson, 1985).

It is perceived through this evolutionary trail that the measurement of student knowledge has evolved from reliance on essay examinations to reliance on objective examinations to measure cognitive problem-solving, and the evolutionary process clearly demonstrates the acceptance and application of both rating scales and chart review to assess students clinical habits.

Abrahamson (1985) referring to the evaluation phenomenon states:

Today we are beset with problems that grow out the solutions to our earlier problems. In addition, we are in midst of trial applications of many forms of assessment of students' clinical performance, or components of students' clinical performance.

The present day tendency to use multiple-choice question solves an earlier problem of lack of objectivity. MCQ's are the main component in many certification procedures, intended to provide evidence that an adequate level of clinical competence has been achieved (Neufeld, 1985). The MCQ format is probably the most widely used method for assessing clinical competence at the present time (Neufeld, 1985). It
is used by the NBME, certifying boards, Flex, ECFMG, MKSAP (Medical Knowledge Self Assessment Program of the American College of Physicians, National Self Evaluation Program of the College of Family Physician of Canada). Robert Ebel is more emphatic to express his support for MCQ. He states: "the [multiple-choice items] are adaptable to the measurement of most important educational outcomes of knowledge, understanding, and judgment; of ability to solve problems, to recommend appropriate action to make predictions" (1979). McGuire and others have concluded that these examinations focus primarily on assessment of general knowledge. McGuire (1983) writes: "the most important innovation has been the increasingly widely accepted practice of designing items to a table of test specifications that includes, preferably, both cognitive skill and content domains." Levine, McGuire and Natress (1970) concluded that multiple-choice tests designed in such a manner are broader and test a more relevant range of competencies.

More recently, Maatsh at Michigan State University reported that the cognitive aspect of clinical competence can be understood to have two main components, a knowledge factor and a performance factor, and that these two factors are quite correlated.

Among the traditional written examinations, McGuire (1983) reports that: "Perhaps the most significant innovation has been the introduction of the Modified Essay Question (MEQ) and the Sequential Management problem (SMP) formats." Felleti (1980) reports in relation to MEQ that: "Tests in these formats are designed to assess clinical problem-solving skills by means of sequential presentation of a sometimes extensive data base about a particular case, accompanied at each stage by a specific
question." Similarly, Knox (1980) has employed MEQ based on actual patients in order to assess components of competence appropriate to the personal doctor providing primary and continuing care. Knox further suggests that MEQ test abilities at various cognitive levels: factual recall, application and synthesis, and that the content includes patient–doctor communication, antenatal care, and patient management including the involvement of health visit. Barrows and Mitchell (1975) at MacMaster University used a similar approach, and describe a learning package as problem box.

Patient Management Problem brought to attention to the medical community about 20 years ago by NBME and Christine McGuire has been incorporated into the assessment efforts of clinical departments in medical schools and as well as in the national standardized examinations procedures (Abrahamson, 1985). Hubbard (1978) describes in a very precise form this technique:

This method simulates a realistic clinical situation in which the physician must face the dynamic challenge of a sick or injured patient. As in real life, the examinee is confronted with a patient about whom brief but carefully selected information is presented. This information must be studied and decisions reached about appropriate action. Laboratory studies and diagnostic procedures may be required; decisions about therapy and management are called for."

McGuire (1963) reports that PMP is seen as an objective instrument capable of evaluating components of competence that could not be assessed by multiple-choice questions or other conventional test instruments. McGuire (1972) also reports that experts that have reviewed PMP's and students and physicians that have worked through them report on systematic inquiry that the PMP closely approximates the clinical setting. Moreover, Feightner (1985) states that there is
reasonable evidence for reliability both in linear and branching formats, and that in general, there is acceptable evidence for content validity. Abrahamson (1985) concludes that: "this technique is now commonly used device, despite its lingering psychometric problems mentioned by Hubbard in a publication in 1965 in the New England Journal of Medicine."

The increasing interest for simulation and the capabilities of today's computer, have produced computerized PMP, eliminating the disadvantages of cueing inherent in written simulations (McGuire, 1983).

Moreover, many of these computerized management problems now include extensive interaction that goes far beyond the rather conventional PMP (Abrahamson, 1985). It is interesting to note that with the increase use of microcomputers faculty members are developing their own PMP's. Norman & Paivin (1985) summarize(s) nicely the role of computer simulation to learning and evaluation:

The area [computer simulation] depend on future developments which maximize the technical potential of the hardware and adequately investigate the measurements properties of the instrument. It is encouraging that most of the studies confirmed a positive attitude toward the computer, but much work needs to be done before computer simulation have a confirmed role in the evaluation of clinical competence."

Although NBME claims to set up 22 computerized testing centers.

The Programmed Patient Technique was studied by Barrow and Abrahamson (1964) and they conclude that simulates patients can overcome difficulties inherent in traditional oral and observational techniques in evaluating clinical competence. The simulated patient technique has been adopted in settings that heretofore have been mostly conventional in all of their educational processes (Barrows & Tamblyn, 1980). The
results are encouraging as more and more medical schools are discovering that human situations can provide so much information about students' clinical performance (Abrahamson, 1985). In contrast, Norman (1985) points out that in spite of the apparent advantages, simulated patients have been infrequently applied to routine clinical evaluation in the undergraduate and post graduate years.

In the medical education field there are aspects other than knowledge that have to be measured - attributes, judgment, rapport with patients, etc., - and the way to assess them is not through written examinations. Rating Scales, though a not novel technique, is now finding sound widespread application in this respect at clinical level. Streiner (1985) states that: "Many medical schools have adopted to use global rating scales to supplement, and in some cases to replace, formal written examinations and course grades." These scales, he adds, most often consist of a series of adjectives or descriptive terms, each of which can be graded along a continuum from very poor to superlative performance. However, caution must be taken because there seems to be some aspects counter balancing the advantages, such as low or poor track of record about validity and reliability, and factors affecting the measurement, such as rater, scale, halo effect, etc. A question can be raised then: should this technique be abandoned or can it be improved? Streiner (1985) states: "the potential of the technique has not been realized and that there are changes that can be implemented that would result in improvement." McGuire states that to evaluate task performance "we shall probably need to continue to employ various techniques" (1983). In this respect Streiner (1985) writes that: "when
behaviors being rated is observable, as in performing an examination, checklists are preferable to global rating scales." Turner (1972) and Weissman (1975) report that the interrater reliability of checklists is unusually significantly higher than for global scales. The potential of checklist has been documented by Andrew (1971, 1977) in a study using the checklist to assess the physical examination skill of physician assistants.

The use of chart reviews/chart audits or medical record review is considered by Payne (1979) a potential source of evaluative data. Abrahamson (1985) states that the interaction that involves the student (or house officer) in such a way as to permit external review of the written documents for which that student (or house officer) can be held responsible or accountable, apriori criteria have been developed and applied in a sound approach to the assessment of certain clinical performance components. McGuire (1983) reports that many methodological problems remain to be solved before audit can become a cost-effective technique for general use in the evaluation of individuals. She further points out the principal problems: the nature of the standards to be used, the criteria and level of standards to be applied, and the number of records that must be audited to get a reliable estimate of performance. Samph and Templeton (1979) in reference to the same issue concluded that even if some of these major methodological difficulties could be resolved, judgments based on audits will continue to require long-term observation of many patients. However, some advantages have been recognized, especially educationally related. Peter Tugell (1985) reports that this technique has advantages that makes it suitable for
both summative and formative evaluation; it is based on real clinical behavior; it is experienced centered and relevant to the patient-care problems of the learner, encourages active participation and problem based approach. He further states: "record review is feasible to apply in the educational setting." However, the impact on physicians is not all clear. Nelson (1976) reports that studies using these methods don't consistently change physicians' behavior. It is important though to stress in this respect that reviews is only a measurement method for assessing clinical competence. Nelson, however believes that record review for medical students and residents has considerable potential.

Tugwell (1985) summarizes the strengths of record review:

1. The medical record is an established component of clinical practice.
2. Actual practice behavior can be assessed.
3. Complete episodes of care can be assessed.
4. It allows application of explicit criteria with high levels of consistency in abstracting the information.
5. It allows the reactive effect of being studied to be avoided of the clinicians are aware beforehand, this effect can still be minimized by keeping those being evaluated blind as to which conditions are to be selected.
6. Relative large numbers of cases with the same problems or diagnosis can be assessed, thus reducing the sampling problem that may result from basing conclusions on very few cases such as an oral examination.

Tugwell (1985) also points out some weaknesses:
1. The fact that records are used more as an aide-de-memoire rather than a documentation of the justification for management decisions, which continues to compromise the validity of the medical record.

2. The inconsistency with which actions are recorded.

3. The difficulty in deciding who made the recorded decisions where more than one health professional discusses the patient before the record is written.

4. The difficulty in maintaining blindness of the assessor to the clinician involved.

5. The logistical difficulties in identifying a homogeneous sample in ambulatory settings, where there is no incentive to implement methods for retrieving charts with similar problems and diagnosis.

6. The failure to emphasize the evidence for the efficacy of the patient care being assessed, which has reduced the credibility and meaningfulness of the results and conclusions.

There are some recent developments in the field of medical education evaluation which have been addressed by Abrahamson (1985) as the "more recent adventures in Assessmentland," even, he adds, in the full knowledge that these instruments and/or approaches are not unknown:

1. CBX. A computerized, interactive form used for the testing of clinical problem-solving and patient management that has been in development for some 15 years, ultimately to be used as assessment of clinical competence by the NBME.
2. **I.S.I.E. (Interpersonal Skills Interaction Evaluation).** The system involves having a candidate videotape an interview (with real or simulated patients) and send that videotape to NBME where it can be scored according to a set of carefully developed interaction-analysis criteria.

3. **Videodisc and Computer.** This combination has produced the capability to offer students the challenge of a clinical problem in both an interactive mode for obtaining information and intervening and visual mode that permits the students to see and hear the patient as well as obtain other visual auditory information about the patient. The same kind of interactive programs can be produced and used on videotape instead of videodisc.

4. **Simulated Patients.** Though introduced more than 20 years ago it is a devise only now coming more widely into its own. The simulated patient permits a standard set of stimuli to be provided for all students who are being "tested" as well as a totally unobtrusive source of information about the students' performance.

5. **OSCE, OSCA.** Technique that is receiving more recognition and attention is the Objective Structured Clinical examination (Objective Structured Clinical Assessment) though it has been used before. This form of assessment is generally limited to one particular component of clinical performance at each of several stations and usually requires a some observational data obtained through rating scales.
EXTERNAL EVALUATION IN MEDICAL EDUCATION

The role of external examinations in the development of medical education programs and students' evaluation is discussed from the North American perspective since at present it constitutes the most appropriate model. The aim of this review is to provide useful information that may be used by Chilean medical educators in their efforts to develop a licensure model.

American medical faculties use different evaluative examinations prepared by external agencies to assess either students' progress, the scope and quality of their educational programs and for licensing purposes. Most faculties are aware that their students must pass a licensing examination that is acceptable to the jurisdiction in which a license is sought to practice medicine. Historically, the responsibility for educating students in medicine and the responsibility to determine who shall be licensed to practice medicine have been separate and distinct (Association of American Medical Colleges, 1981). In the case of medical graduate of schools outside the jurisdiction of USA, faculties and the licensing agencies have relied since 1958 on the certification of these graduates by the Educational Commission for Foreign Medical Graduates (ECFMG). After reviewing the literature it is easy to conclude that there exists an interactive relationship among faculty and the authority to grant the MD and among the state medical boards and the ECFMG, it is easy to conclude that there exists an external examination generally is a comprehensive, multiple-choice examination, usually designed in a paper and pencil format and constructed by an outside source. Such tests are primarily cognitive
but contain the possibilities of the examination of minimal problem-solving skills. Kapelman, et al. (1983) in relation to the use of external examination state:

External examinations are used to evaluate medical students and the curriculum within the medical schools. These externally composed tests are also used to license trainees from foreign schools into the US graduate medical training programs and such examinations frequently are used to qualify the physician initially with appropriate credentials for specialty recognition as well as part of the process to recertify physicians.

He further states that the question of the use of external examination for licensure has come to forefront with the proposed use of Flex I and Flex II, single sequence by the Federation of State Medical Boards (Kapelman, 1983).

Examples of external examinations widely used in the medical community include the three parts of the examinations of the National Board of Medical Examiners (NBME), the Federal Licensure Examination (FLEX), the Educational Commission for Foreign Medical Graduates (ECFMG) examinations used as part of the certification for practice in USA, and the examinations of specialty boards.

In order to ascertain what should be measured by the examination, most licensing and certification agencies conduct some type of studies. For the medical education fields, methodologies for determining and evaluating competences were discussed above.

Among the methods used by licensing agencies Shimberg (1981) reports that a survey of assessment practices of 42 health certifying agencies revealed:

that 37 were using multiple-choice tests. However, many supplemented these with other type of tests. Sixteen agencies used essay tests, 15 used practical (performance) tests; six
used oral examinations, and two used paper and pencil simulations, often referred to in the medical field as 'patient management problems' (National Commission for Health Certifying Agencies, 1979).

It can easily be deduced that the format most commonly used in licensing and certification is the written multiple-choice test. These tests according to Shimberg are highly cost effective, particularly when the number of applicants to be tested is large.

Kapelman (1983) reports that external examinations are used to evaluate medical students and the curriculum within the medical schools. These externally composed tests are also used to license trainees from foreign schools into US graduate medical training programs and such examinations frequently are used to qualify the physician initially with appropriate credentials for specially recognition as well as part of the process to recertify physicians.

External examinations play important roles in the evaluation systems employed in medical education. The AAMC in the Annual Curriculum Directory (1981) reports that medical schools in USA use NBME to a considerable degree to evaluate curriculum. The reports further states that in 1980-1981, 84% of the schools were using the NBME tests to evaluate the overall curriculum. Joseph Sheehan during the AAMC meeting in 1981 warned about relying too heavily on external examination to evaluate curriculum. He insisted that this function of an external examination is only useful when objectives, learning experiences, and examinations are consistent with one another, since he added, it is common to find inconsistencies. Kapelman(1983) addresses to this issue in relation to the impact of external examinations on residency selection. He states that: "A major psychological impact on medical
education is the concern that their performance on external examinations will have significant relevance in their selection for placement in desired residency programs." However, in a poll of 179 residency programs directors in family medicine, pediatrics, internal medicine and surgery, results of the NBME Part I Examination were ranked 23rd among factors considered important in selecting medical students for residencies (Wagoner & Gray, 1979). Spellacy's study (1985) reports that resident's performance on the Council on Residents Education in Obstetrics and Gynecology (CREOG) examination can be predicted by prior performance on the national board examination, especially Part II of NBME. Therefore the study concludes that:

the Residency Review Committee should consider evaluating the scores on Part II of NBME of the residents in a program as a new index of the program's quality. Second, that program directors can use the results of Part II of NBME to select residents who will do well in learning the cognitive aspect of specialties" (Spellacy, 1985).

Eichna (1983) in addressing the effect of external examination on medical students reports that the impact of external examinations on medical students is substantial, although varied, since the stress can be facilitative, but excessive stress becomes destructive. Kapelman (1983) in relation to the same issue states that:

it may sweep across the medical school landscape twice during the students' career. Emanating from a foreign source, the external examination frightens the medical student because it obviously is not only assessing his knowledge but also comparing the teaching effectiveness on his medical school against national normative figures with him as the front basis of the external examination.

Eichna (1980) in the New England Journal of Medicine posed the following question related to the impact of external examination on students: "would we as teachers and Medical Educators want to take an external
examination if we had to do over again?

Kapelman (1980) refers to the use of external examinations for promotion and curricular evaluation in terms of using them moderately and not to abuse the use of them.

In response to those that expect a licensure examination to predict future performance, Kane (1982) reports:

scores on licensure examination [should] be interpreted in terms of specific abilities that are critical for effective performance in practice; abilities are considered critical to the extent that their absence would be a serious limitation in the practice of the profession.

He further suggests that the interpretation of licensure examinations as measures of critical abilities is more consistent with various other governmental activities designed to protect the public than in their interpretation as predictors of future performance. Kane (1982) further explains that the critical abilities for a profession include cognitive abilities involving knowledge and the ability to apply knowledge, as well as psychomotor skills involving the ability to do various things for clients. The abilities may be, he says, quite general such communications skills or quite specific as the ability to carry out a particular diagnostic procedure. La Duca, et al. (1984) comment in this respect:

seen in this light, the professional licensure examination should assess measurable abilities specifically relevant to practice. It follows that the underlying competence definition should be practiced referenced with the examination content reflecting this orientation by manifesting an organization that transcends academic disciplines and curricular frameworks.

