INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

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

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
Taylor, Marcia Beth Baxter

THE DEVELOPMENT OF THE MEASURE OF EPISTEMOLOGICAL REFLECTION

The Ohio State University

University Microfilms International

Copyright 1983 by Taylor, Marcia Beth Baxter All Rights Reserved
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ☑.

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Pages with black marks, not original copy ______
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages ☑
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) _______ lacking when material received, and not available from school or author.
12. Page(s) _______ seem to be missing in numbering only as text follows.
13. Two pages numbered _________. Text follows.
14. Curling and wrinkled pages ______
15. Other _________________________________________________________________
THE DEVELOPMENT OF THE
MEASURE OF EPISTEMOLOGICAL REFLECTION

DISSERTATION

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

By

Marcia Baxter Taylor, B.A., M.A.

* * * * *

The Ohio State University
1983

Reading Committee:
John C. Gibbs, Ph.D.
Anne L. Pruitt, Ph.D.
Robert F. Rodgers, Ph.D.
Robert J. Silverman, Ph.D.

Approved By:

Adviser
Department of
Special Services
ACKNOWLEDGEMENTS

The following individuals have contributed in various ways to the direction and scope of this research. I value their contributions and wish to thank them for their support.

Carole Widick for unknowingly serving as a role model, and whose guidance during my master's work influenced my professional direction.

Bob Rodgers, whose expertise and enthusiasm in the Perry scheme was contagious, thus inspiring the present research.

Bob Silverman, whose ethics, sincerity, and commitment to students are admirable.

Anne Pruitt for her planful questioning which significantly enhanced my thinking and research.

John Gibbs, whose work significantly influenced the present research.

Bill Moore for coordinating the use of the Measure of Intellectual Development, intellectual exchange and general encouragement.

John Heidke and Cynthia Straub for their endless hours of rating and moral support.

Rick Doty for graciously sharing his expertise in data analysis.

William Porterfield, my colleague in instrument and manual construction, whose ideas and energy had a significant impact on this research.

Peter Magolda, whose friendship, caring and support were of great assistance in overcoming the frustrations of research.

Mrs. R. W. Baxter, my mother, for her strength of character which has so significantly effected my personal and professional development.
VITA

October 29, 1951 ................ Born - Elida, Ohio

1974 .............................. B.A., Psychology, Capital University Columbus, Ohio

1975 .............................. Social Worker, Franklin County Children Services, Columbus, Ohio

1976 .............................. M.A., Student Personnel Work The Ohio State University Columbus, Ohio

1977-1980 ....................... Residence Hall Director The Ohio State University Columbus, Ohio

1980-1982 ....................... Area Coordinator, Residence and Dining Halls, The Ohio State University Columbus, Ohio

1982-Present ..................... Assistant Director, Student Development, Residence and Dining Halls, The Ohio State University Columbus, Ohio
## TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii
VITA ............................................................................................... iii
LIST OF TABLES ........................................................................ vi

### Chapter

#### I. INTRODUCTION ............................................................ 1
- Purpose of this Study .......................................................... 1
- Overview of the Perry Scheme ............................................ 1
- Application of the Scheme ................................................. 5
- Initial Assessment Techniques ......................................... 7
- Rationale of this Study ....................................................... 9

#### II. REVIEW OF THE LITERATURE ................................. 14
- Perry Scheme of Intellectual Development ...................... 14
- Measurement ..................................................................... 19
  - Loevinger and Wessler ............................................... 19
  - Kohlberg and Gibbs .................................................. 23
  - Knefelkamp ............................................................... 26
  - Widick and Rodgers .................................................. 27
  - Alverno ........................................................................ 29
  - Syracuse ...................................................................... 30
  - Summary ...................................................................... 31

#### III. METHODS ................................................................. 34
- Introduction ....................................................................... 34
- Research Hypotheses ..................................................... 35
- Construction of the Measure of Epistemological Reflection ........................................................................ 36
- Development of the Preliminary Manual ......................... 38
- Sample ............................................................................. 40
- Administration of Instrumentation .................................. 45
- Revision of the Preliminary Manual ................................. 45
- Raters .............................................................................. 49
- Reliability ........................................................................ 49
- Validity ............................................................................ 50
- Limitations ...................................................................... 52
IV. RESULTS.......................................................................................................54
  Interrater Reliability: Domain Ratings ....................................................... 54
  Interrater Reliability: Total Protocol Ratings ................................................. 63
  Internal Consistency....................................................................................... 66
  Validity............................................................................................................ 70

V. DISCUSSION OF THE RESULTS....................................................................79
  Introduction..................................................................................................... 79
  Review of the Results..................................................................................... 79
    Interrater Reliability..................................................................................... 79
    Internal Consistency.................................................................................... 87
    Validity.......................................................................................................... 90
  Implications..................................................................................................... 95
  Limitations...................................................................................................... 96
  Areas for Further Inquiry............................................................................... 96
  Summary.......................................................................................................... 97

APPENDICES
  A. The Measure of Epistemological Reflection .............................................. 99
  B. Requests for Participation......................................................................... 106
  C. Instructions to Subjects............................................................................ 109
  D. Demographic Data Forms......................................................................... 111
  E. Consent Form............................................................................................ 113
  F. Preliminary Manual.................................................................................... 114
  G. Revised Manual......................................................................................... 141

REFERENCE NOTES.......................................................................................... 183
BIBLIOGRAPHY................................................................................................. 185
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Summary of Current Measures of the Perry Scheme</td>
<td>33</td>
</tr>
<tr>
<td>2. Description of the Sample</td>
<td>42</td>
</tr>
<tr>
<td>3. Socioeconomic Status* by Level of Education</td>
<td>44</td>
</tr>
<tr>
<td>4. Percentage Agreement of Highly Trained Raters Domain Ratings by Sample</td>
<td>55</td>
</tr>
<tr>
<td>5. Percentage Agreement of Highly Trained Raters' Compromise Domain Ratings and Self-Trained Expert Domain Ratings by Sample</td>
<td>57</td>
</tr>
<tr>
<td>6. Percentage Agreement of Rater One Domain Ratings and Self-Trained Expert Domain Ratings</td>
<td>59</td>
</tr>
<tr>
<td>7. Percentage Agreement of Rater Two Domain Ratings and Self-Trained Expert Domain Ratings</td>
<td>60</td>
</tr>
<tr>
<td>8. Percentage Agreement of Self-Trained Raters Domain Ratings</td>
<td>61</td>
</tr>
<tr>
<td>9. Percentage Agreement of Self-Trained Raters Domain Ratings and Highly Trained Raters' Compromise Domain Ratings</td>
<td>62</td>
</tr>
<tr>
<td>11. Correlations of Total Protocol Ratings of Self-Trained with Highly Trained and Self-Trained Expert Raters</td>
<td>65</td>
</tr>
<tr>
<td>12. Percentage Agreement on Modal TPR's for Highly Trained Raters and Self-Trained Expert</td>
<td>67</td>
</tr>
<tr>
<td>13. Percentage Agreement on Modal TPR's for Self-Trained Raters with Highly Trained and Self-Trained</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER ONE
Introduction