Kane's report is cautious in assuming that any specific set of critical abilities would guarantee success in professional practice, but
it assumes that the absence of these abilities in professional practitioners would pose a definite risk to the clients of these practitioners. Kane's reports further suggests that specification of test content in terms of critical abilities doesn't require an exhaustive listing of the abilities required for practice, but each ability should be clearly related to client outcomes. Shimberg (1981) seems to endorse Kane's position when he states: "Neither Licensing nor Certification tests are intended to predict job success."

Testing agencies have concentrated more on relative objectivity, reliability, validity and security in number, leaving to the educational process the unwritten and imprecise obligation to instill that appropriate professional behavior-scientific, social, moral, and ethnical - which marks the physician. Egan (1976) depicts in the next lines what supposes to be the responsibilities of medical schools in relation to evaluation in spite of the existence of external examination agencies: "the faculty of every accredited medical school has the inescapable responsibility to permit the graduation of medical school only if reasonably certain of the competence of the candidate for the MD degree." In the same line of thought Kapelman (1983) states that: "Licensing exams cannot neglect nor denigrate the responsibility of the faculties in the clinical departments in medical schools.

TESTING PROGRAMS

The National Board of Medical Examiners founded in USA as an independent agency to improve the quality of medical licensure examination, has no inherent power to license individuals for medical practice, however over the years most states have come to recognize the
diploma issued by the board.

To become certified as a diplomat, candidates must graduate from a LCME - accredited medical education program and pass Part I, Part II and Part II examination. Part I and Part II are written Multiple-choice Question format and Part III a combination of MCQ with patient management-problem for the clinical area.

In spite the externality of NBME, it has always maintained close relationship with medical school faculties. For example, during 1979-1980, 73 medical schools had one or more faculty members served on the NBME's test committees (AAMC, 1981). This of course has contributed to the content validity of the examination. Faculty members not only prepare test questions but also establish the specifications for the content of the examination and select the items to fulfill the specifications. Thus, faculties have participated in setting the educational standards that must be satisfied for licensure (AAMC, 1981). Over the years the NBME has become a national standard recognized by many medical schools faculties as an additional measure to evaluate their students' educational achievements and/or the adequacy of their institutions' medical education program.

Until recently, candidates for medical licensure who were not diplomats of the NBME were required to pass an examination developed by the jurisdiction in which they sought a license. Essentially this meant that each jurisdiction had two standards for licensure, the NBME's diploma and its own examination. Collectively this meant that a physicians had more than 50 alternative ways to become licensed (AAMC, 1981). In 1968 the Federation of State Medical Boards introduced the
Federation Licensing Examination (FLEX) (FSMB, 1984).

The new FLEX has been designed so that it can provide information to the licensing boards about adequacy of the examinee's medical knowledge and understanding in order to distinguish acceptable from unacceptable examinees, usually described as a pass/fail decision (FSMB 1984). The FSMB gives an accurate description of the FLEX program. It consists of two complementary components. FLEX Components 1 is designed to evaluate measurable aspects of knowledge and understanding of basic and clinical science principles and mechanisms underlying disease and modes of therapy. The clinical aspects assessed by FLEX Component 1 place special emphasis on high-impact diseases and problems encountered in a supervised setting, on an in-patient basis, during the delivery of health care. FLEX Component 2 is designed to sample from another domain that is representative of the additional knowledge and cognitive abilities required of a physician assuming independent responsibility for the general delivery of health care to patients. The second component emphasizes a core of competences involved in the diagnosis and management of selected clinical problems most frequently encountered by the physician licensed for the independent practice of medicine (FSMB, 1984).

In Canada, the licensing programs use similar methods as NNME with common techniques and similar standards. This process generally requires: (1) graduation from an approved medical school; (2) internship in an approved program; (3) the meeting of implicit and unstated requirements for the behavioral professionalization component; (4) testimony relating to moral and ethical standards; and (5) examination
in a range of subjects related to the preparation for practice. Admission to the licensing process requires authorization from some accredited agency. In Canada, a Provincial Medical Council awards the enabling certificate whereas in USA, the NBME and the State Medical Board requires a certificate from an accredited medical schools (Neufeld, 1985).

In the United Kingdom the Licensing process is more closed tied to the educational process than in the systems of USA and Canada. Acquisition of medical degree is tantamount to a license to practice medicine, since the central licensing authority (General Medical Council) establishes national standards and imposes general conformity through accreditation of educational system. Universities grant the degree of MB, ChB or MB-BS upon completion of the prescribed course, Final Examinations held with External Examiners from other Universities set common educational standards (Muller, 1985).

It can easily be deduced that the three models presented rely heavily on the accreditation systems which in Chile it is not employed. This fact also supports, as does the earlier section of the literature review, the initiative to explore the possibility of establishing in Chile a NME. We have seen that the three models employ several ways and techniques to evaluate students before they are granted the license to practice medicine.

CONCLUSIONS OF REVIEW OF RELATED LITERATURE

Since ancient times there has been a genuine concern for evaluation in medical education and much progress has been achieved. Medical educators agree that proper evaluation is as essential from the
standpoint of the individual medical student, intern, or resident as it is for the medical school, teaching hospital, and the community in general. However, there exists the concern that evaluation as presently conducted in many schools of medicine does not satisfy efficiently the needs of the individual being evaluated due to the lack of definition of competences to be assessed, or to the fact that research in many cases about evaluation methods has had limited generalizability because it has been done on highly variable instruments, administered on a time-basis only to a relatively small group under unique circumstances for particular purposes.

Perhaps, if a NME is to be developed in Chile, it is important to make reference to the current thinking related to the issue of evaluating competences which has been summarized by Christine McGuire (1983):

Until we know more about both the structure of competence and the potential of different techniques, we must, for the purpose of triangulation, continue to use multiple methods implemented at the highest quality level to gather data that will assist us in making inescapable decisions about individual competences.

More research is required on methods for determining competencies in medical education and it is imperative to codify the findings of researchers into some generally acceptable standards for achievement testing in medical education.
ENDNOTES


Federation of State Medical Boards. February, 1984.


Hubbard, J. 1971. "Aims of Evaluation." In Gilbert, J. (Ed.) Proceedings of the Conference in Evaluation in Medical Education. Faculty of Medicine, University of Alberta, Canada.


CHAPTER THREE

METHODOLOGY

RESEARCH DESIGN AND PROCEDURES

The two concurrent methodologies employed in this investigation are a combination of descriptive quantitative and qualitative design.

First, the quantitative approach was designed to evaluate Chilean medical educators' perceptions about establishing a National Medical Examination (NME) in Chile as a licensure model. The study was also designed to evaluate: (1) the perceptions about which entity medical educators prefer for the management of a NME; (2) the areas that they believe should be included for assessment in a NME; (3) their attitude toward the evaluation methods used in medical education; (4) the potential of objective-tests; (5) their perceptions about the universities being the entity that should manage a NME if there is one; and (6) their perceptions about employing practical exams with patients in a NME. Through the evaluation of medical educators' preferences for a NME and attitudes towards the other components, the dependent variables, the approach was designed to retrospectively examine the independent variables of university affiliation and academic rank and the demographic variables of years of experience and gender. Because the independent variables are not manipulated by the investigator, the study cannot be termed a true experiment in Campbell & Stanley sense (1966). Therefore, cause and effect interpretations cannot eventuate. Rather,
regarding the independent variables, the investigator is limited to interpretation of association and relation. That is aside from the descriptive findings that will result. At most, the investigator will be able to conclude that an independent variable under study is associated with the dependent variable (Kerlinger, 1971). Second, the qualitative approach was designed to: (1) provide a better and more complete understanding of the setting under study and to identify the factors that may have influenced the development of the quantitative study; (2) answer two research questions. The first concerns the factors that may have inhibited the development of a NME in Chile. The second concerns the understanding of medical educators about licensing, accreditation and certification; (3) gain an understanding of two methodological issues. The first addresses the reluctance of physicians to respond questionnaires and the second addresses whether students should participate in the decision-making process about establishing a NME in Chile; and (4) to analyze the open-ended comments that some respondents of the questionnaires of the quantitative portion of this investigation made.

QUANTITATIVE APPROACH

Dependent Variable

The dependent variable in the study is the perceptions of Chilean medical educators about establishing in Chile a National Medical Examination as a licensure model. There are, however, six major components that constitute the dependent variable: (1) the need for a NME; (2) what entity must exist for management of a NME; (3) what areas to include for assessment in a NME; (4) Evaluation methods in medical
education; (5) objective-test and its potential; and (6) practical examination with patients in a NME.

Medical educators' perceptions about the components that constitute the dependent variable were determined by asking the respondents to complete a Likert-type questionnaire of 25 statements comprising the corresponding grouping subscales. The questionnaire also contained four demographic items.

Content validity of an instrument has been defined by Kerlinger (1971) as "the representativeness or sampling adequacy of the content - the substance, the matter, the topics - of a measuring instrument." For the present study content may be defined as those aspects which are present in the problems of Chilean medical education identified previously in Chapter One, and those aspects which are addressed in the related literature related to issues on evaluation and licensing in medical education. For example, who should manage a NME? Are medical educators satisfied with current evaluation methods that they are using? Are external examinations necessary?

**Independent Variables**

The independent variables in the study are: (1) university affiliation; (2) academic rank; (3) years of experience; and (4) gender. For the key people group the independent variable in the study is institution affiliation.

**Instrument Development and Data Collection**

For the collection of measures of Chilean medical educators' perceptions about establishing a NME as an external evaluation for
licensing medical graduates, a Likert-type questionnaire comprising of 25 statements was constructed. Four demographic-related items were also included. A sample of the questionnaire is found in Appendix A. This type of instrument was chosen following Babbie's recommendation: "it represents a more systematic and refined means for constructing indexes." The development of the instrument was based in the various aspects related to external examinations appearing in the literature and also based in the reported problems of Chilean medical education, especially those related to evaluation methods. Items were grouped into subscales in order to facilitate the measurement of the different components of the dependent variable.

In order to ensure that: (1) each statement or item of the Likert-type questionnaire whether individually or in clusters, was medically accurate and related to the main purpose of the study; (2) that the statements would validly represent medical educator's perceptions about establishing a NME; (3) that statements would validly represent medical educators' attitudes toward all the components of the dependent variable; (4) to refine the questionnaire, a four stage construction and revision process was employed. In stage one, a general input was given by the investigator's graduate committee at The Ohio State University in February, 1985 in relation to the different areas that the instrument should cover. In stage two, statements were reviewed and revised several times by a group of academics of the School of Medicine of the University of Concepcion, by members of the Chilean Association of Medical Colleges and by academics of the Office of Medical Education of the University of Concepcion. The purpose of this
revision was to examine if the content was adequate and related to medical education in Chile. In stage three, a group of medical educators were presented the questionnaires for analysis and comments. Finally, in stage four, the questionnaire was pilot tested with 30 medical educators of the School of Medicine of the University of Concepcion. Some minor modifications were made related to the clarity and precision of the statements as a result of the pilot testing. Some statements were eliminated based on an analysis of proportions. Those statements which obtained 60% or higher agreement or disagreement response were eliminated in order to work with items that provide smaller homogeneity.

Following the refinement of statements and questionnaire final construction procedure a stratified sample of N=291 medical educators (20% of the population) was drawn from the official rosters of each university participating in the study: University of Chile in Santiago, University Austral in Valdivia, University of Valparaiso in Valparaiso, University of Concepcion in Concepcion, and University de la Frontera in Temuco. All five universities provided a population of 1,445 medical educators organized in three academic ranks: Full Professors, Associate Professors, and Assistant Professors. Besides, 21 key people related to the medical education field and health area were purposively selected from the national arena.

The Likert-type questionnaire was chosen as survey technique because it was thought that a high response rate could be achieved, a large population could be sampled efficiently, and because it allows a wide range of statements and is quick to respond, considering that
physicians are always short of time. A sample of the questionnaire is found in Appendix A.

The strategy employed to assure cooperation and high response rate was a combination of mail questionnaire, face-to-face communication, telephone calls and specially effective was the participation of the Offices of Medical Education of the universities participating in the study as well as the Office of the non-participating School of Medicine. The distribution and collection in four universities was done through the offices of medical education. The distribution and collection of the questionnaires in the University of Valparaiso was done through the office of the Dean. Several follow-up letters and telephone calls to the offices of medical education were carried out. The President of the Chilean Association of Medical Colleges conducted a personal follow-up among the deans of the schools of medicine participating in the study.

Population and Sample

Medical educators of five out of six universities that have schools of medicine comprise the population for the study. The target population is 1,445 Chilean medical educators that hold professorial rank, and 21 relevant key people in the Chilean medical education field and health care service. The first group was chosen because it is believed that they possess the necessary knowledge and experience in Chilean medical education and the health arena to give accurate and representative academic opinion about evaluation and certification examinations for medical graduates and undergraduates. Relevant key people were chosen because they have the experience and expertise and
they represent entities closely related to the medical education field at its various levels. They represent scientific associations, medical associations, universities, etc., and they are permanent source for consultation in matters related to medical education and health care.

A stratified random sample of N=291 medical educators was drawn from the official rosters of each school of medicine participating in the study. Because of the large size of the sample selected (20% of the population) it was considered that possible sample error was small.

The population was organized by university affiliation and professional rank. Work experience and gender were included as demographic data. This sampling method was selected because according to Babbie (1973) "it allows to obtain a greater degree of representativeness, decreasing the probable sampling error." Table 1 shows the sampling plan in detail.
TABLE 1

SAMPLING PLAN

Population = 1,445
Sample Size N=291

<table>
<thead>
<tr>
<th></th>
<th>U. Chile</th>
<th>U. Austral</th>
<th>U. Valparaiso</th>
<th>U. Concepcion</th>
<th>U. Frontera</th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>p</em> s+ startum</em>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Prof.</td>
<td>230</td>
<td>46</td>
<td>14</td>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>Assc. Prof.</td>
<td>287</td>
<td>57</td>
<td>27</td>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>485</td>
<td>97</td>
<td>39</td>
<td>74</td>
<td>15</td>
</tr>
</tbody>
</table>

p* = population
s+ = sample
** = It was thought that a sample less than one was not to be used.

Relevant key people was purposively selected according to the judgment of the investigator about their influential role in the development of Chilean medical education and health care system. Purposive or judgemental sampling according to Delbert Miller (1967) "can be used when practical considerations preclude the use of probability sampling." Table 2 shows the sampling plan for the key people.
TABLE 2

SAMPLING PLAN FOR KEY PEOPLE

<table>
<thead>
<tr>
<th>Organization</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chilean Association of Medical Colleges</td>
<td>3</td>
</tr>
<tr>
<td>Ministry of Health</td>
<td>3</td>
</tr>
<tr>
<td>CONACEM (Board of Certification)</td>
<td>1</td>
</tr>
<tr>
<td>Ministry of Health Concepcion</td>
<td>1</td>
</tr>
<tr>
<td>Medical Association Concepcion</td>
<td>2</td>
</tr>
<tr>
<td>Concepcion Health Service</td>
<td>3</td>
</tr>
<tr>
<td>Board of Trustees U. of Chile</td>
<td>1</td>
</tr>
<tr>
<td>Chilean Academy of Medicine</td>
<td>2</td>
</tr>
<tr>
<td>Institute of Public Health</td>
<td>1</td>
</tr>
<tr>
<td>National Medical Association</td>
<td>1</td>
</tr>
<tr>
<td>Medical Education Association</td>
<td>2</td>
</tr>
<tr>
<td>Past Dean two periods at the U. Concepcion</td>
<td>1</td>
</tr>
</tbody>
</table>

N=21

Data Analysis

Factor analyses were performed several times in order to identify and confirm the grouping of items in the subscales or factors of the questionnaire. As another indication of how well the items did group together, Cronbach's alpha was calculated for each item.

Correlations of demographic data, work experience and gender were calculated. The actual type of correlations were point biserial and Spearman rank. Compilation of descriptive tables containing percentages, means and standard deviations was carried out. Two way
analyses of variance were performed to study possible differences related to rank and university in relation to the responses. Post hoc multiple comparison by Scheffe's method were computed for those instances where there was a difference.

Limitations of the Study

As stated before in Chapter One this study could have been more representative if the entire population had included all practicing physicians of Chile drawing from there the corresponding sample. However, due to the magnitude of such enterprise, plus the cost involved, this was impossible to do. Perhaps a multidisciplinary team could replicate this study at a larger scale. Another limiting factor was the exclusion of students as part of the study. This problem was anticipated by the investigator at the time of presenting the proposal for the study and was investigated in the qualitative approach of this study. The third apparent limiting factor was the refusal of the Dean of the Catholic University School of Medicine to permit his academicians to answer the questionnaire. His action produced a "blockage" of the study and the questionnaires were withheld in the deans offices of the participating schools of medicine for almost a month before allowing the instruments to be distributed among the sampled subjects.