Purpose of the Study

The purpose of this study is to describe the development and initial validation of an instrument to assess intellectual development on the Perry scheme (Perry, 1970). The instrument, entitled the Measure of Epistemological Reflection (MER), is an attempt to provide an accurate measure of development on the Perry scheme that is suitable for practical use. The Perry scheme, developed on the basis of interviews with Harvard college students, describes ways persons make meaning of experiences. The scheme has potential for application in higher education to enhance the learning experiences and foster the personal growth of students. Application and refinement of the scheme to date has been affected by the absence of assessment tools that can be utilized in a practical manner. Thus the MER is intended to provide an improved method of assessment to in turn enable further use and refinement of the Perry scheme.

Overview of the Perry Scheme

In the early 1950's the staff of the Bureau of Study Counsel at Harvard College became interested in the variety of responses which students made to the academic and social challenges of their college experience. William G. Perry, Director of the Bureau, began a study of
these differences. The preliminary data prompted him to expand the study in an attempt to define the course of development that seemed to appear in the initial sets of data. After nearly 500 interviews, Perry (1970) outlined a theory of intellectual and ethical development that he felt represented the students' ways of making meaning out of their college experiences. The theory is cognitive developmental in its nature (Knefelkamp, Widick and Parker, 1978; Rodgers, 1980). Perry focuses on cognitive structures as the way individuals make meaning of their experience. These structures are consistent groups of perceptions about the world and form the reasoning behind individuals' perceptions and organization and evaluation of their perceptions. Development is characterized by a progression from holistic, increasingly differentiated and integrated structures. Distinctly different structures exist over the course of development each of which characterizes a stage or position in the developmental scheme. Positions in the Perry scheme adhere to the criteria for cognitive developmental stages (Knefelkamp, Widick and Parker, 1978; Perry, 1970; Rodgers, 1980). They are hierarchical and integrative, with each new structure representing a qualitative change in thinking. The positions follow an ordered sequence and regression on the scheme is rare.

Change in cognitive structures occurs through equilibration, a Piagetian notion (Piaget, 1950; Piaget, 1960). When experiences are incongruent with one's way of viewing the world, the discrepancy is initially assimilated into the current cognitive structure. However, when sufficient and appropriate dissonance is experienced repeatedly, the cognitive structure often is altered to accomodate the new
experience and restore equilibrium between the cognitive structure and perceived reality.

Perry identifies nine qualitatively different positions in his scheme of intellectual and ethical development. Perry groups the nine positions into three categories: dualism, relativism and commitment. The dualistic world view includes the first three positions and is a dichotomous perspective, interpreting the world in either-or, right-wrong terms. The Position One dualist believes that truth is known by authorities, whose role it is to teach the truth to others. The realization that some authorities do not seem to know the truth leads to Position Two. The dilemma is solved by labelling authorities who do not know the truth as bad authorities thus preserving the notion that truth is known, at least by good authorities. When good authorities, however, admit that they do not know all of the truth, the Position Two perspective is challenged. Position Three is characterized by an acknowledgement that some truths are unknown, accompanied by the belief that the uncertainty is only temporary. The truth will be known in the future.

In the transition from Position Three to Position Four the dualistic viewpoint gives way to relativistic thinking. Upon discovering the extent to which absolute truth is unknown, the dualist changes from an "either-or" explanation of the world to an "everything is uncertain" viewpoint. All persons have a right to their own opinion, and no one's opinion is any better than anyone else's. The key in Position Four is that no criteria exist to make judgment since truth does not exist and authorities can no longer provide the right answer.
At this point the making of meaning shifts from finding the right answer to the process of making meaning. This opens the way for Position Five which entails supporting one's opinions in a context. This results in the ability to judge and choose again, even though choices are now made on a different basis. This basis involves the rules of evidence in different fields and paradigms of inquiry. Movement into Position Six is accompanied by the realization that the world, including self, is indeed uncertain. This realization is accompanied by an awareness that commitment without certainty is necessary in order to act in a relativistic world. The making of initial commitments characterizes Position Seven. This step may initially be seen as resolving uncertainty; however, soon one becomes aware that each commitment opens up another set of issues that in turn require exploration, judgments and commitment without certainty. The realization and grappling with these implications are the focus of Position Eight. The attempt to balance the implications continues into Position Nine where the development of commitments within commitments becomes possible. Position Nine also brings the realization that choices are on-going and commitments are re-evaluated. Perry describes a few variations from orderly progression through the scheme. Temporizing involves hesitating in one position, exploring the implications at length before risking further growth. Perry indicates that one can escape in the middle stages or retreat into dualistic thinking.

Positions Six through Nine involve the development of commitment, a somewhat different focus than the first five positions. Heffernan (1971) and King (1978), among others, have noted that the first five
positions focus on cognitive developmental phenomena whereas the latter four seem to focus on identity issues and may not be cognitive in nature. Hence, it may be the case that the cognitive developmental aspects of Perry's theory are completed at Position Five.

Perry's discovery of the positions of intellectual and ethical development of college students emerged from rich interview data, collected in a highly unstructured format. His method was to ask students what in their experience of the preceding year stood out for them. They were asked to elaborate and clarify as necessary. The collective responses led Perry to outline the scheme described above. In an attempt to verify the existence of the scheme in the data, Perry introduced the scheme to six individuals who later evaluated the interview data. After studying the scheme, these six individuals listened to the recorded interviews and assigned ratings to indicate the level of development evidenced in the data. The agreement among the raters when rating four year protocols, single interviews and excerpts was sufficiently high to validate that the developmental scheme was observable in the data.

Applications of the Scheme

Following publication of the scheme in 1969, various studies arose focusing on descriptions of students at various types of colleges and application of the theory to practice. Rodgers (1975-1979) studied populations in liberal arts colleges and large state universities. Widick studied liberal arts colleges and rural college-age non-college populations. Knefelkamp studied university students and adults. The early applications focused on the classroom setting. Widick and
Knefelkamp (1974, 1975) were the first to report use of the scheme in developmental instruction at the college level. Their instructional methods were developed on the basis of the Perry scheme and employed in an English literature course. The authors reported higher subject mastery and development under these methods than in a comparable course taught with traditional methods. Meanwhile Copes, beginning with a doctoral dissertation in 1974, reported use of the scheme in the teaching of mathematics. Stephenson and Hunt (1977) reported a study supporting the Widick and Knefelkamp developmental instruction method. Other instructional areas in which the scheme was used include English drama (Sorum, 1976), counselor education (Knefelkamp and Cornfeld, 1977), history (Widick and Simpson, 1978), and engineering (Culver and Hackos, 1981).