QUALITATIVE APPROACH

Purpose of the Qualitative Approach

The main purposes of employing the qualitative approach were: (1) to provide a better and more complete understanding of the setting under study and to identify the factors that may have influenced the
development of the survey study; (2) to answer two research questions. The first concerns to the factors that may have inhibited the development of a NME in Chile. The second concerns the understanding of medical educators about licensing, accreditation and certification of specialties; (3) to explain two methodological issues. The first addresses the reluctancy of physicians to respond questionnaires and the second addresses whether students should participate in the decision-making process about establishing NME in Chile; and (4) to analyze the open-ended comments that some respondents of the questionnaires of the quantity study made. An additional question was added later to the interview schedule in order to find out some explanations for the Catholic University refusal attitude.

This qualitative approach has been described by Patton (1980): "as a naturalistic inquiry effort whose purpose is to understand naturally occurring phenomena in their naturally occurring states." Guba & Lincoln (1981) understand naturalistic inquiry as "a discovery oriented process." Williams and Raush (1969) define naturalistic inquiry as "the investigation of phenomena within and in relation to their naturally occurring context."

Greene (1985) states that there are signals of greater acceptance of the naturalistic perspective - or at least of qualitative methods - within the evaluation community. Mathew Miles and Michael Huberman (1984) state that: "qualitative data are attractive, they are a source of well-grounded, rich descriptions and explanations of process occurring in local contexts."

Miles (1984) summarizes the advantages of qualitative data when he
writes "words, specially when they are organized into incidents or stories, have a concrete, vivid, meaningful flavor that often proves far more convincing to a reader - another researcher a policy-maker, a practitioner - than pages or numbers."

**Methodology**

The methodology employed in this qualitative approach follows to some extent the patterns of the technique identified by Patton, Miles, Greene as triangulation. Berdnaiz (1983) cited by Greene (1985) explains the concept of triangulation as "the multiple employment of sources data, observers, methods or theories in investigation or the same phenomenon." Greene further states that the goal of triangulation methods is to strengthen the validity of the overall findings through congruence and/or complementary of the results from each method. Congruence means, according to Greene, similarities, consistency or convergence of results, whereas complementary refers to one set of results enriching, expanding upon clarifying, or illustrating the other.

The triangulation type approach in this study permitted the comparison of the responses obtained in the survey and the open-ended comments with the response which it generated patterns obtained via interviews both comparison made in relation to the research questions posed for this qualitative approach.

Sampling was done purposively on the basis of the research questions. Miles (1984) states in this respect that choices for sampling are driven in large part by the conceptual framework and research questions. In this study the investigator in his judgment determined that a selected group of 35 experienced and knowledgeable
medical educators, Deans, Directors of Offices of Medical Education, President of Chilean Association of Medical Colleges, Chilean Medical Association and Faculty was an appropriate sample and representative in quality.

A standardized open-ended interview was employed to obtain responses for the two research questions and explanation for the two methodological issues. Appendix B contains the interview schedule employed. The investigator chose the standardized open-ended interview because it was in the interest of the study to ask all interviewees the same basic questions, in order to increase comparability of responses (Patton, 1980), and to facilitate organization and analysis of the data, risking flexibility as Patton puts it. Interviewing was thought to be further effective following Patton's idea: "the fundamental principle of qualitative interviewing is to provide a framework within which respondents can express their own understanding in their own terms."

An informal conversation interview (Patton, 1980) was conducted during the design of the entire study in order to obtain feedback data for the design of the study. This technique according to Patton (1980) is particularly useful when the evaluator can stay in the situation for some period of time. To accomplish this the investigator traveled to the location of the various schools of medicine participating in the study and conferred with medical educators who contributed the revision of content, and strategies for distribution and collection of the survey instruments.

Observation was another technique employed, especially for the description of the setting. Patton (1980) states that the purpose of
observational analysis is to take the reader into the setting that was observed. He further states that "the basic criterion to apply to a recorded observation is whether or not that observation permits the reader to enter the situation."

The final technique employed in this approach was the content analysis of the open-ended comments made by the respondents of the survey instruments on the questionnaires in addition to the responses they gave to them.

The Setting

The study at a national level was conducted in five out of six Chilean universities that have schools of medicine in 1985. The University of Chile, the largest in the country is located in the capitol, Santiago and the other four are located in provinces. The University of Valparaiso, the closest in distance to Santiago, is about 75 miles Northwest, and the other three are located more than 300 miles south of Santiago.

The Catholic University of Chile refused to participate in the study arguing that the study was endangering the university autonomy. This action initiated a national turmoil and as a result the Chilean Association of Medical Colleges withheld the survey instruments for almost a month. Finally, the survey instruments were able to reach destiny by November, 1985. However, the acting dean of the Catholic University School of Medicine did not allow the distribution of the questionnaires. The investigator was informed that a letter explaining this unusual situation was going to be sent to the investigator of the study. Such document never arrived. The investigator in order to find
helpful information to interpret this attitude of the dean of the School of Medicine of the Catholic University added one off-the-record question to the interview schedule. The findings are detailed further in the study.

In spite of the difficulties in conducting this study, it can be said that it has produced a deep reflection among the different institutions related to medical education and health care system - ministry of health, universities, and medical associations. Letters of support and congratulations for the effort were received from the various areas of the medical education and health area.

Prior to this turmoil within the medical education field social unrest and student agitation were increasing in intensity which one may have thought was going to affect the return rate of the questionnaires. Communication was difficult among the schools of medicine and the development of the entire study was slow and uncertain.

Data Analysis

Patton (1980) states that "there are no formal, universal rules to following analyzing, interpreting and evaluating qualitative data." However, Miles (1984) suggests three concurrent flows of activity when analyzing qualitative data: (1) data reduction; (2) data display; and (3) conclusion drawing/verification. This qualitative approach followed Miles' suggestions. Data reduction and data display are found in Table 27 through 33 in Chapter Four where patterns were identified in organized assemble of information to facilitate conclusion drawing and action taking in relation to the two research questions posed for this qualitative approach: (1) what are the factors that may have inhibited
the development of a NME in Chile? the opinions about (2) licensing; (3) certification; and (4) accreditation? The responses to the two methodological issues and the responses to the refusal attitude of the Catholic University are also found in Chapter Four.


CHAPTER FOUR

RESULTS

QUANTITATIVE APPROACH

The following analyses were carried out in this investigation: (1) factor analyses; (2) Cronbach alpha; (3) point biserial correlation and Spearman rank correlations; (4) compilation of descriptive tables containing percentages, means and standard deviations; (5) two-way analysis of variance; and (6) post hoc multiple comparisons.

Examination of the Return Rate

A stratified random sample of size N=291, and a purposive sample of 21 key people were selected. Questionnaires comprising 25 Likert-type statements and 4 demographic items were sent and distributed to the subjects within the different universities and participating institutions in the study. It can be observed that of 291 questionnaires sent, 217 were returned. The highest return rate occurred among full professors and associate professors, 95.16% and 95.12% respectively. The return rate among assistant professors was much lower, 54.42%. Particularly at the University of Chile followed by the University La Frontera. Table 3 shows the return rate information.
### TABLE 3

**RETURN RATE INFORMATION**

Population = 1,445  
Sample Size = 291  
Return - 217 (91%)

<table>
<thead>
<tr>
<th>Population</th>
<th>20% sample</th>
<th>return</th>
<th>return rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Chile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>stratum</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Prof.</td>
<td>230</td>
<td>46</td>
<td>43</td>
</tr>
<tr>
<td>Assoc. Prof.</td>
<td>287</td>
<td>57</td>
<td>53</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>485</td>
<td>97</td>
<td>37</td>
</tr>
<tr>
<td>University Austral</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Prof.</td>
<td>14</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Assoc. Prof.</td>
<td>27</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>39</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>University of Valparaiso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Prof.</td>
<td>28</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Assoc. Prof.</td>
<td>36</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>74</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>University of Concepcion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Prof.</td>
<td>26</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Assoc. Prof.</td>
<td>51</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>95</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>University La Frontera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Prof.</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Assoc. Prof.</td>
<td>15</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>41</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Rank</td>
<td>Population</td>
<td>20% Sample</td>
<td>Return</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>------------</td>
<td>--------</td>
</tr>
<tr>
<td>Full Prof.</td>
<td>302</td>
<td>62</td>
<td>59</td>
</tr>
<tr>
<td>Assoc. Prof.</td>
<td>416</td>
<td>82</td>
<td>78</td>
</tr>
<tr>
<td>Asst. Prof.</td>
<td>734</td>
<td>147</td>
<td>80</td>
</tr>
</tbody>
</table>

The total return rate of 91% was unexpectedly high due to the general conditions under which the study at a national level was conducted. The key people response rate was 90%. Nineteen out of 21 responded the questionnaires.

Factor analyses were performed several times in order to identify and confirm the grouping of statements or items into the subscales of the questionnaire. Table 4 summarizes the results of the factor analysis that was deemed to be the most interpretable.
TABLE 4
FACTOR ANALYSIS SUMMARY TABLE RESULTS

<table>
<thead>
<tr>
<th>Item</th>
<th>Loading</th>
<th>Item</th>
<th>Loading</th>
<th>Item</th>
<th>Loading</th>
<th>Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>.80</td>
<td>18</td>
<td>.78</td>
<td>16</td>
<td>.81</td>
<td>20</td>
<td>.80</td>
</tr>
<tr>
<td>5</td>
<td>.79</td>
<td>22</td>
<td>.71</td>
<td>17</td>
<td>.78</td>
<td>23</td>
<td>.74</td>
</tr>
<tr>
<td>15</td>
<td>.73</td>
<td>8</td>
<td>.70</td>
<td>25</td>
<td>.58</td>
<td>19</td>
<td>.38</td>
</tr>
<tr>
<td>1</td>
<td>.73</td>
<td>11</td>
<td>.62</td>
<td>7</td>
<td>.38</td>
<td>24</td>
<td>.36*</td>
</tr>
<tr>
<td>9</td>
<td>.72</td>
<td>6</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>.69</td>
<td>12</td>
<td>.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.57</td>
<td>2</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Factor 5
Practice/School Management

<table>
<thead>
<tr>
<th>Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>.62+</td>
</tr>
<tr>
<td>4</td>
<td>.59+</td>
</tr>
</tbody>
</table>

Factor 6
Objective-test

<table>
<thead>
<tr>
<th>Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>.75</td>
</tr>
</tbody>
</table>

*Although this loading is not large and in fact this item loaded slightly higher on another factor, the decision to include it in this factor was made on the basis of the investigator knowledge of content.

+Although the analysis indicated that these two items could be grouped together the investigator deemed them to be sufficiently different to analyzed separately.
In Appendix C the items corresponding to each factor are found grouped. Factor 1, corresponds to the dimension named by the investigator Need for a NME. Factor 2, corresponds to Entity, the institution that the respondents believe should manage a NME. Item 4 also is related to the management of a NME, though being more precise about the entity. Factor 3, grouped the items corresponding to Areas that respondents believe should be assessed by a NME. Factor 4 groups those items or statements related to respondents opinion with respect to the current evaluation methods used in medical education. Factor 5, comprises two items which the analysis put together. However, according to the investigator knowledge of the content they are not related. Item 4 was mentioned above and item 3 is related to whether the respondents favor a NME that includes a practical examination with patients. Finally, factor 6 composed of only item 14 is related to whether the respondents believe that objective-tests measure mainly knowledge.

As another indication of how well the items of the questionnaire did group together into the four major factors, the Cronbach's Alpha was calculated for each of them. Cronbach's for the 10 item factor 1 need was .896; for the 4 item factor 2 entity was .691; for the 4 item factor 3 areas was .667; and for the 4 item factor Methods was .611.

Correlations of demographic data of experience and gender were calculated. Point biserial correlations of sex with individual item responses indicated conclusively that sex was not related to responses. Similarly, Spearman rank correlations of work experience with item responses indicated that experience was also not related to responses. Therefore, these two demographic variables were not included in further
analyses on a subgroup basis. However, the data for which university they came from and rank was considered to be potentially different across sub-groups, therefore two-way analyses of variance were performed to further study possible differences related to rank and university. These analyses showed no differences due to rank. Therefore, in the interest of simplicity, the descriptive data to be presented has been pooled for all ranks and shown only by the five university groups.

Descriptive Response Data and Tables

For the purpose of economy of space, the names of the participating universities appearing in the tables are abbreviated as follows:

- University of Chile = U/Chi
- University Austral = U/Aus
- University of Valparaiso = U/Val
- University of Concepcion = U/Con
- University La Frontera = U/Fron

Perceptions of the respondents about establishing in Chile a National Medical Examination were determined by items 1, 2, 5, 6, 9, 10, 13, 15, 21 that comprise factor need. Table 5 shows percentages of agreement and disagreement on the need for establishing in Chile a NME. It is observed that 64% of the respondents agree and 11% disagree. The distribution by university is also found in Table 5.
### TABLE 5

PERCENTAGES OF AGREEMENT AND DISAGREEMENT BY UNIVERSITY RESPECT TO THE NEED FOR A NME

<table>
<thead>
<tr>
<th>Item</th>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+A</td>
<td>D+</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>1</td>
<td>95</td>
<td>3</td>
<td>75</td>
<td>6</td>
<td>50</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>22</td>
<td>38</td>
<td>31</td>
<td>38</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>79</td>
<td>8</td>
<td>69</td>
<td>25</td>
<td>38</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>71</td>
<td>17</td>
<td>63</td>
<td>25</td>
<td>25</td>
<td>63</td>
</tr>
<tr>
<td>9</td>
<td>52</td>
<td>34</td>
<td>38</td>
<td>44</td>
<td>17</td>
<td>67</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>5</td>
<td>94</td>
<td>6</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>12</td>
<td>81</td>
<td>8</td>
<td>81</td>
<td>13</td>
<td>38</td>
<td>33</td>
</tr>
<tr>
<td>13</td>
<td>61</td>
<td>26</td>
<td>69</td>
<td>13</td>
<td>33</td>
<td>42</td>
</tr>
<tr>
<td>15</td>
<td>59</td>
<td>26</td>
<td>56</td>
<td>25</td>
<td>25</td>
<td>63</td>
</tr>
<tr>
<td>21</td>
<td>56</td>
<td>24</td>
<td>56</td>
<td>25</td>
<td>25</td>
<td>63</td>
</tr>
<tr>
<td>AV</td>
<td>64</td>
<td>18</td>
<td>59</td>
<td>24</td>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

+\(A\) = Represents the sum of the categories strongly agree plus agree.
+\(D\) = Represents the sum of the categories strongly disagree plus disagree.

Means and standard deviations by university of the items that constitute the need factor is shown in Table 6.
TABLE 6
MEANS AND STANDARD DEVIATIONS BY UNIVERSITY OF ITEMS THAT CONSTITUTE THE NEED FACTOR

<table>
<thead>
<tr>
<th>Item</th>
<th>U/Chi X</th>
<th>SD</th>
<th>U/Aus X</th>
<th>SD</th>
<th>U/Val X</th>
<th>SD</th>
<th>U/Con X</th>
<th>SD</th>
<th>U/Fron X</th>
<th>SD</th>
<th>Total X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.6</td>
<td>0.72</td>
<td>4.0</td>
<td>0.89</td>
<td>3.3</td>
<td>1.37</td>
<td>4.5</td>
<td>0.89</td>
<td>4.6</td>
<td>0.96</td>
<td>4.2</td>
</tr>
<tr>
<td>2</td>
<td>3.4</td>
<td>1.06</td>
<td>3.1</td>
<td>0.95</td>
<td>2.9</td>
<td>1.31</td>
<td>3.6</td>
<td>1.07</td>
<td>3.0</td>
<td>0.81</td>
<td>3.3</td>
</tr>
<tr>
<td>5</td>
<td>4.1</td>
<td>0.99</td>
<td>3.5</td>
<td>0.96</td>
<td>3.0</td>
<td>1.35</td>
<td>4.1</td>
<td>1.11</td>
<td>3.9</td>
<td>0.87</td>
<td>3.7</td>
</tr>
<tr>
<td>6</td>
<td>3.9</td>
<td>1.09</td>
<td>3.5</td>
<td>1.03</td>
<td>2.5</td>
<td>1.17</td>
<td>3.8</td>
<td>1.38</td>
<td>3.8</td>
<td>1.03</td>
<td>3.5</td>
</tr>
<tr>
<td>9</td>
<td>3.3</td>
<td>1.20</td>
<td>2.9</td>
<td>1.02</td>
<td>2.4</td>
<td>1.01</td>
<td>3.2</td>
<td>1.25</td>
<td>3.1</td>
<td>0.87</td>
<td>3.0</td>
</tr>
<tr>
<td>10</td>
<td>4.2</td>
<td>0.84</td>
<td>4.1</td>
<td>0.68</td>
<td>3.0</td>
<td>1.35</td>
<td>4.3</td>
<td>0.89</td>
<td>4.0</td>
<td>0.66</td>
<td>3.9</td>
</tr>
<tr>
<td>12</td>
<td>4.1</td>
<td>0.90</td>
<td>3.9</td>
<td>0.92</td>
<td>3.1</td>
<td>1.05</td>
<td>4.1</td>
<td>0.89</td>
<td>4.0</td>
<td>0.47</td>
<td>3.9</td>
</tr>
<tr>
<td>13</td>
<td>3.6</td>
<td>1.26</td>
<td>3.7</td>
<td>1.07</td>
<td>3.0</td>
<td>1.25</td>
<td>3.4</td>
<td>1.32</td>
<td>3.0</td>
<td>1.24</td>
<td>3.3</td>
</tr>
<tr>
<td>15</td>
<td>3.6</td>
<td>1.22</td>
<td>3.4</td>
<td>1.03</td>
<td>2.6</td>
<td>1.13</td>
<td>3.4</td>
<td>1.35</td>
<td>2.9</td>
<td>0.87</td>
<td>3.2</td>
</tr>
<tr>
<td>21</td>
<td>3.4</td>
<td>1.13</td>
<td>2.6</td>
<td>0.71</td>
<td>2.5</td>
<td>1.02</td>
<td>3.4</td>
<td>1.32</td>
<td>2.9</td>
<td>1.10</td>
<td>2.9</td>
</tr>
<tr>
<td>AV</td>
<td>3.8</td>
<td>3.5</td>
<td>2.8</td>
<td>3.8</td>
<td>3.5</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.5</td>
</tr>
</tbody>
</table>

The key people overall opinion toward establishing in Chile a NME was in general in favor, with the exception of the member of the Board of Trustees of the University of Chile and a Past Dean for two periods at the University of Concepcion, College of Medicine, which were slightly opposed. Table 7 shows the detailed information about this group of selected representatives of the medical education in Chile.
TABLE 7
MEANS OF KEY PEOPLE CORRESPONDING TO THE ITEMS THAT CONSTITUTE THE NEED FACTOR

<table>
<thead>
<tr>
<th>A** B C D E F G H I J K L</th>
</tr>
</thead>
<tbody>
<tr>
<td>N 2 3 1 1 2 3 1 1 1 1 2 1</td>
</tr>
</tbody>
</table>

Items | X
--- | ---
1 | 3.0 4.7 4.0 5.0 5.0 4.3 3.0 4.0 2.0 4.0 4.5 5.0
2 | 4.0 4.0 4.0 3.0 3.0 4.0 4.0 4.0 5.0 4.0 4.0 2.0
5 | 3.0 3.0 4.0 2.0 5.0 3.3 4.0 4.0 4.0 2.0 4.5 2.0
9 | 2.5 3.0 4.0 4.0 1.5 2.3 1.0 4.0 2.0 2.0 3.5 2.0
10 | 4.0 3.3 4.0 4.0 4.0 3.6 4.0 4.0 5.0 4.0 4.0 2.0
12 | 4.0 3.7 4.0 4.0 3.5 3.7 2.0 5.0 4.0 4.0 4.0 4.0
13 | 3.0 2.3 3.0 3.0 5.0 3.7 5.0 5.0 5.0 2.0 4.0 4.0
15 | 3.0 2.3 4.0 2.0 5.0 3.0 2.0 5.0 5.0 2.0 4.0 2.0
21 | 3.0 3.7 3.0 4.0 5.0 3.0 2.0 4.0 2.0 4.0 4.5 2.0
AV | 3.2 3.4 3.7 3.4 4.2 3.4 2.9 4.4 3.6 3.2 4.2 2.9

** Key
A = Chilean Association of Medical Colleges - 3
B = Ministry of Health - 3
C = CONACEM (Board of Certification) - 1
D = Ministry of Health Concepcion - 1
E = Medical Association Concepcion - 2
F = Concepcion Health Service - 3
G = Board of Trustees U. of Chile - 1
H = Chilean Academy of Medicine = 2
I = Institute of Public Health - 1
J = National Medical Association - 1
K = Medical Education Association - 2
L = Past Dean two periods at the U. Concepcion - 1

The perceptions of medical educators about which entity should manage a NME were determined by items 8, 11, 18, 22 that comprise factor entity. Table 8 shows percentages of agreement and disagreement of responses on the entity factor and the distribution by university is also found in Table 8. It is observed that 59.3% of the responses disagree on governmental interference or interference of any other
external agency in evaluating medical graduates' competences, only 12% disagree with respect the university being exclusively the organization to deal with evaluation of competences. For item 22 the scoring scale was reversed in order to reflect the wording. Agreement on item 22 meant disagreement on items 8, 11 and 18.

TABLE 8
PERCENTAGES OF AGREEMENT AND DISAGREEMENT BY UNIVERSITY RESPECT TO THE ENTITY FOR THE MANAGEMENT OF A NME

<table>
<thead>
<tr>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>8 9 77 13 81 4 92 18 77 10 50</td>
<td>10 78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 50 35 31 44 33 58 29 53 50 30</td>
<td>44 41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 24 55 13 81 8 75 29 59 40 30</td>
<td>34 39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22+ 11 81 6 75 8 83 12 68 30 40</td>
<td>12 77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AV 24 62 16 70 13 77 22 64 33 38</td>
<td>25 59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+For this item the scoring scale was reversed in order to reflect the wording. Agreement on item 22 means disagreement on items 8, 11 and 18.

Means and standard deviations by university of items that constitute entity factor is found in Table 9.
<table>
<thead>
<tr>
<th>Item</th>
<th>U/Chi X</th>
<th>SD</th>
<th>U/Aus X</th>
<th>SD</th>
<th>U/Val X</th>
<th>SD</th>
<th>U/Con X</th>
<th>SD</th>
<th>U/Fron X</th>
<th>SD</th>
<th>Total X</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.9</td>
<td>1.05</td>
<td>2.1</td>
<td>0.92</td>
<td>1.6</td>
<td>0.92</td>
<td>2.2</td>
<td>1.29</td>
<td>2.5</td>
<td>0.84</td>
<td>2.1</td>
</tr>
<tr>
<td>11</td>
<td>3.2</td>
<td>1.44</td>
<td>2.8</td>
<td>1.32</td>
<td>2.5</td>
<td>1.35</td>
<td>2.7</td>
<td>1.48</td>
<td>3.1</td>
<td>1.10</td>
<td>2.9</td>
</tr>
<tr>
<td>18</td>
<td>2.6</td>
<td>1.27</td>
<td>2.1</td>
<td>1.08</td>
<td>2.0</td>
<td>0.90</td>
<td>2.6</td>
<td>1.50</td>
<td>3.0</td>
<td>1.05</td>
<td>2.5</td>
</tr>
<tr>
<td>22</td>
<td>1.8</td>
<td>1.13</td>
<td>1.9</td>
<td>0.95</td>
<td>1.8</td>
<td>1.17</td>
<td>2.1</td>
<td>1.20</td>
<td>2.7</td>
<td>1.15</td>
<td>2.0</td>
</tr>
<tr>
<td>AV</td>
<td>2.39</td>
<td>2.22</td>
<td>1.98</td>
<td>1.98</td>
<td>2.38</td>
<td>2.38</td>
<td>2.83</td>
<td>2.83</td>
<td>2.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The key people overall opinion toward the type of organization or entity that should manage a NME was in general against governmental interference in evaluating medical graduate's competences. However, a mean of 3.0 was obtained for the item indicating that the responsibility of evaluating competences should be left exclusively in the hands of the university. Table 10 shows the means for this key people.
TABLE 10
MEANS OF KEY PEOPLE CORRESPONDING TO ITEMS THAT CONSTITUTE THE ENTITY FACTOR

<table>
<thead>
<tr>
<th>A** B C D E F G H I J K L</th>
</tr>
</thead>
<tbody>
<tr>
<td>N  2 3 1 1 2 3 1 1 1 1 2 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2.0 2.0 2.0 2.0 5.0 2.0 2.0 3.0 4.0 2.0 3.0 2.0</td>
</tr>
<tr>
<td>11</td>
<td>2.0 2.5 4.0 4.0 3.5 2.5 5.0 4.0 2.0 4.0 4.5 2.0</td>
</tr>
<tr>
<td>22</td>
<td>3.0 4.0 2.0 2.0 3.5 4.0 2.0 4.0 4.0 2.5 1.0</td>
</tr>
<tr>
<td>AV</td>
<td>3.2 2.8 2.5 2.7 4.0 2.6 2.7 3.5 3.0 3.0 3.0 1.8</td>
</tr>
</tbody>
</table>

** Key
A = Chilean Association of Medical Colleges - 3
B = Ministry of Health - 3
C = CONACEM (Board of Certification) - 1
D = Ministry of Health Concepcion - 1
E = Medical Association Concepcion - 2
F = Concepcion Health Service - 3
G = Board of Trustees U. of Chile - 1
H = Chilean Academy of Medicine = 2
I = Institute of Public Health - 1
J = National Medical Association - 1
K = Medical Education Association - 2
L = Past Dean two periods at the U. Concepcion - 1

The perceptions of Chilean medical educators about the areas of competence that a NMR should intend to assess were determined by items 7, 16, 17, 25 which comprise the areas factor. Table 11 shows percentages of agreement and disagreement of responses on the area factor. It can be seen that a 79% agreement on the areas stated on the questionnaire and only 10% disagreement was found.
TABLE 11

PERCENTAGES OF AGREEMENT AND DISAGREEMENT BY UNIVERSITY RESPECT TO THE ITEMS CONSTITUTE AREAS FACTOR

<table>
<thead>
<tr>
<th>Item</th>
<th>A</th>
<th>D</th>
<th>A</th>
<th>D</th>
<th>A</th>
<th>D</th>
<th>A</th>
<th>D</th>
<th>A</th>
<th>D</th>
<th>A</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>66</td>
<td>14</td>
<td>56</td>
<td>13</td>
<td>33</td>
<td>25</td>
<td>68</td>
<td>18</td>
<td>80</td>
<td>20</td>
<td>63</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td>99</td>
<td>2</td>
<td>88</td>
<td>6</td>
<td>92</td>
<td>4</td>
<td>94</td>
<td>3</td>
<td>100</td>
<td>0</td>
<td>96</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>89</td>
<td>3</td>
<td>69</td>
<td>6</td>
<td>75</td>
<td>17</td>
<td>88</td>
<td>6</td>
<td>80</td>
<td>10</td>
<td>85</td>
<td>6</td>
</tr>
<tr>
<td>25</td>
<td>75</td>
<td>10</td>
<td>69</td>
<td>19</td>
<td>58</td>
<td>29</td>
<td>74</td>
<td>24</td>
<td>80</td>
<td>10</td>
<td>73</td>
<td>15</td>
</tr>
<tr>
<td>AV</td>
<td>83</td>
<td>8</td>
<td>70</td>
<td>11</td>
<td>64</td>
<td>19</td>
<td>81</td>
<td>12</td>
<td>85</td>
<td>10</td>
<td>79</td>
<td>10</td>
</tr>
</tbody>
</table>

Means and standard deviations per item/university that constitute areas factor is shown in Table 12.

TABLE 12

MEANS AND STANDARD DEVIATIONS PER ITEM BY UNIVERSITY THAT CONSTITUTE AREAS FACTOR

<table>
<thead>
<tr>
<th>Item</th>
<th>X</th>
<th>SD</th>
<th>X</th>
<th>SD</th>
<th>X</th>
<th>SD</th>
<th>X</th>
<th>SD</th>
<th>X</th>
<th>SD</th>
<th>X</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>3.8</td>
<td>1.14</td>
<td>3.7</td>
<td>1.01</td>
<td>3.1</td>
<td>1.05</td>
<td>3.7</td>
<td>1.10</td>
<td>3.9</td>
<td>1.10</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4.6</td>
<td>0.57</td>
<td>4.5</td>
<td>0.89</td>
<td>4.5</td>
<td>0.93</td>
<td>4.5</td>
<td>0.82</td>
<td>4.3</td>
<td>0.48</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4.3</td>
<td>0.78</td>
<td>3.9</td>
<td>0.88</td>
<td>4.0</td>
<td>1.19</td>
<td>4.2</td>
<td>0.92</td>
<td>3.8</td>
<td>0.78</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>3.9</td>
<td>0.95</td>
<td>3.6</td>
<td>0.95</td>
<td>3.3</td>
<td>0.96</td>
<td>3.6</td>
<td>1.15</td>
<td>3.8</td>
<td>0.78</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>AV</td>
<td>4.16</td>
<td>3.92</td>
<td>3.70</td>
<td>4.02</td>
<td>3.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The key people's overall opinion about the areas of competence that a NME should intent to assess is in favor of the areas mentioned.
Table 13 shows the detailed information.

**TABLE 13**

**MEANS STANDARD DEVIATIONS OF KEY PEOPLE CORRESPONDING TO THE ITEMS THAT CONSTITUTE THE AREAS FACTOR**

<table>
<thead>
<tr>
<th>A** B C D E F G H I J K L</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N 2 3 1 1 2 3 1 1 1 1 2 1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>4.5 3.6 2.0 2.0 5.0 3.0 2.0 4.0 4.0 2.0 3.5 2.0 3.1</td>
</tr>
<tr>
<td>16</td>
<td>4.0 4.0 4.0 4.0 5.0 4.3 5.0 4.0 4.0 3.0 5.0 4.5</td>
</tr>
<tr>
<td>17</td>
<td>4.0 4.3 4.0 4.0 5.0 4.0 5.0 4.0 4.0 4.5 4.0 4.3</td>
</tr>
<tr>
<td>25</td>
<td>3.5 2.7 4.0 5.0 2.5 4.0 4.0 4.0 4.0 3.5 4.0 3.8</td>
</tr>
<tr>
<td>AV</td>
<td>4.1 3.6 3.5 3.5 5.0 3.5 4.0 4.5 4.0 4.0 3.5 3.7 3.9</td>
</tr>
</tbody>
</table>

**Key**

A = Chilean Association of Medical Colleges - 3
B = Ministry of Health - 3
C = CONACEM (Board of Certification) - 1
D = Ministry of Health Concepcion - 1
E = Medical Association Concepcion - 2
F = Concepcion Health Service - 3
G = Board of Trustees U. of Chile - 1
H = Chilean Academy of Medicine = 2
I = Institute of Public Health - 1
J = National Medical Association - 1
K = Medical Education Association - 2
L = Past Dean two periods at the U. Concepcion - 1

The attitudes of medical educators toward the adequacy of current evaluation methods employed in Chilean medical education to assess students' competence was determined by items 19, 29, 23 and 24 which comprise the methods factor. Table 14 shows percentages of agreement and disagreement on the method factor. It can be observed a 51% agreement on that the adequacy of evaluation methods is poor and a 28% disagreement.
### TABLE 14

PERCENTAGES OF AGREEMENT AND DISAGREEMENT BY UNIVERSITY RESPECT TO THE ADEQUACY OF CURRENT EVALUATION METHODS

<table>
<thead>
<tr>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Response Category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>19</td>
<td>57</td>
<td>26</td>
<td>56</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>20</td>
<td>45</td>
<td>30</td>
<td>57</td>
<td>31</td>
<td>25</td>
</tr>
<tr>
<td>23</td>
<td>43</td>
<td>36</td>
<td>44</td>
<td>31</td>
<td>33</td>
</tr>
<tr>
<td>24</td>
<td>73</td>
<td>1</td>
<td>56</td>
<td>6</td>
<td>50</td>
</tr>
<tr>
<td>AV</td>
<td>55</td>
<td>26</td>
<td>53</td>
<td>25</td>
<td>35</td>
</tr>
</tbody>
</table>

Means and standard deviations per item by university is shown in Table 15.