While these studies focused on teaching methods and organization of individual courses, studies on the broader aspects of teaching and organization emerged. Kovacs (1977) explored use of Perry's scheme in curriculum design while Parker (1978) addressed the notion of improving instruction through faculty consultation and the Perry scheme. Knefelkamp and Cornfeld (1978) applied the scheme to graduate curriculum in counseling. Nowakowski and Laughney (1978) applied the scheme to training in health careers.

Career education and development comprised a second area of application. The initial work in this area was reported by Widick and Stulck (1975) and Slepitza and Knefelkamp (1976). Exploration of applications in career development has continued by these authors and others (Touchton, 1978; Wiley and Yiesar, 1979). Finally, application
has been extended to include areas beyond the traditional instructional setting. Knefelkamp, Widick and Stroad (1976) have reported use of the scheme in the counseling of women. Similarly Chickering (1976) utilized Perry's notions in the advising of adult learners at Empire State College. Clement (1977) pursued peer training by way of the Perry scheme.

Initial Assessment Techniques

In order for applications of the scheme to occur, some method of assessment was required. Attempts to measure development on the Perry scheme began in two directions. One direction involved interviewing while the other focused on paper and pencil instrumentation. A second important distinction among the techniques under study involved production versus recognition tasks.

Initial interview formats were of the recognition task variety. In 1975, Clinchy and Zimmerman developed a structured interview involving sets of statements representative of groups of Perry positions. The interviewer asked the student to choose one of the statements from each set and talk about that choice. Probing techniques were employed to gain an understanding of the student's thinking. Thus the choice of a statement gave an indication of position, and elaboration provided additional assessment data. On the basis of Perry's explanation of development it is reasonable to expect that a student would choose a statement within the position or one position beyond his/her present level of thinking. The probing technique therefore becomes an essential part of the assessment in order to determine whether the students can describe the rationale for their choice of statements. The interviews
were recorded and analyzed by the authors to determine major and minor positions present in the responses. This method thus calls for trained interviewers and trained raters.

Kurfiss (1977) developed an assessment technique involving paraphrasing of paragraphs presented to the respondent. The rationale for assessment in this manner was that students would accurately paraphrase paragraphs representative of their developmental position but be unable to do so with positions above their level of thinking. The paragraphs focused on five topics: making moral decisions, the role of the counselor or advisor, how essays should be graded, responsibilities of instructors in teaching and the nature of academic knowledge. Respondents were asked to paraphrase both orally and in writing. Responses were scored as "hits" or "misses" depending on whether they matched accurate paraphrases included in the scoring manual. In addition to assessing degree of comprehension, Kurfiss also asked respondents to rank the items that they had paraphrased to assess whether positions preferred in each topic added to the position assessment. This technique also requires trained interviewers to probe paraphrased responses and trained raters.

Kitchener and King (1977) have explored the commitment phases of Perry's theory, and subsequently developed their own Reflective Judgment Scheme. They introduced a more structured interview format for assessing positions. Interviewees are presented with a dilemma about questions of knowledge and asked to express their views about the dilemma. Probing questions are used by the interviewers to clarify responses and encourage elaboration of the students' thinking. This
method has produced ratable data but also requires both trained interviewers and trained raters in order to complete an assessment. Their method differs from the previous two in that the response is completely produced by the subject.

The initial work on paper and pencil measurement was done by Widick and Knefelkamp in conjunction with their doctoral dissertations mentioned earlier. They developed the KneWi, which originally consisted of ten sentence stems and two essay questions. From their piloting of various sentence stems they chose five which elicited the most ratable data. The stems and essays purport to sample students' thinking since the self generated responses should be representative of students' developmental position. The authors of the KneWi developed a preliminary rating manual through which raters could be trained to evaluate data. The KneWi has been the stimulus for further work in instrumentation, much of which has been done by the original authors and their colleagues. Although these techniques are paper and pencil tools, all require trained raters to evaluate data. In addition the stimuli sometimes produce data that require substantial inference to rate. Since these measures serve as the background against which the current instrument has been developed, they are explained in detail in Chapter Two. The literature review will focus primarily on written assessment techniques requiring production tasks since they are most relevant to the present study.

Rationale for this Study

The purpose of this research is to develop the Measure of Epistemological Reflection and thereby provide an improved measure of
intellectual development on the Perry scheme. A major issue related to the Perry scheme revolves around the attempt to move from theory to practice. Perry presented his theory as descriptive of student development and did not advance his scheme as a method by which to facilitate development. When the latter notion occurred to other researchers, Perry argued that one could not force development. He opposed prescriptive uses of the theory and wrote a paper in which he contended that assessment on the Perry scheme was unreasonable and unethical. Concern around the forcing of development has abated due to the results of the studies mentioned earlier. The studies have resolved two issues relevant to use of the Perry scheme. First, the scheme can be used to effect cognitive development in a general sense with preliminary assessment techniques. Perry expressed encouragement in his chapter in *The Modern American College* that teachers' basic awareness of the scheme enables them to identify the broader categories of students' thinking and thus respond more effectively. Second, the work on application has shown that use of the Perry scheme to facilitate development is not a manipulative process in a negative sense. Rather it is a process of organizing experience and communication with students that is close enough to their making of meaning to be understood, yet beyond their perspective enough to challenge their thinking.

Thus at present there appears to be a substantial body of evidence supporting the existence and validity of the theory and its use in facilitating development through various forms of practice. Perry's research has not been replicated elsewhere. King and Kitchener's (1977) work on the Reflective Judgment Scheme and Clinchy and Zimmerman's
A 1977 study of women comes closest to verifying the existence of Perry's cognitive structures or ones similar to Perry's. Research studies utilizing different assessment techniques also verify aspects of the theory but cannot be used collectively to refine the theory due to their inconsistent measurement techniques.

Kohlberg's model of phases in developmental research (Gibbs, 1981) would suggest that in order for the Perry scheme to be utilized effectively in facilitating student development, it is essential that more rigorous assessment methods be developed. Kohlberg suggests that the initial phase of developmental research involves working with large samples of individuals to identify the structural patterns of development. The completion of this phase rests upon the researcher's satisfaction that the sequence of development has been sufficiently identified. It would appear that the work of Perry, King and Kitchener and Clinchy and Zimmerman, supported by the remaining studies, indicate sufficient completion of this phase as relates to Perry's scheme.