### TABLE 15

MEANS AND STANDARD DEVIATIONS OF ITEMS BY UNIVERSITY THAT CONSTITUTE METHOD FACTOR

<table>
<thead>
<tr>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Item</strong></td>
<td>X</td>
<td>SD</td>
<td>X</td>
<td>SD</td>
<td>X</td>
</tr>
<tr>
<td>19</td>
<td>3.5</td>
<td>1.13</td>
<td>3.3</td>
<td>1.01</td>
<td>2.8</td>
</tr>
<tr>
<td>20</td>
<td>4.6</td>
<td>0.57</td>
<td>4.5</td>
<td>0.89</td>
<td>4.5</td>
</tr>
<tr>
<td>23</td>
<td>3.1</td>
<td>1.12</td>
<td>3.1</td>
<td>0.88</td>
<td>2.9</td>
</tr>
<tr>
<td>24</td>
<td>3.8</td>
<td>0.91</td>
<td>3.6</td>
<td>0.72</td>
<td>3.2</td>
</tr>
<tr>
<td>AV</td>
<td>3.4</td>
<td>3.3</td>
<td>3.0</td>
<td>3.4</td>
<td>3.3</td>
</tr>
</tbody>
</table>

The key people overall attitude toward the adequacy of current evaluation methods employed by Chilean medical educators in general
favor the fact that these methods are not efficiently assessing the necessary students' competences with the exception of representatives of Chilean Association of Medical Colleges, Board of Trustees of University of Chile and Health Service. Table 16 shows the means for these key people.

** TABLE 16 **

MEANS OF KEY PEOPLE CORRESPONDING TO ITEMS THAT CONSTITUTE THE METHOD FACTOR

| A** | B   | C   | D   | E   | F   | G   | H   | I   | J   | K   | L   |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| N   | 2   | 3   | 1   | 1   | 2   | 3   | 1   | 1   | 1   | 2   | 1   |
| Items | X   |     |     |     |     |     |     |     |     |     |     |     |
| 19   | 3.0 | 3.0 | 5.0 | 2.0 | 4.5 | 2.6 | 1.0 | 2.0 | 4.0 | 4.0 | 2.0 | 2.0 |
| 20   | 2.5 | 3.0 | 3.0 | 4.0 | 5.0 | 3.3 | 1.0 | 3.0 | 4.0 | 4.0 | 4.0 | 4.0 |
| 23   | 3.0 | 3.0 | 4.0 | 3.0 | 5.0 | 2.6 | 2.0 | 3.0 | 5.0 | 4.0 | 2.0 | 2.0 |
| 24   | 3.0 | 3.6 | 4.0 | 5.0 | 5.0 | 2.3 | 4.0 | 4.0 | 2.0 | 4.0 | 4.5 | 4.0 |
| AV   | 2.8 | 3.2 | 4.0 | 3.5 | 4.0 | 2.7 | 2.5 | 3.0 | 3.7 | 4.0 | 3.1 | 3.0 |

** Key **
A = Chilean Association of Medical Colleges - 3
B = Ministry of Health - 3
C = CONACEM (Board of Certification) - 1
D = Ministry of Health Concepcion - 1
E = Medical Association Concepcion - 2
F = Concepcion Health Service - 3
G = Board of Trustees U. of Chile - 1
H = Chilean Academy of Medicine = 2
I = Institute of Public Health - 1
J = National Medical Association - 1
K = Medical Education Association - 2
L = Past Dean two periods at the U. Concepcion - 1

The attitude of medical educators with respect to include a practical examination with patients in a NME was determined by item 3. Table 17 shows percentages of agreement and disagreement of responses.
It is observed that only a 44% of agreement was obtained and a 39% disagreement.

### TABLE 17

PERCENTAGES OF AGREEMENT AND DISAGREEMENT BY UNIVERSITY RESPECT TO THE INCLUSION OF PRACTICAL EXAMINATION IN A NME

<table>
<thead>
<tr>
<th></th>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>Item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>58</td>
<td>63</td>
<td>25</td>
<td>33</td>
<td>58</td>
</tr>
</tbody>
</table>

Means and standard deviations by university of item 3 that constitute practical examination factor is shown in Table 18.

### TABLE 18

MEANS AND STANDARD DEVIATIONS BY UNIVERSITY OF ITEM 3 THAT CONSTITUTE PRACTICAL EXAMINATION FACTOR

<table>
<thead>
<tr>
<th></th>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>X</td>
<td>SD</td>
<td>X</td>
<td>SD</td>
<td>X</td>
<td>SD</td>
</tr>
<tr>
<td>Item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.1</td>
<td>1.22</td>
<td>3.6</td>
<td>1.31</td>
<td>2.6</td>
<td>1.31</td>
</tr>
</tbody>
</table>

The key people opinion toward including a practical examination with patients in a NME was in general little over neutral with a mean of 3.36. Table 19 shows the detailed information.
TABLE 19

MEANS OF KEY PEOPLE CORRESPONDING TO ITEM 3 THAT CONSTITUTE THE PRACTICAL EXAMINATION FACTOR

<table>
<thead>
<tr>
<th></th>
<th>A**</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Pract

** Key
A = Chilean Association of Medical Colleges - 3
B = Ministry of Health - 3
C = CONACEM (Board of Certification) - 1
D = Ministry of Health Concepcion - 1
E = Medical Association Concepcion - 2
F = Concepcion Health Service - 3
G = Board of Trustees U. of Chile - 1
H = Chilean Academy of Medicine - 2
I = Institute of Public Health - 1
J = National Medical Association - 1
K = Medical Education Association - 2
L = Past Dean two periods at the U. Concepcion - 1

The perceptions of Chilean medical educators about the University becoming the organization responsible for the design and administration of a NME were determined by item 4. Table 20 shows percentages of agreement and disagreement on item 4. It is observed 83% in favor and 6% against.
TABLE 20

PERCENTAGES OF AGREEMENT AND DISAGREEMENT RESPECT TO THE UNIVERSITY BECOMING THE ORGANIZATION RESPONSIBLE OF THE DESIGN AND ADMINISTRATION OF A NME

<table>
<thead>
<tr>
<th>Response Category</th>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item A D A D A D A D A D A D</td>
<td>4 83 8 81 0 75 13 94 0 70 0 83 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Means and standard deviations of item 4 that constitute school-management factor is shown in Table 21.

TABLE 21

MEANS AND STANDARD DEVIATIONS OF ITEM 4 THAT CONSTITUTE THE SCHOOL-MANAGEMENT FACTOR

<table>
<thead>
<tr>
<th>Item</th>
<th>X SD</th>
<th>X SD</th>
<th>X SD</th>
<th>X SD</th>
<th>X SD</th>
<th>X SD</th>
<th>X SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4.2</td>
<td>0.94</td>
<td>4.4</td>
<td>0.80</td>
<td>4.0</td>
<td>0.99</td>
<td>4.4</td>
</tr>
</tbody>
</table>

The key people’s attitude toward the University becoming the organization responsible for the design and administration of a NME is in favor with a mean of 4.0. Table 22 shows the detailed information.
TABLE 22
MEANS OF THE KEY PEOPLE CORRESPONDING TO ITEM 4 THAT
CONSTITUTE THE SCHOOL-MANAGEMENT FACTOR

<table>
<thead>
<tr>
<th></th>
<th>A**</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4.5</td>
</tr>
<tr>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>5.0</td>
<td>4.5</td>
</tr>
<tr>
<td>4.0</td>
<td>3.6</td>
</tr>
<tr>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>5.0</td>
<td>2.0</td>
</tr>
<tr>
<td>4.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

**Key**
- A = Chilean Association of Medical Colleges - 3
- B = Ministry of Health - 3
- C = CONACEM (Board of Certification) - 1
- D = Ministry of Health Concepcion - 1
- E = Medical Association Concepcion - 2
- F = Concepcion Health Service - 3
- G = Board of Trustees U. of Chile - 1
- H = Chilean Academy of Medicine - 2
- I = Institute of Public Health - 1
- J = National Medical Association - 1
- K = Medical Education Association - 2
- L = Past Dean two periods at the U. Concepcion - 1

The perceptions of medical educators in relation to whether objective-tests measure mainly cognitive knowledge were determined by item 14 which comprises objective-test factor. Table 23 shows percentages of agreement and disagreement on item 14. It is observed that 87.5% agrees on that objective tests measures mainly cognitive abilities. Only a 5% disagrees.
## TABLE 23

**PERCENTAGES OF AGREEMENT AND DISAGREEMENT BY UNIVERSITY RESPECT TO THE USEFULNESS OF OBJECTIVE TESTS**

<table>
<thead>
<tr>
<th></th>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response Category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>14</td>
<td>90</td>
<td>2</td>
<td>88</td>
<td>6</td>
<td>79</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>6</td>
<td>70</td>
<td>10</td>
<td>87</td>
<td>5</td>
</tr>
</tbody>
</table>

Means and standard deviations by university of item 14 that constitutes objective-tests factor is shown in Table 24.

## TABLE 24

**MEANS AND STANDARD DEVIATIONS BY UNIVERSITY OF ITEM 14 THAT CONSTITUTE OBJECTIVE-TEST FACTOR**

<table>
<thead>
<tr>
<th></th>
<th>U/Chi</th>
<th>U/Aus</th>
<th>U/Val</th>
<th>U/Con</th>
<th>U/Fron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>X</td>
<td>SD</td>
<td>X</td>
<td>SD</td>
<td>X</td>
<td>SD</td>
</tr>
<tr>
<td>14</td>
<td>4.2</td>
<td>0.68</td>
<td>3.9</td>
<td>0.68</td>
<td>3.8</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>4.1</td>
<td>0.88</td>
<td>3.6</td>
<td>0.69</td>
<td>3.2</td>
<td></td>
</tr>
</tbody>
</table>

The key people's opinion about agrees on that objective tests measures mainly capacity to recall information. A mean of 4.09 was obtained. Table 25 shows the detailed information.
TABLE 25
MEANS OF KEY PEOPLE CORRESPONDING TO ITEM 14 THAT CONSTITUTE OBJECTIVE-TEST FACTOR

<table>
<thead>
<tr>
<th></th>
<th>A**</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Items</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>4.0 4.0 4.0 4.0 3.5 3.6 5.0 4.0 5.0 4.0 4.0 4.0</td>
</tr>
<tr>
<td>Pract</td>
<td>X</td>
</tr>
</tbody>
</table>

** Key
A = Chilean Association of Medical Colleges - 3
B = Ministry of Health - 3
C = CONACEM (Board of Certification) - 1
D = Ministry of Health Concepcion - 1
E = Medical Association Concepcion - 2
F = Concepcion Health Service - 3
G = Board of Trustees U. of Chile - 1
H = Chilean Academy of Medicine - 2
I = Institute of Public Health - 1
J = National Medical Association - 1
K = Medical Education Association - 2
L = Past Dean two periods at the U. Concepcion - 1

The two-way analysis of variance for the total score on the need factor showed a significant (<.0001) effect for university. The effect for rank as well as the interaction effect were not significant. The analysis of variance summary table for these results is found in Appendix D.

The post hoc multiple comparison by Scheffe's method, showed that the means for the faculty at the University of Chile, and University of Concepcion are both significantly higher than the means for the University of Valparaiso. None of other comparisons are significant.

For the contrast between the University of Chile and University of
Valparaiso, an F value of 8.57 was obtained exceeding the critical value of 4.80 (p<.001); and for the contrast between the University of Concepcion and the University of Valparaiso an F value of 5.72 was obtained exceeding the critical value of 4.80 (p<.001).

The two-way analysis of variance for the total score on the entity factor showed no significant effect for university. The effect for rank as well as the interaction effect were not significant. The analysis of variance summary table for these results is found in Appendix E.

The two-way analysis of variance for the total score on the areas factor showed a significant (p<.04) effect for university. The effect for rank as well the interaction effect were not significant. The analysis of variance summary table is found in Appendix F.

The post hoc multiple comparison by Scheffe's method, showed that the means for the faculty at the University of Chile are significantly higher than the means for the faculty of the University of Valparaiso. None of other comparisons are significant.

For the contrast between the University of Chile and University of Valparaiso, an F value of 2.54 was obtained exceeding the critical value of 2.41 (p<.05).

The two-way analysis of variance for the total score on the methods factor showed no significant effect for university. The effect for rank as well as the interaction effect were not significant. The analysis of variance summary table for these results is found in Appendix G.

The two-way analysis of variance for the total score on both practice and school-management items showed no significant effects for university. The effects for rank as well as the interaction effects
were not significant. The analysis of variance summary table for these results is found in Appendix H.

The two-way analysis of variance for the total score on objective-test factor showed a significant (p<.003) effect for university. The effect for rank as well as the interaction effect were not significant. The analysis of variance summary table for these results is found in Appendix I.

The post hoc multiple comparison by Scheffe's method, showed that none of the pairwise comparisons of university means were significantly different.

Table 26 shows summary for the effect of university for each factor/item.

**TABLE 26**

**SUMMARY FOR THE EFFECT OF UNIVERSITY FOR EACH FACTOR/ITEM**

<table>
<thead>
<tr>
<th>Factor/Item</th>
<th>F Ratio</th>
<th>p-Value</th>
<th>Post Hoc Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need</td>
<td>7.20</td>
<td>.001</td>
<td>Chile, Concepcion &gt; Valparaiso</td>
</tr>
<tr>
<td>Entity</td>
<td>1.91</td>
<td>.110</td>
<td>No differences</td>
</tr>
<tr>
<td>Areas</td>
<td>2.57</td>
<td>.039</td>
<td>Chile &gt; Valparaiso</td>
</tr>
<tr>
<td>Methods</td>
<td>1.38</td>
<td>.240</td>
<td>No differences</td>
</tr>
<tr>
<td>Practical Exam</td>
<td>1.71</td>
<td>.150</td>
<td>No differences</td>
</tr>
<tr>
<td>School-management</td>
<td>1.25</td>
<td>.290</td>
<td>No differences</td>
</tr>
<tr>
<td>Objective-test</td>
<td>4.17</td>
<td>.003</td>
<td>No pairwise differences</td>
</tr>
</tbody>
</table>

Discussion of the results presented in this chapter are found in Chapter Five of this investigation.
QUALITATIVE APPROACH

Interviews were conducted with 30 out of 35 sampled Chilean medical educators. The interview technique's aims were: (1) to provide a better and more complete understanding of the setting under study and to identify the factors that may have influenced the development of the survey study; (2) to answer two research questions. The first one related to the factors that may have inhibited the development in Chile of a NME. The second one related to the understanding of medical educators about licensing, accreditation and certification of specialties; (3) to answer two methodological issues. The first one related to the reluctancy of physicians to respond questionnaires and the second one related whether students should participate in the decision-making process about establishing in Chile a NME; (4) to analyze the open-ended comments that some respondents of the questionnaires of the survey study made; and (5) to find explanation about the Catholic University refusal to participate in the study. Tables 27 through 34 contain the detailed information.
### TABLE 27
FACTORS IDENTIFIED AS POSSIBLE INHIBITORS OF THE DEVELOPMENT OF A NME IN CHILE

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Lack of leadership</td>
</tr>
<tr>
<td>8</td>
<td>Universities don't want to assume responsibility</td>
</tr>
<tr>
<td>6</td>
<td>High cost of operation and lack of infrastructure</td>
</tr>
<tr>
<td>5</td>
<td>Fear to be evaluated comparatively by means of a NME</td>
</tr>
<tr>
<td>4</td>
<td>Fear to loose university autonomy</td>
</tr>
<tr>
<td>3</td>
<td>Lack of interest on the part of the universities</td>
</tr>
<tr>
<td>2</td>
<td>Lack of curricular homogeneity</td>
</tr>
<tr>
<td>2</td>
<td>The objectives for a NME have never been stated clearly</td>
</tr>
<tr>
<td>Frequency of citation by interviewees</td>
<td>Opinions</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>9</td>
<td>A NME could be a Licensing mechanism</td>
</tr>
<tr>
<td>7</td>
<td>Licensing is determined by the political context and is controlled by the government. It is a legislative issue</td>
</tr>
<tr>
<td>3</td>
<td>Licensing is a good mechanism to raise standards in medical education</td>
</tr>
<tr>
<td>2</td>
<td>The concept needs to be studied in depth. It is not advisable now due to political factors</td>
</tr>
<tr>
<td>2</td>
<td>Licensing should be exploratory and volunteer</td>
</tr>
<tr>
<td>1</td>
<td>Licensing should be analyzed and considered by the new Health Law to be approved and put into practice in 1986</td>
</tr>
<tr>
<td>1</td>
<td>Licensing in Chile is granted with the degree and there is no need to worry</td>
</tr>
<tr>
<td>2</td>
<td>Gave no opinion</td>
</tr>
</tbody>
</table>
TABLE 29

OPINIONS ABOUT CERTIFICATION OF SPECIALTIES

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>Important for defining specialties, assessing patient care service, prevent malpractice due to complexity of medicine</td>
</tr>
<tr>
<td>3</td>
<td>Excellent mechanism to assure competency in specialties</td>
</tr>
<tr>
<td>2</td>
<td>Must be compulsory</td>
</tr>
<tr>
<td>1</td>
<td>Important that exists an exam to certify</td>
</tr>
<tr>
<td>1</td>
<td>Suggests a recertification procedure is needed</td>
</tr>
<tr>
<td>1</td>
<td>The certification mechanism in Chile needs to be evaluated</td>
</tr>
<tr>
<td>1</td>
<td>Must be done through residency</td>
</tr>
</tbody>
</table>
TABLE 30
OPINIONS ABOUT PRE-MED ACCREDITATION

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Important because helps to set requisite and standards. Also helps to improve</td>
</tr>
<tr>
<td>5</td>
<td>Aware that it doesn't exist in Chile and that it only exists in postgraduate education</td>
</tr>
<tr>
<td>4</td>
<td>Convenient, but not indispensable to maintain adequate level of medical education</td>
</tr>
<tr>
<td>2</td>
<td>Before 1981 U. Chile controlled medical education in Chile and it performed a kind of accreditation</td>
</tr>
<tr>
<td>1</td>
<td>Gave no opinion</td>
</tr>
</tbody>
</table>

During the development of the interview the investigator was able to pose an unwritten question related to the interviewees' attitude toward establishing a NME in Chile. Nearly all showed markedly preference for the development of a NME. However, they made clear their concern for the type of entity which would manage an NME and about the competences that such an exam ought to measure.

Data Analysis for the Methodological Issues

Two main methodological issues were of concern of the investigator: (1) should students participate in the decision-making process about establishing in Chile a NME? and (2) do physicians have a positive attitude toward responding survey instruments? Tables 31 and 32 contain a summary of the opinions.
### TABLE 31
OPINIONS ABOUT STUDENTS’ PARTICIPATION IN DECISION-MAKING PROCESS ABOUT ESTABLISHING IN CHILE A NME

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Students should not participate in the decision-making process about establishing in Chile a NME</td>
</tr>
<tr>
<td>4</td>
<td>Perhaps graduates should be consulted</td>
</tr>
<tr>
<td>4</td>
<td>Students should express their opinions but should not participate</td>
</tr>
<tr>
<td>2</td>
<td>It is not the right time to consult students about this issue. They are highly politicized</td>
</tr>
</tbody>
</table>

The reason to include this methodological issue within the interview schedule grew out from the need to confirm the decision of the investigator not to include students among the population for the survey research. The investigator stated at the moment of presenting the proposal for the survey that consulting students was not advisable due to the general situation within the universities.
TABLE 32
OPINIONS ABOUT PHYSICIANS RELUCTANCY TO RESPOND QUESTIONNAIRES

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>There is a tendency not to respond questionnaires or any kind of written document</td>
</tr>
<tr>
<td>6</td>
<td>Respond depending on the purpose of the investigation, on the motivation posed by the topic and on the usefulness of the study</td>
</tr>
<tr>
<td>6</td>
<td>Physicians do not like to make public their opinions and attitudes. They are reluctant to state an opinion outside their profession field</td>
</tr>
<tr>
<td>2</td>
<td>It implies a personal definition of what one believes. This results in rejection of the questionnaire</td>
</tr>
<tr>
<td>1</td>
<td>General public like to talk not to write</td>
</tr>
</tbody>
</table>

The investigator after spending 14 years in the medical field strongly believed that physicians were reluctant to respond questionnaires because they are always short of time. This concern has been addressed by Sudman (1985) when he states:

For many evaluations studies responses will be needed from professionals such as physicians, social workers, lawyers, business executives, teachers or college professors. Because these professionals are often too busy and difficult to locate for a personal interview, mail procedures are especially efficient methods for reaching them.

He further comments: "among many such groups, one sees detailed verbal objections to participating, which sometimes break out into formal group refusals to cooperate." Refusals by professionals are not a new phenomenon. In 1956, Deutscher, in describing physicians relations to a
mailed questionnaire called it a study in "Resistentialism." In an interesting but unfortunately forgotten article, he described some reasons for refusal based on comments on questionnaires that were returned. These include the following:

- Resistance to stereotyping and generalizing
- Resistance to restrictions imposed by multiple-choice questions
- Resistance to questions that "don't make sense."

Sudman (1985) reports that there are four main reasons why professionals are noncooperative. First, is that the respondent is very busy and the time spent filling out the survey could be better spent on other, more important (or lucrative) professional tasks. The second reason, which is highly related to the first, is that the value of the survey is not clear to the respondent, or that it is clear but the value is low. The third reason for noncooperation is concern about confidentiality of the results. The final reason is that individual questions in the questionnaire appear to the respondent to be biased or not to allow the respondent the choice of a full range of answers.

Professor Juan Manso an experienced sociologist of the University of Concepcion in relation to this issue stated: "in answering surveys there is a psychological reaction to defend themselves [physicians] because there is a strong cultural force given by tradition of their institution [medical profession]." He further adds that: "in Chile social science research did not develop more rapidly due to the rejection attitude of the general public toward being asked how they think or how they are doing." He finishes his comments by stating that: "an answered questionnaire is like a mirror where we can see ourselves just the way we are or think."
Content Analysis of Open-ended Comments

Respondents of the survey instruments of the quantitative approach of the study were given the opportunity to express additional comments on the reverse part of the questionnaire. Nearly 35 respondents made some remarks which are summarized on Table 33.

TABLE 33
OPEN-ENDED COMMENTS MADE BY RESPONDENTS OF THE QUESTIONNAIRE

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Psychomotor skills should be evaluated during the medical education process in a formative way, and at the entrance to the medical education process</td>
</tr>
<tr>
<td>6</td>
<td>A NME is not the appropriate method to improve evaluation techniques</td>
</tr>
<tr>
<td>5</td>
<td>Simulated patients and written simulation should be used instead of real patients in evaluation</td>
</tr>
<tr>
<td>5</td>
<td>The public doesn't have the means to find out who has approved a competence examination. The public is incapable to understand evaluation of physicians competence</td>
</tr>
<tr>
<td>5</td>
<td>A NME should be designed and managed by a joint commission composed of: universities, medical associations and ministry of health</td>
</tr>
<tr>
<td>4</td>
<td>Practical examination with patients is useful to assess practitioner-patient relationship</td>
</tr>
<tr>
<td>4</td>
<td>Attitudes should be developed and assessed during the process of medical education. Attitudes are difficult to evaluate</td>
</tr>
</tbody>
</table>

An additional question was added to the interview schedule as a
result of the Catholic University refusal to participate in the study. The investigator considered of great importance to find out a possible explanation. Results of such a question are found in Table 34.

**TABLE 34**
OPINIONS ABOUT THE REFUSAL ATTITUDE OF THE CATHOLIC UNIVERSITY TO PARTICIPATE IN THE STUDY

<table>
<thead>
<tr>
<th>Frequency of citation by interviewees</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Fear of being evaluated comparatively</td>
</tr>
<tr>
<td>3</td>
<td>Difficult to understand such unbelievably antiscientific attitude</td>
</tr>
<tr>
<td>3</td>
<td>Suspicions about a NME</td>
</tr>
<tr>
<td>2</td>
<td>Fear of lowering their educational standard specially because they favor tertiary medicine</td>
</tr>
<tr>
<td>2</td>
<td>Fear of losing autonomy</td>
</tr>
</tbody>
</table>
CHAPTER FIVE
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

INTRODUCTION

The two concurrent methodologies employed in this investigation were a combination of descriptive and qualitative design. First, the quantitative study was designed to evaluate: (1) Chilean medical educators' perceptions about establishing a NME in Chile as a licensure model; (2) the perceptions about which entity Chilean medical educators would prefer for the management of a NME; (3) the areas that they believe should be included for assessment in a NME; and (4) their attitudes toward the evaluation methods used currently in Chilean medical education. The second methodology, the qualitative study, was designed: (1) to provide a better and more complete understanding of the setting under study and to identify the factors that may have influenced the development of the entire investigation; (2) to answer two research questions. The first regarding the factors that may have inhibited the development of a NME in Chile, and the second regarding the understanding of Chilean medical educators about licensing, accreditation and certification; (3) to find explanation to two methodological issues. The first addressed the reluctance of physicians to respond to questionnaires and the second addressed whether the students should participate in the decision-making process about establishing a NME in Chile; and (4) to analyze the open-ended comments.
that some respondents in the quantitative approach made.

The major findings of this investigation will be discussed in relation to the research questions posed for the quantitative and qualitative approach.

**QUANTITATIVE APPROACH**

The questions posed for the quantitative approach are: first, do Chilean medical educators perceive a need for a national medical examination? Second, what type of entity do they believe should manage a NME if there is one? Third, what areas of competence should a NME measure? Fourth, are they satisfied with the current medical students' evaluation methods?

The response rate for this quantitative investigation, 91%, is considered to be excellent by Babbie (1973) for survey research. For those failing to respond there was no record. The reasons given by the offices of medical education for the 9% of those failing to respond were: that some of the sampled subjects were no longer members of their faculty, others were away from the city, some have died and that a few did not have time to respond to the questionnaire. It is interesting to note the low return rate of assistant professors of the University of Chile, 38.1%. At least two possible explanations can be made for this finding which was provided by academics of the offices of medical education. First, the relatively low return rate is due, in part, to the fact that assistant professors work in different institutions and they have to move a great deal, therefore, it is difficult to locate them in order to deliver and collect the questionnaires. Second, it is believed that assistant professors are not motivated to participate in
this type of study, especially when dealing with attitudes and opinions, because they are not entitled to participate in their own universities in the decision-making processes.

The factor analyses of the instrument confirmed and identified the grouping of statements or subscales: need, entity, areas, methods, and three independent items, 3, 4 and 14. Table 35 shows a summary of means for each factor.

<table>
<thead>
<tr>
<th>Factor</th>
<th>U. Chile</th>
<th>U. Austral</th>
<th>U. Valparaiso</th>
<th>U. Concepcion</th>
<th>U. Frontera</th>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Need</td>
<td>3.81</td>
<td>3.47</td>
<td>2.83</td>
<td>3.78</td>
<td>3.52</td>
<td>3.51</td>
</tr>
<tr>
<td>2. Entity</td>
<td>2.39</td>
<td>2.22</td>
<td>1.98</td>
<td>2.38</td>
<td>2.83</td>
<td>2.80</td>
</tr>
<tr>
<td>3. Areas</td>
<td>4.16</td>
<td>3.92</td>
<td>3.70</td>
<td>4.02</td>
<td>3.95</td>
<td>4.23</td>
</tr>
<tr>
<td>4. Methods</td>
<td>3.38</td>
<td>3.33</td>
<td>2.96</td>
<td>3.36</td>
<td>3.30</td>
<td>3.29</td>
</tr>
<tr>
<td>Practice</td>
<td>3.09</td>
<td>3.62</td>
<td>2.62</td>
<td>2.94</td>
<td>2.50</td>
<td>3.36</td>
</tr>
<tr>
<td>School</td>
<td>4.19</td>
<td>4.37</td>
<td>3.95</td>
<td>4.44</td>
<td>4.00</td>
<td>4.05</td>
</tr>
<tr>
<td>Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective-</td>
<td>4.21</td>
<td>3.93</td>
<td>3.75</td>
<td>4.05</td>
<td>3.60</td>
<td>3.75</td>
</tr>
<tr>
<td>test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion of Research Question 1

Do Chilean medical educators perceive a need for establishing in Chile a NME? It was anticipated that this study would reveal medical educators' opinions toward the establishment of a NME, and that the schools from provinces would be more in favor of such an exam than the schools from Santiago. By studying Table 35 it can be concluded that with the exception of the University of Valparaiso with a mean of 2.83, the other universities favored the establishment of a NME. The means and percentages of the other four universities although in favor, did not reveal a strong body of agreement for a NME. According to the investigator's judgment this may be due to the considerable controversy of the entire evaluation issue among the medical education community in Chile. Key people appeared to be markedly in favor of a NME. This is in part due to the fact that they are in a better position to judge more objectively the process of medical education since they look at the process from another angle. The qualitative analysis revealed that the interviewees are very much in favor of a NME. Of the individual items for this factor it is interesting to note that item 1 obtained the highest score, and this is possibly due to the fact that this item was the one that measured most directly the need for an NME. There was less than 50% agreement in four universities on item 5 which relates to the public involvement in judging medical practice. This may be explained from a sociological perspective since it is known that physicians do not accept the public's judgments about their performance. This statement has been supported by the analysis of the open-ended comments which revealed that some physicians think that the public doesn't understand
evaluation of medical practice. This perception seems to agree with the report of Segal and Burnet (1980) which states, "since patients are typically laymen confronting a professional this evaluation becomes problematic." The report further maintains that "a true profession, according to the generally acknowledged characteristics of professionalism, is free from lay evaluation and control."

Item 9 also obtained less than 50% agreement and with a total mean of 3.00. This statement was related to the measuring of psychomotor skills in a NME. These findings reveal that in spite of the overall agreement in the need for a NME they are not quite sure about including psychomotor skills in the exam. The analysis of the open-ended comments revealed that the respondents agree that psychomotor abilities should be evaluated formatively during the process of medical education and at the end of the process. Key people show higher means in item 1 and 10. University of Valparaiso means were consistently lower than the other means in this need factor.

The score of item 21 is lower than the total score for the factor need. This item is related to whether a NME would protect patient's health. The findings may be suggesting that the respondents do not link a development of an examination with the patient's health protection.

Interviewees (90%) appeared to be in favor of a NME. However, they showed concern about the type of test management entity, the areas to be assessed and the type of examination.

It can be concluded that among the medical educators there is a positive attitude toward establishing a NME in Chile. But, the analysis of the items individually revealed that they would like a more precise
clarification of the purposes of a NME.

The two-way analysis of variance for the total score on the need factor showed a significant (<.0001) effect for university. The post hoc multiple comparison by Scheffe's method showed that the means for the faculty at the University of Chile and University of Concepcion were both significantly higher than the means for the University of Valparaiso. A possible explanation is that the issue of a NME had been discussed before responding to the questionnaires, and therefore a kind of general consensus could have resulted. It is interesting to note that the School of Medicine of the University of Valparaiso was the only one in which the instruments were distributed and collected by the Dean's office. In addition, the dean of this same school questioned the procedure of random sampling.

Discussion of Research Question #2

What type of entity should manage a NME if there is one? The findings related to this question revealed that there is a strong body of opinions against an independent extrauniversity entity dealing with evaluation of medical graduates competencies (78% disagreement). Findings for item 11 show a divided attitude toward the Government being the entity to license to practice medicine after graduates demonstrate competence (44% agreement and 41% disagreement). The findings for item 18 also revealed a split decision whether a specialized agency of the government should evaluate professional competence of medical graduates. The findings for item 22 reveal 77% in favor of the universities exclusively being the entity to evaluate professional competence of the medical graduates. The findings for item 4, although
not grouped in factor 2 by the factor analysis, reveal a strong body of opinion showing preference for the universities being the entity to design and administer a NME (83%). A possible explanation to this is that items comprising the factor entity were not as precise as item 4 in measuring and specifying the type of entity. It is evident that the respondents do not want the government to be involved in managing a NME. Forty-four percent (44%) of the respondents indicated agreement with government involvement, with 34% in agreement on giving the responsibility to a specialized agency controlled by the government. On the other hand, 77% agreed that the universities have exclusive responsibility to evaluate medical graduates' competence. One explanation for this is that it is possible that medical educators believe that other institutions could participate in the evaluation as well. However, item 4 stated that the universities would be the entity for the design and management of a NME and consequently a higher percentage of agreement was obtained, 83% with a mean of 3.39. Key people in this same item scored a mean of 4.02. It was expected also that key people would be cautious about entity since they represent organizations which may be part of a NME program or committee.

An explanation for these disparate findings for factor entity is better understood by looking at the historical development of events related to entities and medical education. Until 1981, the University of Chile was in charge of granting the MD degrees for the entire country. The schools of medicine struggled long for autonomy which came in 1981. The Chilean Medical Association used to have control over peer physicians. In a way this control functioned as a kind of professional
self-regulation. Newly graduated physicians had to register in the medical association before engaging in the practice of the profession. This was abolished by the Ministry of Justice in 1981.

The two-way analysis of variance for the total score on the entity factor showed no significant effect for university. One possible explanation for this is that the universities are proud of their autonomy and do not wish interference from the government, and faculty believe that evaluation is their business. Another possible explanation for rejecting the involvement of the government or any other external agency is that in future times these institutions may not make good use of their prerogatives.