Kohlberg's second phase involves the development of instrumentation to elicit data specific to the theory in order to test the validity and reliability of the scheme in a more scientific fashion. It is in this phase that the broad range of assessment techniques hinders progress to the third phase for the Perry scheme. This last phase, as Kohlberg describes it, is the process of further defining and refining the original theory. In order to effectively engage in phase three one needs to be assured that phase two measurement is sound and will not hinder the refinement of the theory. Current assessment methods for the Perry scheme may not be generating data that are specific enough to use
in refining the scheme. Rating procedures have not been rigorously
developed and the training of raters has not been systematically
conducted.

Another compelling reason exists to establish a more specific
assessment technique. In order to apply the Perry theory in student
development practice, an accurate and usable method of measurement is
necessary. Intensive interviews, as indicated earlier, are too time
consuming to conduct and require highly trained raters for accurate
developmental assessment. Written instruments are much more realistic
when large numbers of students are involved, as is evidenced by their
use in numerous applications of the Perry scheme. Even the written
instruments, however, are subject to numerous difficulties. Those that
provide sentence stems and essay questions provide a stimulus to which
the respondent reacts. Since respondents may choose to elaborate in
varying degrees, data are not always ratable or rating can be
systematically low if minimum data are provided. The possibility exists
that verbal expressiveness and test taking ability could affect ratings,
an argument Rest (1979) cited in reviewing Kohlberg's assessment in the
area of moral judgment. The capacity to objectively score data rests
upon the completeness of the rating manual, which in some cases allow
for considerable interpretation and potential scorer bias. These issues
and others related to measurement on the Perry scheme were included in
the topics of discussion at a June 1981 conference of proponents of the
scheme. Rating of data done by the Maryland, Alverno and Syracuse
groups at the conference revealed mixed correlations. The Maryland and
Alverno ratings correlated more highly with each other than did the
Syracuse ratings (Rodgers, 1982). Differences in the stimuli and techniques utilized by these groups are detailed in Chapter Two. Finally the current methods are unstructured to the point that content of thinking and justification of reasoning behind thinking may not be discernable from one another in the response. Should this be the case, the degree of interpretation necessary to assess the reasoning structures is increased. All of these concerns indicate the need for more structured written instrumentation. In addition, the concerns of extensive time to rate data, rater training and accuracy of rating still exist with current measures.

The significance of concerns related to assessment of intellectual development on the Perry scheme reaches beyond a scholarly concern and application issues. The widespread use of the scheme for facilitating student development from classroom instruction to career choice underscores the responsibility of researchers to work toward more standardized and objective methods of assessment. The present project is an attempt to move one step closer to more efficient and accurate measurement of intellectual development on the Perry scheme.
Perry's Scheme of Intellectual Development

Prior to exploring assessment of intellectual development on the Perry scheme, a more detailed description of the scheme is necessary. As noted in Chapter One, Perry's scheme shares the basic assumptions of cognitive developmental theory (Knefelkamp, Widick and Parker, 1978; Rodgers, 1980). The concept of cognitive structures is central to the Perry scheme. A structure is a pattern or filter through which an individual interprets the external world. This process has been described as a way of making meaning (Perry, 1970) and processing information via patterns for experiencing the world (Bieri, 1970). Additional characteristics of cognitive developmental theory focus on how change from one structure to another occurs. Movement is expected to be in the direction of increasingly differentiated and integrated structures. The structures are hierarchical in nature with each representing a qualitative change in thinking. Once the increased complexity has been woven into one's way of thinking it generally remains. Thus regression to less complex stages is rare. Cognitive developmental theory explains the stimulus for movement from one structure to another as external. When aspects of the environment are discrepant to the individual's way of thinking, cognitive conflict
(Rodgers, 1980) or disequilibrium is experienced. This conflict serves as the stimulus to reorganize the filter through which the external world is viewed in order to reduce the discrepancy. In cases where the discrepancy is small, the individual alters the structure to accommodate the experience and reduce the conflict. However, if the disequilibrium is too great, fixation in the current structure could occur (Maves, 1972).

Perry's scheme consists of nine positions, each representing a qualitatively different structure or way of thinking. The positions are categorized into three areas based on similarities in the organization of the structures. The scheme is outlined below.

**Dualism.** The first three positions are characterized by a dichotomous structure from which the world is viewed in absolute either-or terms. Truth is assumed to be known and information is processed to fall neatly into one of the two categories: right or wrong, good or bad, et cetera. The student perceiving the world from one of these structures views authority figures as the holders of truth from whom he must learn. Thus reliance on authority rather than one's own ideas is the order of the day in learning and decision-making.

**Position One.** Perry (1970) describes this position as a pure, closed structure. The world as seen from this structure is free of conflict since uncertainty is not perceived. Everything is absolute to the point of inability to imagine the existence of more than one answer to a question. Truth is accepted without question from the authorities, since they are always right.
Position Two. Diversity and alternate possibilities are recognized in this structure in a limited fashion. The truth still exists and is known to authorities. Two new twists appear however. Authorities introduce uncertainty as a way of helping students learn to find the truth on their own. Perry (1970) notes that this uncertainty is seen as "temporary, good for the mind, resolvable and therefore ultimately unreal" (p. 78). Learning to find one's own answers is dealt with through a focus on quantity and hard work. Complexity as such also represents the freedom to explore in the process of acquiring the truth. A second change is that some authorities do not seem to know the answers, and in accord with the dualistic structure are categorized as wrong.

Position Three. The major discovery in this position is that some truth is not known at present, even to authorities. Uncertainty is more uncomfortable since it cannot be immediately resolved by appeal to authority. Confusion over evaluation of work arises since authorities cannot be expected to know the answers in all areas. In the attempt to relieve anxiety the student focuses on quantity of work and holds tightly to the notion that the authority will lead the way to finding the truth in the future.

Relativism. Positions Four, Five and Six represent a view of knowledge as relative. The dualistic structure is discarded in favor of recognition of numerous possibilities. Uncertainty replaces absolutism as the order of the day. This change releases authority from its previous all-knowing position and elevates the student's ideas to an equal status. The search for absolute right answers is abandoned as the structure calls for a variety of answers dependent on the context.
Position Four. Position Four represents the beginning of the shift from certainty to uncertainty as the way reality is perceived. Some truth remains in limited areas. For the most part, there are no answers and no criteria with which to distinguish one opinion from another. Learning focuses on thinking about thinking, or practicing the way the authorities want us to think with their guidance.

Position Five. All knowledge is now viewed as relative and contextual. Opinions can be judged better or worse depending on the existence of supportive evidence. Authorities are redefined as those who have experience and competence in groping for the truth in a given context. Introspection and detachment are possible for the first time. The ability to evaluate and subsequently choose creates the fear of narrowing one's options.

Position Six. The acceptance of a truly relativistic world in which infinite contexts exist is accompanied by the realization that choosing is essential to avoid disorientation. One must affirm one's identity through establishing a context in which to invest oneself. The Position Six individual recognizes this need for commitment but has not yet acted upon it.

Commitment in Relativism. Positions Seven, Eight and Nine do not represent restructuring as do the previous positions. They are characterized by the development of commitment. The individual takes responsibility for making choices and affirming his or her identity in numerous contexts or areas of life. Perry notes that this process addresses both the content and style of identity.
Position Seven. Position Seven marks the point of initial commitment in some aspect of life such as career, relationships, or values. The commitment is an examined act, presupposing a genuine exploration of alternatives. Commitment is initially viewed as a settling act, restoring some order to one's life.

Position Eight. Position Eight is accompanied by the emergence of additional choices regarding how to carry out initial commitments. The major focus is on the style with which to implement the commitments as well as the issue of responsibility. The recognition of the infinite nature of choosing comes with the acceptance of relativism in its fullest sense.

Position Nine. This constitutes an integration of commitments and one's lifestyle as determined by implementation of commitments. Commitment is recognized as an ongoing activity, never to be settled in a final sense.

In addition to the main line of development Perry notes three alternatives to growth. Although evidence of these occurring is minimal, they warrant mention as possibilities. Temporizing involves a pause in growth for a period of a year or more. Perry describes these students as gathering their energy to be more ready for the next challenge. Temporizing could occur at any point along the scheme when students find themselves unable to take the next step. Retreat refers to the actual regression to an earlier position, often as a defense against the implications of more complex positions. Retreat is most likely to be back to dualism in order to maintain the needed structure for emotional control. Retreat is possible in more complex positions,
but escape is a more likely alternative. Escape involves an abandonment of responsibility through detachment. The advanced structures are more conducive to detachment by their very nature. Perry identifies two types of escape. The first, labelled dissociation, involves a rather passive delegation of all responsibility to fate. Rather than participating in exploring growth, the individual drifts along, waiting for fate to guide his or her identity. This type of escape is most likely to occur in the stage of Multiplicity. The second type, encapsulation, is more likely to occur in the position of Relativism due to the greater complexity of thinking at this level. This involves submersion in activity to avoid the implications of establishing one's values. The individual maintains an identity through doing rather than choosing an identity in a relativistic world.

Measurement

Loewinger and Wessler

Numerous studies of measurement of cognitive developmental theory are contained in the literature. The work of Loewinger and Wessler (1970) has emerged as one of the most comprehensive models of measurement and the prototype for many of the others. Their work focuses on ego development, a theory that entails many assumptions similar to those found in the Perry scheme. Ego development involves a progression through increasingly integrated stages that represent qualitatively different frames of reference. The central function of the ego is the search for meaning in experience, and movement occurs when external observations call for integration into the frame of reference. The ego as described by Loewinger and Wessler is relatively
stable since individuals selectively attend to experiences that are consistent with their frame of reference. This explanation approximates Perry's description of intellectual development in which an individual assimilates new information into his or her current view of the world until the discrepancy is too extensive to handle in this manner. Both theories share the qualities of the cognitive developmental family of theory and both focus on the task of making meaning. Ego development appears to be a broader conceptualization than intellectual development in its inclusion of the affective realm as well as the cognitive. Loevinger and Wessler comment that many strands of development occur simultaneously and that correlations between ego and intellectual development are almost certain particularly in the early years. The similarity in theoretical bases for these two theories legitimizes the use of Loevinger and Wessler's measurement methodology as a criterion against which to compare and further develop measurement on the Perry scheme.

Loevinger and Wessler (1970) believe that measurement of ego development can only occur through projective methods that allow the respondent to project his or her own frame of reference. Thus they developed the 32 item Sentence Completion Test. Their major contribution to measurement is in the area of manual construction. The method of manual construction and scoring paradigm they present serves as a model for measurement in areas other than ego development. The basic strategy espoused by Loevinger and Wessler involves the identification of qualitative differences in the successive stages of ego development. This results in the ability to match subject response
to the sequence of model criteria and examples in a rater's manual, and determine the level it most closely matches. The construction of the manual serves to identify the qualitative differences across the stages. The manual describes these differences for each sentence stem, increasing the degree of specificity and the probability of matching responses accurately. The evolution of the Loevinger and Wessler manual began with an initial version that contained examples of stem responses for the four ego levels that they originally used. A second manual was developed through grouping responses within a level with similar content into categories with representative titles. Since some content was found in more than one level, a third type of manual was constructed in which all differences among the categories were rationalized on the basis of the theory.

The construction of the manual was guided by three basic rules. Rating every response was seen as essential, in part due to minimizing the psychometric difficulties encountered in having varying numbers of responses. In addition, rating every response involved pushing the theorists insights to the limits so as not to verify the theory prematurely or lose sight of unexpected information. A second basic rule was to stay with the level of meaning presented by the respondent. Guessing what the respondent meant is acceptable to the extent that the inference is toward what the respondent would recognize as her meaning. Inferences beyond this are not acceptable. Inherent in this rule is the notion that the whole response must serve as the basis for the rating. The third rule emphasized use of plain English as opposed to theoretical jargon. Loevinger and Wessler emphasized that
the parameters of categories are arbitrary even though placement of categories in levels is done empirically. The final underlying assumption in this measurement methodology is that an individual functions, and can therefore be assigned, in one core level.

Loevinger and Wessler (1970) also present a paradigm for empirical verification of the scoring manual. Beginning with a categorized or rationalized manual, two raters work independently to rate a sample of data. Preparation of the data for rating includes typing a list of responses for each stem, removing them from the context of other stems and eliminating demographic data. The raters complete one stem for all protocols and discuss differences prior to proceeding to the next stem. Following this, total protocol ratings (TPRs) are done by three raters working independently. Only demographic data are withheld at this stage. When two raters agree and the third is within one half step, the rating on which the first two agree is automatically assigned. In all other cases the protocol is read and discussed in order to assign a compromise rating. Upon completion of rating each response is listed in a category as assigned by the raters along with the protocol number and TPR. This decoding report is a tool for examining the categories to determine necessary changes. This step is important in determining the best parameters and limits of categories. The manual revised by way of this method will eventually contain only observed categories. Theoretical categories in earlier versions will be confirmed by observation or eliminated. Loevinger and Wessler describe this on-going revision of both the manual and theory as a bootstrap operation that serves to create an increasingly defined form of
assessment. In discussing the methodological difficulties inherent in representing persons with TPR's, Loevinger and Wessler describe a method similar to the item-sum method, but free of the assumption of interval scale. Called ogive rules, it involves determining the frequency of individual item ratings to form a cumulative distribution. Cut off points designate ego levels which are assigned by comparing the protocol to the ogive.