The analysis of the open-ended comments made evident the fear concerning test security in the sense that some people might obtain the questions before hand and pass the examination without being qualified for it.

Discussion of Research Question #3

What areas should a NME measure? Factor 3 grouped four statements that indicated the areas that medical educators thought a NME should include. The areas posed by the questionnaire are by no means complete, however, it suggests areas of competence compiled from the review of medical education literature. There is a 79% agreement for the areas mentioned in the study. The findings for the key people reveal a mean of 3.92 which tells us that they are aware of the importance of determining the competencies of medical students to be evaluated. It is important to point out that item 3 is closely related to factor 3, which is, however, the analysis grouped it separately. An agreement of 44%
was showed that nearly half the respondents indicated that a practical exam with patients should not be included in a NME. These findings seem to support decisions of The National Board of Medical Examiners (Part III) as well as other medical specialties to discontinue this modality of examination (Muzzin, 1985).

The high percentages obtained in this factor 3 area clearly indicate the importance that medical educators give to the determination of competencies and clinical abilities as reported in Chapter Two on determining competencies. This finding is important because if a NME is in the mind of planners and authorities it will be feasible to undertake studies in this direction.

The findings for item 7 related to the assessment of attitudes by a NME report that a mean of 3.63 was obtained for all universities which is considered good since the attitude development assessment and in medical education has been always a controversial issue. Sixty-three percent (63%) of the respondents indicated agreement to include attitude assessment in a NME. The mean for the same item reported for the key people was 3.13, this seems to confirm the value feelings of attitude assessment. The findings from the analysis of the open-ended comments reveal also a serious concern for the assessment and development of attitudes. Several comments suggest that attitudes should be developed and measured during the process of medical education due to its importance. These findings are consistent with McGuire's study (1983) on techniques of assessing attitudes: "Educators and public alike have been concerned about the assessment and prediction of attitudes in the recruitment, selection, training, and licensure of health professionals"
The two-way analysis of variance for the total score on the areas factor showed a significant ($p < 0.4$) effect for university. Post hoc multiple comparison by Scheffe's method showed that the means for the faculty of the University of Chile are significantly higher than the means for the faculty of the University of Valparaiso. Although there is a statistically significant difference, it can be concluded that has a relative small practical importance, and that it can be attributed to the consistency of the University of Valparaiso in having low means which can be considered as a overall attitude. It is interesting to note that in this occasion the mean for the University of Valparaiso unusually exceeded the value of 3.00. A 64% of agreement for the factor was obtained by this same university. For the key people a mean of 3.92 was obtained expressing this way support for these possible areas to be assessed in a NME.

**Discussion of Research Question #4: Methods**

The items comprising this factor determined to some extent medical educators' attitudes toward the evaluation methods being employed in the medical education process. The content of this factor was obtained from the seminars on medical education held in Chile and from the study of Repeto and Salazar (1984) which reported that faculty of the schools of medicine weren't familiar with evaluation techniques. The findings for this factor named methods, indicate low evidence about medical educators being quite aware of the deficiencies of evaluation methods. The findings are encouraging in sense because medical educators do recognize that something is wrong with the evaluation procedures they are
currently using. Nearly 51% seem to have agreed with Repeto & Salazar's findings. Interestingly again the University of Valparaiso showed the lowest percentage of agreement and a mean of only 2.62.

Findings for item 19 showed that the respondents agree on that current evaluation methods employed in medical schools need some kind of evaluation supplement to assure a better assessment of students' competencies. The findings for item 20 reveal that current methods of evaluations are not intended to evaluate the capacity of students to take responsibility in the patients' care.

Findings for item 23 agree with the belief that current evaluation methods assess mainly capacity for recall. This fact shows that faculty are aware that other and better methods for evaluating competencies are needed. Key people mean of 3.32 shows little evidence that they agree with the statements that evaluation methods are not sufficient to evaluate medical competencies. It can be concluded that further studies are needed in this area in order to determine the strategies to follow.

The analysis of item 24 showed a surprisingly 70% agreement in circumstances that the issue of malpractice in Chile is a relatively new concept and not often is linked to medical education process. These findings support the general principle of having efficient evaluation methods and the process of medical education can be improved, consequently the output and medical care.

**Discussion of Item 14**

This item, though closely related to factor labeled methods, was by factor analysis grouped separately. The findings for this item indicate that there is a strong body of opinion among Chilean medical educators
(87%) agreeing that objective-tests measure chiefly the cognitive domain. This suggests that they are aware that other methods are needed. These findings agree also with Repeto & Salazar's study and with the report of the national seminar on medical education held in 1984. These findings seem to agree well also with Phillipson's report (1971) that evaluation based solely on knowledge of facts and memorization of bits of information, and not on application of knowledge to the patient, may in fact be detrimental rather than beneficial. The variable key persons reported a mean of 4.42 showing a high level of agreement among Chilean medical educators that answered the questionnaires. These findings are important since most external examinations are usually designed in a multiple-choice format or objective-tests, therefore it has implications for future developments of such kind of exams, especially if a NME is developed.

Summary of Findings

The findings of this study seem to indicate that there is a positive attitude toward establishing in Chile a NME. However, there is strong evidence that there exists serious concern about the entity that would design and manage a possible NME. In part an explanation for such concern lies in the politics of jurisdiction and control. There is evidence of the concern for the areas to be assessed and this evidence is well supported by the various studies reported earlier in Chapter Two related to the determination of competencies. Also there is substantial evidence that the issue is of interest and that medical educators are willing to participate and debate. The findings also show interest for improving the evaluation methods employed in medical education.
However, the findings do not agree with the anticipated assumption that schools of medicine from provinces would be more in favor for a NME than the ones from Santiago. University of Valparaiso and University of Santiago are two such examples. The first one is opposed to the NME and the second one is in favor.

As a result of this investigation, one cannot make categorical statements regarding medical educators preference for the establishment of a NME in Chile. Indeed, to do so would surely result in increasing controversy and conflict among the scholars. Rather, any statements which are made regarding the development of a NME in Chile have to keep in mind that Chilean medical educators are open to discuss and debate the issue. Chilean medical educators do not reject the idea of NME but they are not so strongly in favor to adopt a NME without careful analysis.

QUALITATIVE APPROACH

Conclusion Drawing and Verification

Miles (1984) explains nicely this activity: "From the beginning of the study and data collection the investigator decides what things mean, notes regularities, patterns, explanations, possible configurations, causal flows and proportions."

The findings of this qualitative approach are analyzed from the perspective of the research questions posed for this approach, from the methodological issues, the Catholic University School of Medicine refusal to authorize medical educators to respond to the questionnaire and from the perspective of its relation to the overall findings of the quantitative approach.
The findings related to the factors that may have inhibited the development of a NME in Chile show that to some extent the issue has been present in the mind of medical educators all this time and that if only an adequate leadership would have developed the idea something would have resulted. It is interesting to note that the lack of interest on the part of the universities was mentioned a few times. However, the findings of the quantitative approach revealed that the faculty of the universities participating in the study were rather in favor of a NME. Perhaps the problem lies within the authority level. It is interesting to note that the fear of losing autonomy was mentioned four times as a possible inhibiting factor. This same reason was given by some faculty in relation to the refusal of Catholic University to participate in the study. These findings tell us that any intent to develop a NME should take into consideration this fact and therefore make sure that autonomy is not in danger.

In relation to the interviewee's opinion about licensing, certification of specialties and accreditation of pre M.D. education, the possibility that a NME be considered a licensing mechanism was mentioned nine times and three times was mentioned the fact that a licensing examination would raise the standards of medical education. These findings support the findings obtained for the quantitative approach in that medical educators if not strongly favor a NME are not against one. However, other opinions are to be taken into consideration if dealing with the issue in the future. Medical educators interviewed seem to understand the concept of licensing which is not unusual since most physicians in Chile keep abreast with USA medical education and
In relation to the concept of certification of specialties, it was mentioned 20 times that it is an important mechanism for defining specialties, assessing patient care services, and preventing of malpractice. These findings agree very well with the importance given to this subject by the schools of medicine in Chile. In April, 1985, the first Seminar of Graduate Medical Education in Medical Specialties was held in the School of Medicine of the University of Concepcion, and the topics under discussion were related to curriculum, accreditation in graduate medical education, certification of specialties, research and evaluation in graduate medical education (Ramirez, 1985).

In relation to accreditation in pre M.D. medical education, it was mentioned 15 times that this mechanism is important to help set requisites and standards, and that it was also a mechanism for improvement of the medical education process. These findings are important since the model of Chilean medical education is the no-accreditation, no-licensing examination, based on completion of course programs. Other models of licensing were explained in Chapter One in the conceptual framework section. It is interesting to note that some interviewees mentioned that when University of Chile used to have control of granting the MD degrees, this was a kind of accreditation that has been lost. It was explained in the section corresponding to the identification of the problem that in 1981 the universities obtained autonomy to grant MD degrees. This fact was mentioned a problem factor that may suggest the development of a NME.

Two main methodological issues were of concern of the investigator:
(1) should students participate in the decision-making process about establishing in Chile a NME? and (2) do physicians have a negative attitude toward responding questionnaires. In relation to the first issue, interviewees seem to support the investigator in his assumption that it was not appropriate to consult students about the issue because the conditions did not advise to do it yet. It was stated on 27 occasions by the interviewees that students should not participate in the decision-making process about establishing in Chile a NME. However, some agree that maybe graduate students could give their insights to the problem. Others agree that it is possible to ask students' opinions but they do not participate in the decision-making process itself. These findings are important because if a NME is going to be developed, it is necessary to find the right time to seek students' cooperation.

In reference to whether physicians are reluctant to respond questionnaire the findings appear to be in contrast to the return rate obtained in the quantitative approach (91%). It can be concluded with confidence then that further studies of this nature can be conducted among medical educators in Chile.

SUMMARY OF CONCLUSIONS

Based upon the results of the present study it can be concluded that:

1. Chilean medical educators of five out of six Universities that have medical schools are in favor of the establishment of a National Medical Examination. However, they showed concern about who would design and administer the proposed NME.

2. Chilean medical educators of the schools of medicine participating
in the study agreed that the University is the organization that should manage a NME.

3. Faculty members in the schools of medicine are aware of the need for the development of new assessment methodologies to assess students' clinical performance.

4. The interviewees who participated in the qualitative study agreed that the lack of leadership and the lack of interest on the part of the Chilean medical schools of medicine were the main factors that have inhibited the development of a NME in Chile.

5. Chilean medical educators are concerned about the quality of assessment of competencies of their graduates. Faculty members appear to be open to further study and debate of the issue.

6. Interviewees in the qualitative study agreed that accreditation, certification and licensure are useful processes to assure the protection of the public's health.

7. Chilean medical educators are aware that any intent to develop a NME in Chile should define, very clearly, the objectives of such an examination.

8. Similar studies of the nature of the present investigation can be conducted among medical educators as it was indicated by the high return rate obtained in the quantitative approach.

RECOMMENDATIONS

1. If Chilean medical education assessment continues to rely on the current model based only on completion of course program and faculty evaluation to grant MD degrees it is recommended that implementation of an accreditation process be developed that will
help achieve the following goals: document the characteristics of medical education programs; course content and educational strategies; and the judgment of faculties prior to awarding the MD degree. To be accredited should mean then that the recipients of the degree have studied in an acceptable curriculum and have been evaluated and judged by an acceptable faculty.

2. If Chilean medical education decide to explore a new model based on a licensing type model, then a systematic process is suggested: (1) discuss the purpose of a NME at national level including physicians of the national health service and private practice, and students; (2) design and implement methodologies to determine content of medicine to be evaluated, including general knowledge and clinical skills, and attitudes; (3) organize a task force to study models of licensing exams; (4) develop pilot programs; (5) determine the entity that will manage the NME, and (6) study and develop evaluation methodologies to be employed by the NME.

IMPLICATIONS FOR FUTURE STUDIES

1. A feasibility study should be conducted to identify the problems and determine the potential for development of a NME in Chile.

2. The present study should be replicated, if possible, among the faculty of the Chilean Catholic University School of Medicine.

3. A series of studies should be undertaken on the following areas: (1) determination of competencies in medical education; (2) evaluation of the effectiveness of the Chilean medical education admission process as a mechanism for the assurance of students' competencies; (3) the evaluation of attitudes and interpersonal
skills; and (4) the evaluation of psychomotor skills through a NME.

4. To conduct studies on faculty development programs as a means for assuracing medical students' competencies.


Phillipson, E. 1971. Evaluation — How is it in a Professional Education in Gilbert, J. (Ed). *Proceedings of Competence on Evaluation in Medical Education*. Faculty of Medicine, University of Alberta, Canada.


BIBLIOGRAPHY


Federation of State Medical Boards. February, 1984.


Flexner, A. 1922. *Medical Education in the United States and Canada: A Report to the Carnegie Foundation for the Advancements of Teaching.*


Hubbard, J. 1971. "Aims of Evaluation." In Gilbert, J. (Ed.) *Proceedings of the Conference in Evaluation in Medical Education.* Faculty of Medicine, University of Alberta, Canada.


Publicación Oficial del Gobierno de Chile. 1981. Diario Oficial Enero 3, DFL N1, Santiago, Chile.


Dear Professor:

In Chile since a few years ago the issue of establishing a National Medical Examination as an evaluation mechanism of medical graduates' competence for practicing medicine has been discussed. Moreover, during the National Medical Education Seminar held in Santiago in June, 1984 it was advanced that the issue was once again calling the attention of several institutions related to medical education and health care.

For the purpose of contributing to the analysis of the problem we have designed a study which aim is to evaluate Chilean medical educators' perceptions about establishing in Chile a National Medical Examination. This study has received support from The Ohio State University, University of Concepcion, and the Chilean Association of Medical Colleges.

Please find enclosed to this letter a questionnaire designed to complete the study. Your opinions will be very valuable. If you wish to give additional opinions in the form of comments, please use the reverse of the questionnaire. We hope to complete the study by October 30, 1985. For this reason your promptness in responding will be very useful.

Thank you very much.
QUESTIONNAIRE

Instructions: In front of each statements draw a circle around the number-option that best represents your opinion.

1. Chilean medical graduates in order to the MD degree must demonstrate competence through a uniform evaluation that certifies it. 5 4 3 2 1

2. General public rely more on those professionals that have demonstrated competence in the profession through a competence examination. 5 4 3 2 1

3. A national medical examination should include practical exam with patients. 5 4 3 2 1

4. Chilean medical schools are capable enough for the design and management of a NME to assure medical graduates' competences. 5 4 3 2 1

5. A national medical examination would be a means for evaluating effectiveness of the medical education programs of the medical schools. 5 4 3 2 1

6. Only with the establishment of a national medical examination would it be possible to homogenzie the minimum curriculum of the medical schools. 5 4 3 2 1

7. A national medical examination should be able to determine presence or absence of required attitudes to practice medicine. 5 4 3 2 1

8. To evaluate medical graduates' competences before engag in in practice, must be a function of a specialized, independent extrauniversity entity. 5 4 3 2 1

9. The achievement of the necessary abilities to practice medicine is best assured through a national medical examination. 5 4 3 2 1
10. Evaluation of Chilean medical education output through a national medical examination would improve medical education programs.

5 4 3 2 1

11. The government must be responsible for licensing the medical graduates once they have proved certification of professional competence.

5 4 3 2 1

12. A national medical examination would homogenize medical graduates' minimum competencies to practice medicine in primary level of care.

5 4 3 2 1

13. Only a national medical examination would provide equality of opportunity to medical graduates for applying to the National Health Service positions.

5 4 3 2 1

14. Objective tests are primarily effective in evaluating the cognitive area.

5 4 3 2 1

15. The deficiencies of evaluation methods in medicine makes necessary the creation of a national medical examination.

5 4 3 2 1

16. Medical graduates in order to be granted a MD degree must demonstrate medical competence in decision-making, knowledge, clinical abilities and attitudes.

5 4 3 2 1

17. Medical graduates in order to be granted a MD degree should demonstrate through an examination, ability for diagnosis and patient management.

5 4 3 2 1

18. Medical graduates' professional competence should be assessed by a specialized governmental entity.

5 4 3 2 1

19. Current evaluation methods of medical students are not sufficient for assuring professional competence.

5 4 3 2 1

20. Current evaluation of students achievement does not assess the capacity to take responsibility for the patient's health care.

5 4 3 2 1

21. A better patient's health protection can be assured by evaluating medical competence through a national medical examination.

5 4 3 2 1
22. Universities should be in charge of the evaluation of professional competence of medical graduates.