Kohlberg and Gibbs

While Loevinger's work serves to guide the construction of rating manuals, work in the area of moral development provides further insight into instrument development and manual construction. Kohlberg's scheme of moral development evolved, like Perry's, from intensive interviews. Kohlberg's approach was more structured in that it involved presentation of moral dilemmas and an exploration of the subject's moral reasoning. Kohlberg and his colleagues at the Center for Moral Development at Harvard used data from twenty years of longitudinal study in order to construct the manual for the Moral Judgment Interview (Colby, Kohlberg, Candee, Gibbs, Kaufman, Hewer, Power & Speicher-Dubin, in press). During the course of analyzing the longitudinal data, Kohlberg and his colleagues determined that the earlier aspect scoring system was inadequate in identifying the structure of thinking separate from the content. A new scoring system was developed in order to identify the structure underlying a respondent's thinking in relation to moral issues. The Standard Issue Scoring (Kohlberg 1976, Kohlberg et al. in press) involves separating a response to a dilemma on the basis of the issues it addresses. Within each issue the rater identifies norms or
value clusters the respondent uses to support the chosen issue. Within each norm elements are identified that clarify the way the respondent uses the norm(s). Thus the scoring procedure is an attempt to determine the basic structure of the respondent's thinking. The way an element is used within a norm to support a moral issue is called the criterion judgment, and is the unit of analysis. The rating manual is an account of how norms and elements are utilized in various stages of moral development. The manual also explains stage structures and how they differ from one another. Thus the rater identifies the criterion judgment in a response and matches it to criterion judgments in the manual.

Gibbs' (Gibbs, Widaman & Colby, 1982) interest in alternatives for assessing moral judgment led to the development of the Sociomoral Reflection Measure (SRM). Like Kohlberg's Standard Issue Scoring, the SRM contains an emphasis on subject generated responses and justifications for reasoning. The SRM, however, is an attempt to further standardize the collection and evaluation of data. It incorporates methodological aspects of Kohlberg's Standard Issue Scoring, namely the standardization of issues included in a dilemma and the use of the criterion judgment as the unit of analysis. The standardization of issues within a dilemma eliminates variation on the basis of subjects or interviewers, thus focusing the data collection. Since the criterion judgment refers to the justification for reasoning as distinguished from the content of the response, it is the structure of reasoning that is being scored while the question provides the content to elicit the reasoning structure.
The SRM is a paper and pencil test. The questions, however, are presented in a manner that probes the respondent's thinking, much the same way as would occur in an interview, but in a more systematic fashion. Each question focuses the respondent's thinking on a particular norm or issue and proceeds through a series of questions. The series flows from calling for concrete action responses to more abstract reflection upon the issue. At the same time the probe questions serve to elicit the justification for reasoning. The probe questions are varied in such a way as to push the respondent to think through his or her justification and provide better data. This variation is included systematically in each question. In addition to providing justification for reasoning, the SRM also makes scoring easier. The focus on a particular norm eliminates the need for raters to classify responses by norm since the instrument does this. Scoring can begin at the point of stage assessment. Learning to score consequently becomes easier and can be accomplished without intensive workshop training. Gibbs notes that these advantages are tempered by limitations. The increased structure of the SRM focuses the respondent's frame of reference, reducing the probability of identifying new structures. He believes that sufficient research has been done to reasonably identify moral development structures. The SRM still requires training for raters since the production of responses by the subject is maintained. Gibbs comments that objective measures that do not require production of justification such as Rest's Defining Issues Test, measure reasoning in an evaluative rather than a structural sense. Gibbs and his colleagues are working on an objective test of sociomoral reasoning.
The reliability and validity of the SRM has been studied by Gibbs and his colleagues (Gibbs & Widaman, 1981; Gibbs, Widaman & Colby, 1982; Gibbs, Widaman, Colby & Fenton, 1981). Parallel form and test-retest reliability indicate that the SRM generates consistent results across time and form. Internal consistency among questions is high, as is interrater agreement. Comparison of the SRM to the Moral Judgment Interview serves as the initial test of concurrent validity, with results indicating 75% modal stage agreement. Construct validity has been assessed through correlation with age and education (correlations in the high .70's) and socioeconomic status (.23). Finally, two interventions using the SRM have shown that it discriminates differences in sociomoral development.

Knefelkamp

As was the case with ego and moral development, the need for paper and pencil measurement devices arose in relation to Perry's intellectual development scheme. The earliest instrument developed by Knefelkamp and Widick (1974) consisted of ten sentence stems and two essays. In Knefelkamp's later work, two changes occurred. A revised version of five stems and two essays was utilized in focusing on career development (Knefelkamp and Slepitza, 1978). The stems in the original instrument were then abandoned in favor of essays. Rating is done through the use of a manual consisting of explanations of each position along with examples of responses representative of each position. In addition a set of cues for each position assists in assigning responses to positions. Knefelkamp (1978) reports that the career version correlates highly with the interviews (.77). Knefelkamp's instrument is now titled
the Measure of Intellectual Development (MID). The rating process for the MID involves contextually applying criteria for the positions that have been developed by Knefelkamp and colleagues including Mentkowski (1981). Two raters work independently, then meet to reach consensus on protocols. Each essay receives a three digit rating, thus providing for stable as well as transitional positions (Moore, 1982). Although the Maryland group agrees that persons can be reasoning from different levels in different domains, they believe their measurement with the MID allows for accurate rating in the specific realms it addresses.

Validity data include a .51 correlation of Perry level measured by the MID with the Paragraph Completion Test (Widick, 1975) and a .45 correlation with scores on the Defining Issues Test (Meyer, 1977). MID data from various colleges indicate that differences have been observed in scores of freshmen and seniors (Moore, 1982). Freshmen have been reported to be between Positions Two and Three while seniors have been reported to be between Positions Three and Four. No differences have been reported on the basis of sex. Experimental studies in which the MID has been used in a pre-post manner to assess gains in Perry position due to developmental interventions have revealed gains in the predicted direction.

Studies of reliability of the MID have been conducted by Mentkowski at Alverno College and by Knefelkamp at Maryland. Moore (1982) reports that absolute agreement among Maryland raters is 57.9%, rising to 83.1% for dominant position. The Alverno teams rater agreement is reported as slightly lower, 43.4% and 74.4%, respectively.

Widick and Rodgers
Widick also continued work on the Perry Protocol, or the KneWi, as it came to be known at The Ohio State University in conjunction with Rodgers from 1974 to 1977. The Ohio State version eventually consisted of five stems and two essays from the original version. The rating manual for this version calls for first assessing structural cues to assign the response to a position. Secondly, the rater is to look for attitudinal or behavioral correlates that relate to the position. The response is then examined for cues in the style of language used including complexity, active versus passive voice and concrete versus abstract word usage. Finally the response is examined to identify qualitative differences in the use of concepts. The manual specifies the content of structural and behavioral correlate cues for each position, as well as statements common to the earlier positions. The quantification strategy requires the assignment of each stem response to one stage, that which is viewed by the rater as modal. A second stage may be rated as a subscript, but is not utilized in determining the final score for the protocol. Each essay is assigned three scores to represent potential combinations of positions present in the response. Rodgers (1979) suggests that a position ought to comprise at least 25% of the response before being included in the rating. Should the response be rated as representative of only one position, that number is assigned three times. Total protocol ratings are derived in two ways. One approach is to obtain the mean of the eleven numbers which result from the stems and essays. The second is to indicate the dominant overall stage by use of the position number and non-dominant stages by use of their number in parentheses. Dominant stage is defined as
representing 50% or more of the total response. Rodgers is currently reorganizing the manual by question. The specificity afforded by describing the developmental sequence in relation to each question in the instrument is expected to be of assistance in rating. Rodgers and Widick, as well as other researchers, have used this instrument in numerous studies conducted at California Polytechnical Institute, The Ohio State University, St. John's University, Kent State University and Kenyon College. The studies conducted at The Ohio State University (Rodgers, 1975-1979) revealed a relationship between level of education and developmental level. These studies reported most freshmen in dualistic positions and most graduate students in relativistic positions. Heidke and Omahan (1981) reported differences in developmental levels between freshmen and seniors. No sex differences were reported. Further reliability and validity studies on the instrument are as of yet unavailable.

Alverno

A version of the KneWi is in use at Alverno College. The version consists of three written essays: Best class, Decision, and Career. The instrument has been given to 737 students (Mentkowski, Moeser & Strait, 1981). The Alverno group has critiqued and refined cues from the work of Perry (1970), Knefelkamp (1978) and Knefelkamp and Slepitza (1978), resulting in 300 criteria for use in rating. The criteria were evaluated on a continuing basis similar to the bootstrapping utilized by Loevinger and Wessler (1970), Colby (1978) and Gibbs and Widaman (1981). Work has also been done in refining the rating process with the new criteria, improving interrater agreement and agreement with an
external expert assessor (in this case Knefelkamp). Interrater agreement correlations of 76% are reported as well as expert external assessor agreement of 67%. A full report, expected to be available in 1983, outlines the rating procedure and addresses issues the Alverno group encountered as a result of their studies.

**Syracuse**

The type of rating described in the Maryland, Ohio State and Alverno work is also used by the Syracuse Rating Group. The Syracuse Group utilizes a more flexible approach in regard to instrumentation. They define a protocol as a written instrument stimulus consisting of two or more essay questions (Zachary, 1982). They recommend that instrumentation of a sample occur under the same conditions, such as in a classroom. The Syracuse Rating Procedure rests on the use of cues believed to be common to Perry data. They have developed cue sheets from their experiences which are refined on a continuous basis. Ratings are derived through considered judgments, attempting to balance the theoretical framework and the cues. The Syracuse Group provides rating service for any data that conform to their definition of a protocol. Raters, the number required by the particular study, rate the protocols independently with the demographic data removed. Each essay receives three scores, a method noted earlier in the Maryland and Ohio State approaches. Upon completion of each question, the raters meet to discuss ratings, reach consensus, record ratings and discuss the subject's overall responses. Methods of consensus are determined by adding the three scores each rate assigned. If the totals differ by two points, the raters review their cues and rationale together in order to
agree. Larger differences result in rereading and rerating the essays independently. In cases where consensus cannot be reached, additional raters review the protocols and the total group decides on a consensus score.

A West Virginia University study (Allen, 1982) is utilizing the Syracuse Group Rating service to rate an instrument developed for the study. The study focuses on the intellectual development and the learning of science. The instrument includes two essays, the first on how to evaluate educational experience as worthwhile and the second calling for reaction to a hypothetical disagreement in a science class. The instructions emphasize encouraging respondents to be as complete as possible in order to increase the ratability of the data. Plans for validating the instrument with interviews are in progress.

Summary

In summary, current measures share a number of characteristics. The Maryland, Ohio State and Alverno approaches are most similar in their stimuli and rating manuals. They are also currently under revision to improve rating procedures and empirical verification of manuals. Table 1 summarizes the state of measurement on the Perry scheme in relation to criteria for rigor taken from the work of Loevinger and Wessler (1970), Gibbs and Widaman (1981) and Tamashiro (1981). As can be seen from the chart, none of the current approaches meet the criteria regarding stimuli. Empirical evidence exists as a result of the studies using these measures that indicates subjects omit some areas of thought when the questions are presented in essay form at the top of the page. The state of rating manuals is also less than ideal. The
Alverno group in conjunction with Knefelkamp seem to be making the most marked progress in this area. The revision underway at Ohio State will serve to enable the process of empirical verification to be initiated. Even with the improvements underway in the area of rating, stimuli remain inadequate. The Measure of Epistemological Reflection is an attempt to improve both the stimuli to elicit cognitive reasoning on the Perry scheme and the method by which it can be assigned to categories. At the present time these improvements seem essential to further utilization of the scheme.
### Table 1

Summary of Current Measures of the Perry Scheme

<table>
<thead>
<tr>
<th>Measurement Criteria</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maryland</td>
</tr>
<tr>
<td>Stimuli serve to separate content areas</td>
<td>no</td>
</tr>
<tr>
<td>Question format promotes response to each question</td>
<td>no</td>
</tr>
<tr>
<td>Stimuli directly elicit justification</td>
<td>dependent on respondent</td>
</tr>
<tr>
<td>Stimuli elicit adequate data for rating</td>
<td>dependent on respondent</td>
</tr>
<tr>
<td>Rating manual question specific</td>
<td>no*</td>
</tr>
<tr>
<td>Empirical verification of manual</td>
<td>yes</td>
</tr>
<tr>
<td>Specific, detailed rater instructions</td>
<td>no*</td>
</tr>
<tr>
<td>Reliability, validity studies</td>
<td>yes**</td>
</tr>
</tbody>
</table>