24. The absence of effective mechanisms of evaluation medical students increases the risks of malpractice.

25. Medical graduates before engaging in practice should demonstrate through an evaluation that they possess the necessary psychomotor skills.

Demographic Information Section

Introduction

In order to enrich the study we are asking you in the following items information that will be handled confidentially. Please answer the following items by writing the corresponding number in brackets () in the space provided [] on the right side of the page.

26. Your university

University of Chile (1)
University Austral (2)
University of Valparaiso (3) [ ]
University of Concepcion (4)
University of Frontera (5)

27. Your academic rank

Full Professor (1)
Associate Professor (2) [ ]
Assistant Professor (3)

28. Your teaching experience

Less than 5 years (1)
From 5 to 10 years (2) [ ]
From 10 to 20 years (3)
More than 20 years (4)
29. Your sex

Feminine (1) [ ]
Masculine (2)

NOTE: Please return the questionnaire to the Office of Medical Education of your School of Medicine.

Thank you.
Senor
Prof. Dr.
Depto.

PRESENTE/

En Chile desde hace algunos anos se ha venido discutiendo la necesidad de implantar un Examen Nacional de Medicina como mecanismo de evaluacion de la competencia medica de los egresados de medicina del pais. Mas aun, durante el seminario de Educacion Medica realizado en Junio de 1984 en Santiago, se adelanto la noticia de que el tema estaba nuevamente captando la atencion de varias instituciones relacionadas con la ensenanza de la medicina. Con el proposito de contribuir al analisis del problema hemos disenado un estudio cuyo objetivo es evaluar la percepcion de un grupo seleccionado de educadores medicos chilenos acerca de la implantacion en Chile de un Examen Nacional de Medicina.

Este estudio cuenta con el respaldo de la Universidad del Estado de Ohio, Universidad de Concepcion y Asociacion Chilena de Facultades de Medicina.

Adjuntamos a esta carta de presentacion un cuestionario disenado para completar el estudio. Sus opiniones sin duda, seran muy valiosas. Si Ud. desea incluir algun comentario adicional le rogamos hacerlo en el reverso del cuestionario.

Esperamos completar el estudio alrededor del 30 de Octubre de 1985, por lo que su prontitud en responder sera de much a utilidad.

Mucas gracias.
EVALUACIÓN DE LA PERCEPCIÓN DE LOS EDUCADORES MÉDICOS CHilenos
Sobre la Implantación de Un Examen Nacional de Medicina

Cuestionario de Opinión

Instrucciones: Frente a cada aseveración haga un círculo alrededor del número que mejor represente su opinión.

1. Los egresados de medicina del país, para optar al título de Médico-Cirujano deben demostrar que están aptos para ello, a través de una evaluación uniforme para todos que lo acredite.  
   MA  A  I  D  MD
   5  4  3  2  1

2. El público en general confía más en aquellos profesionales que han demostrado dominio de la profesión a través de algún examen de competencia.  
   5  4  3  2  1

3. Un examen nacional de medicina no necesita incluir evaluación práctica con pacientes.  
   5  4  3  2  1

4. Las Facultades de Medicina Chilenas están capacitadas para diseñar y administrar un examen nacional que acredite competencia profesional de los egresados de medicina del país.  
   5  4  3  2  1

5. Un examen nacional de medicina sería el medio para evaluar la efectividad de los programas de educación médica de las Universidades chilenas.  
   5  4  3  2  1

6. Solo la implantación de un examen nacional de medicina uniformará el currículum mínimo de las Facultades Chilenas de Medicina.  
   5  4  3  2  1

7. Un examen nacional debería ser capaz de determinar la presencia o ausencia de las actitudes requeridas por la profesión médica.  
   5  4  3  2  1

8. Evaluar la competencia profesional de los egresados de medicina antes de ejercer, es función de una entidad independiente especializada, extrauniversitaria.  
   5  4  3  2  1

9. El logro de las habilidades destrezas necesarias para ejercer medicina se garantizaría mejor a través de un examen nacional.  
   5  4  3  2  1
10. Evaluando el producto de la educación médica chilena a través de un examen nacional, se lograría un mejoramiento de los programas de formación médica.

11. El Estado debe ser la entidad responsable de autorizar el ejercicio de la profesión médica de los egresados, una vez que acrediten competencia profesional.

12. Un examen nacional de medicina lograría homogenizar las competencias mínimas de los egresados del país para ejercer medicina en un nivel de atención.

13. Un examen nacional mejoraría las posibilidades de todos los egresados de medicina para optar a cargos del Servicio de Salud y formación de Post-Grado.

14. Las pruebas objetivas son efectivas principalmente en la evaluación del área cognoscitiva.

15. Las deficiencias de las evaluaciones del aprendizaje en medicina, hacen necesario la creación de un examen nacional para optar al título de médico cirujano.

16. Para optar al título de médico-cirujano se debe demostrar competencia médica en las áreas de toma de decisiones, conocimiento, habilidades clínicas y actitudes.

17. Los egresados de medicina para optar al título de médico cirujano deben demostrar a través de un examen la habilidad de diagnóstico y manejo de paciente.

18. La competencia profesional de los egresados de medicina debe ser evaluada por una entidad especializada del estado.

19. Las evaluaciones de los estudiantes en las carreras de medicina, por sí solas, son insuficientes para garantizar la competencia profesional de los egresados.

20. Las evaluaciones del aprendizaje, actualmente en práctica en medicina, excluyen la determinación de la capacidad de asumir responsabilidad en la atención del paciente.
21. Evaluando a través de un examen nacional se aseguraría una mayor protección de salud de la población.

22. Evaluar la competencia profesional de los egresados de medicina es función exclusiva de las universidades.

23. Las evaluaciones del aprendizaje en medicina, actualmente vigentes, miden básicamente capacidad de recuerdo.

24. La ausencia de mecanismos efectivos de evaluación del aprendizaje en medicina aumenta los riesgos de malpraxis.

25. Los egresados de medicina antes de ejercer deberían demostrar a través de una evaluación, que poseen las destrezas psicomotoras necesarias.

Continúa página siguiente
SECCION DE DATOS PERSONALES

Con el propósito de enriquecer el análisis de los cuestionarios, solicitamos a continuación información que será manejada con discreción.

Por favor resinda a los siguientes ítems escribiendo el número correspondiente en el recuadro a la derecha de la hoja.

26. Su universidad:

Università de Chile (1)
Università Austral (2)
Università de Valparaíso (3) [ ]
Università De Concepcion (4)
Università de la Frontera (5)
Università Católica (6)

27. Su categoría académica de acuerdo a su Universidad:

Profesor (1)
Profesor Adjunto/Asociado (2) [ ]
Profesor Auxiliar/Asistente (3)

28. Su experiencia docente:

Menos de 5 años (1)
De 5 a 10 años (2) [ ]
De 10 a 20 años (3)
Mas de 20 años (4)

28. Se sexo:

Femenino (1) [ ]
Masculino (2)

NOTA: Por favor devolver cuestionario a la Oficina de Educación Médica de su Facultad.

Muchas gracias.
INTERVIEW SCHEDULE

Introduction

In Chile since several years ago the issue about establishing a National Medical Examination as an External mechanism to evaluate graduates' competences to practice medicine has been discussed. Moreover, in 1977 during the III National Seminar of Medical Education it was recommended the establishing of one, and during the Seminar on Medical Education held in Santiago, June, 1984, this issue was again discussed.

1. I would like to know what is your opinion in relation to the possible factors that may have inhibited the development of a NME as it were recommended in 1977.

2. Now, I would like your opinion about the participation of the students in the decision-making process about establishing in Chile a NME.

3. Based on your experience, which is the attitude of physicians toward responding survey instruments, specifically questionnaires?

4. I would like to know your comments about the concepts of licensing [to practice medicine], certification [of specialties] and accreditation [pregrade].

5. As you may recall the purpose of this study is to evaluate Chilean medical educators perceptions about establishing a national medical examination in Chile. The population selected for the study is composed of full, associate and assistant professors of each of the six Chilean Universities that have medical schools. Five out of the six schools of medicine participated in the study. The Catholic University School of Medicine by decision of the acting dean refused to participate arguing that the study threatened the university's autonomy. Now I would like to know what is your opinion in relation to the refusal attitude of Catholic University School of Medicine.
Desde hace varios anos se ha venido discutiendo en Chile, formal e informalmente la idea de implantar un examen nacional de medicina como mecanismo externo de evaluación de competencia profesional de los egresados de medicina chillanos. Mas aun, en 1977 durante el Seminario de Educación Medica se recomendo su implantacion y durante el Seminario de 1984 nuevamente se planteo la misma inquietud.

1. Me gustana conocer su opinion en relacion a factores que pudieron haber inhibido la implantacion de un Examen Nacional de Medicina, recomendado en 1977.

2. Ahora, que opina ud sobre el papel que le corresponde a los estudiantes en la toma de decisiones sobre la implantar eno un examen nacional?

3. Basado en su experiencia ynal cree ud que es la predisporcion de los medicos en Chile para responder cuestionaires?

4. Me gustaria conocer su opinion sobre los conceptos de licencia para ejercer la profesion, certification de especialidades y acreditacion de pregarado.

5. Me gustaria conocer su opinion acerca de la actitud de la Facultad de Medicina de la Universidad Catolica de no participan en el estudio.
ITEMS THAT CONSTITUTE FACTOR NEED FOR A NME.

FACTOR 1

1. Chilean medical graduates to obtain the MD degree must demonstrate competence through a uniform evaluation that certifies it.

2. General public rely more on those professionals that have demonstrated competence in the profession through a competence examination.

5. A national medical examination would a means for evaluating effectiveness of the medical education programs of the medical schools.

6. Only with the establishment of a national medical examination would be possible to homogenize the minimum curriculum of the medical schools.

9. The achievement of the necessary abilities to practice medicine is best assured through a national medical examination.

10. Evaluation of Chilean medical education output through a national medical examination, would improve medical education programs.

12. A national medical examination would homogenize medical graduates' minimum competencies to practice medicine in primary level of care.

13. Only a national medical examination would provide equality of opportunity to medical graduates for applying to the National Health Service positions.

15. The deficiencies of evaluation methods in medicine makes it necessary the creation of a national medical examination.

21. Evaluating medical competence through a national medical examination a better patients' health care protection can be assured.

ITEMS THAT CONSTITUTE FACTOR ENTITY FOR THE MANAGEMENT OF A NME.

FACTOR 2

8. To evaluate medical graduates' competences before engaging in practice, must a function of a specialized, independent extrauniversity entity.

11. The government must be responsible for licensing the medical graduates once they have proved certification of professional competence.
18. Medical graduates' professional competence should be assessed by a specialized governmental entity.

22. Universities should be in charge of the evaluation of professional competence of medical graduates.

ITEMS THAT CONSTITUTE FACTOR OF AREAS TO BE ASSESSED BY A NME.

FACTOR 3

7. A national medical examination should be able to determine presence or absence if required attitudes to practice medicine.

16. Medical graduates in order to be granted a MD degree they must demonstrate medical competence in decision-making, knowledge, clinical abilities and attitudes.

17. Medical graduates in order to be granted a MD degree should demonstrate through an examination, ability for diagnosis and patient management.

25. Medical graduates before engaging in practice should demonstrate through an evaluation that they possess the necessary psychomotor skills.

ITEMS THAT CONSTITUTE ADEQUACY OF EVALUATION METHODS EMPLOYED BY CHILEAN MEDICAL EDUCATORS

FACTOR 4

19. Current evaluation methods of medical students are not sufficient for assuring professional competence.

20. Current evaluation of students achievement does not assess the capacity to take responsibility for the patient's health care.


24. The absence of effective mechanisms of evaluation medical students increases the risks of malpractice.

ITEM RELATED TO PRACTICAL EXAM AS PART OF A NME.

FACTOR 5

3. A national medical examination should include practical exam with patients.
ITEM RELATED TO THE SCHOOLS OF MEDICINE AS BEING THE ORGANIZATION TO DESIGN AND MANAGE A NME.

FACTOR 5

4. Chilean medical schools are capable enough for the design and management of a NME to assure medical graduates' competences.

ITEM RELATED TO THE MEASUREMENT POTENTIAL OF OBJECTIVE-TESTS IN MEDICAL EDUCATION

FACTOR 6

14. Objective-tests are primarily effective in evaluating the cognitive area.
**TABLE 36**

**ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE NEED FACTOR TOTAL**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>1597.655</td>
<td>4</td>
<td>399.414</td>
<td>7.20</td>
<td>0.000</td>
</tr>
<tr>
<td>Rank</td>
<td>187.638</td>
<td>2</td>
<td>93.819</td>
<td>1.69</td>
<td>0.187</td>
</tr>
<tr>
<td>Interaction</td>
<td>530.365</td>
<td>8</td>
<td>66.296</td>
<td>1.19</td>
<td>0.304</td>
</tr>
<tr>
<td>Error</td>
<td>11211.691</td>
<td>202</td>
<td>55.503</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 37

ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE ENTITY FACTOR TOTAL

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>96.078</td>
<td>4</td>
<td>24.020</td>
<td>1.91</td>
<td>0.110</td>
</tr>
<tr>
<td>Rank</td>
<td>56.935</td>
<td>2</td>
<td>28.468</td>
<td>2.26</td>
<td>0.107</td>
</tr>
<tr>
<td>Interaction</td>
<td>97.346</td>
<td>8</td>
<td>12.168</td>
<td>0.97</td>
<td>0.463</td>
</tr>
<tr>
<td>Error</td>
<td>2542.155</td>
<td>202</td>
<td>12.584</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 38

ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE AREAS FACTOR TOTAL

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>69.674</td>
<td>4</td>
<td>17.418</td>
<td>2.57</td>
<td>0.039</td>
</tr>
<tr>
<td>Rank</td>
<td>3.356</td>
<td>2</td>
<td>1.678</td>
<td>0.25</td>
<td>0.781</td>
</tr>
<tr>
<td>Interaction</td>
<td>56.722</td>
<td>8</td>
<td>7.090</td>
<td>1.05</td>
<td>0.402</td>
</tr>
<tr>
<td>Error</td>
<td>1368.656</td>
<td>202</td>
<td>6.776</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G
### TABLE 39

#### ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE METHODS FACTOR TOTAL

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>45.414</td>
<td>4</td>
<td>11.354</td>
<td>1.38</td>
<td>0.240</td>
</tr>
<tr>
<td>Rank</td>
<td>2.164</td>
<td>2</td>
<td>1.082</td>
<td>0.13</td>
<td>0.876</td>
</tr>
<tr>
<td>Interaction</td>
<td>56.130</td>
<td>8</td>
<td>7.016</td>
<td>0.86</td>
<td>0.555</td>
</tr>
<tr>
<td>Error</td>
<td>1656.322</td>
<td>202</td>
<td>8.200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 40
ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE PRACTICE FACTOR TOTAL

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>11.402</td>
<td>4</td>
<td>2.851</td>
<td>1.71</td>
<td>0.150</td>
</tr>
<tr>
<td>Rank</td>
<td>4.394</td>
<td>2</td>
<td>2.197</td>
<td>1.31</td>
<td>0.271</td>
</tr>
<tr>
<td>Interaction</td>
<td>8.755</td>
<td>8</td>
<td>1.094</td>
<td>0.65</td>
<td>0.731</td>
</tr>
<tr>
<td>Error</td>
<td>337.661</td>
<td>202</td>
<td>1.672</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 41
ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE SCHOOL-MANAGEMENT FACTOR TOTAL

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>4.087</td>
<td>4</td>
<td>1.021</td>
<td>1.25</td>
<td>0.290</td>
</tr>
<tr>
<td>Rank</td>
<td>2.468</td>
<td>2</td>
<td>1.234</td>
<td>1.51</td>
<td>0.222</td>
</tr>
<tr>
<td>Interaction</td>
<td>2.970</td>
<td>8</td>
<td>0.371</td>
<td>0.45</td>
<td>0.886</td>
</tr>
<tr>
<td>Error</td>
<td>164.848</td>
<td>202</td>
<td>0.816</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>SS</td>
<td>df</td>
<td>MS</td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>University</td>
<td>8.955</td>
<td>4</td>
<td>2.239</td>
<td>4.17</td>
<td>0.003</td>
</tr>
<tr>
<td>Rank</td>
<td>0.293</td>
<td>2</td>
<td>0.146</td>
<td>0.27</td>
<td>0.761</td>
</tr>
<tr>
<td>Interaction</td>
<td>5.201</td>
<td>8</td>
<td>0.650</td>
<td>1.21</td>
<td>0.294</td>
</tr>
<tr>
<td>Error</td>
<td>108.369</td>
<td>202</td>
<td>0.536</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>