*under revision

**unpublished
CHAPTER THREE

Methods

Introduction

A number of methodological difficulties are inherent in the measurement of cognitive development. A few noted by Loevinger and Wessler (1970) provided guidance for the format of the instrument discussed here. The difficulties Loevinger and Wessler noted generally led them to recommend that researchers use as much precision and specificity as possible. They indicated that ego, moral, and intellectual development occur simultaneously, making the separation of these developmental phenomena difficult. They emphasized that error is bound to exist in matching behavioral examples to underlying development and distinguishing signs of one strand of development from correlated strands. The probability that subjects exhibit behaviors representative of more than one developmental level further complicates assessment. Finally, due to the nature of developmental theory, they indicated that signs that appear in early stages reappear in more complex form in later stages. All of these difficulties implied that at least three major considerations were essential to the development of the present instrument. First, clear and specific descriptions of the qualitative differences across Perry's positions of intellectual development are necessary. Secondly, the stimuli used in the instrument need to elicit these same specific qualitative differences in sufficient detail to
enlighten the particular strand of development in question, intellectual development in this case. Third, a comprehensive rating manual must accompany the instrument in order to increase the accuracy of distinguishing probable signs across the scheme.

Research Hypotheses

The purpose of the construction of the Measure of Epistemological Reflection (MER) was to minimize the methodological difficulties described above. The format of the instrument and rating manuals is based on research contributions described in Chapter Two. Subsequently the study of the development of the MER included exploration of hypotheses regarding rater consistency, the relationship of the MER to other measures of intellectual development, and the relationship of the MER ratings to level of education and sex. The major hypotheses of the current study were:

1. There will be a moderate to high correlation between the ratings of the two highly trained raters.
2. There will be a moderate to high correlation between the ratings of the two highly trained raters and the self-trained expert rater.
3. There will be a moderate correlation between the ratings of the two highly trained raters and the two self-trained raters.
4. Interrater reliability for the cross-validation sample will be higher than for the derivation sample.
5. The MER will exhibit a moderate degree of internal consistency across domains.
6. MER ratings will indicate differences between levels of education in the direction of increasing complexity on the Perry scheme.

7. MER ratings will not indicate differences on the basis of sex.

8. There will be a positive correlation between the MER and the Measure of Intellectual Development.

Construction of the Measure of Epistemological Reflection

The Perry scheme and research elaborating on the scheme collectively identify specific areas that seem to demonstrate underlying structures of thinking. Utilizing Gibbs (1981) norm framing concept, the instrument was designed to focus specifically on these structures, one per domain. These domains included choosing or decision-making, role of the learner, role of the instructor, role of peers, evaluation, and view of knowledge, truth or reality. The domains are explored in the context of the formal learning process. A second consideration of the instrument concerns an attempt to balance the degree of structure imposed by the stimuli and the requirement of the subject to project his or her own frame of reference or thinking structure in the responses. One of the concerns about available Perry measures is that their stimuli tend to leave the specificity and completeness of response entirely in the hands of the subject. This has led to incomplete responses in some protocols and large inferences in the rating process. Thus the current instrument focuses the respondent's thinking on a domain and requests elaboration through follow-up questions. This structure may increase the probability that the respondent will explore the domain in detail. As Gibbs indicates, this may serve to limit the discovery of new
structures. This may not be a major issue in this study, however, due to the scope of previous research on the Perry scheme. In addition, research on the scheme indicates that dualists may need this structure in order to formulate their responses, while relativistic thinkers can work within or around the structure. Self-generated response is the focus of the MER. This appears to be particularly important given the tendency for respondents to choose higher cognitive developmental stages on a recognition task than they could actually generate (Rest, 1973; Rest, Turiel, & Kohlberg, 1969). A fourth characteristic of the instrument is its focus on the role of reasoning used in responses rather than the content. The focus on justification or reasoning is central to cognitive developmental research (Gibbs & Widaman, 1981; Ginsburg and Opper, 1969; Kohlberg, 1976). As mentioned previously similar content may appear at various levels in qualitatively different form. Study of the Perry scheme would seem to imply that the differences are due to the reason respondents think the way they do rather than in what they think (King, 1977). The follow-up questions on the MER were thus designed to elicit reasoning or thinking structure of the respondents in each domain. In summary the MER format contains the following characteristics.

a) Exploration of six domains in the Perry scheme, each explored separately.

b) Each domain is explored through a series of follow-up probes, separated to increase the chance of response to each question.

c) The production aspect of response to domains and questions is maintained.
d) The follow-up probes elicit reasoning for the respondent's thinking.

A pilot study of the specific domains and questions in the instrument was conducted with a convenient sample of 32 Ohio State University students prior to proceeding with the present study. The pilot sample included subjects from both undergraduate and graduate levels of education. The sample consisted of 18 undergraduates enrolled in a career development course, 6 residence hall student staff and 8 graduate students in Student Personnel Work. Minor revisions in the stimuli were made on the basis of the review of these results. These data also served as the source of examples for the preliminary manual.

Development of the Coding Strategy and Preliminary Manual

The development of the rating procedure followed guidelines for planning coding strategies found in Tamashiro (1981) and Fox (1969). The unit of analysis was defined similar to Kohlberg's (1976) criterion judgment. The reasoning structures used here differ from criterion judgments in that they are not derived within the framework of norms, issues, and elements. Categories for assignment followed the original Perry scheme (1970) through Position Five. The rating process was designed to focus on Position One through Position Five since development beyond Position Five has yet to be very clearly defined by current research. The first five positions were perceived to be a reasonable portion of the Perry scheme on which to test the instrument and rating process. These positions are sufficiently clear in the research, and account for the cognitive structures described in the Perry scheme. Since relativism serves as the basic perspective
underlying higher positions in Perry's description of the scheme, Position Five was designated as a logical boundary for developing an initial rating process. However, the process of inductive derivation of new reasoning structures from data collected could assist in further clarification of development beyond Position Five.

The method of assigning responses to categories followed the Loevinger and Wessler's (1970) emphasis on manual specificity including rules regarding inferences and defining new categories. There are two important assumptions in the scoring algorithm. First it was assumed that within a given domain a subject could be assigned to one position. It seemed reasonable to expect that a subject might show evidence of opening toward a new position or coming from a previous one, however the subject's modal reasoning would be representative of only one position. As a result the scoring algorithm forced the assignment of one position for each domain. This did not eliminate the possibility of the subject receiving varying position ratings across domains.

The second assumption concerns the total protocol rating. Research to date does not indicate that certain domains related to the Perry scheme occur in any order. There is no evidence that one domain consistently serves as the entrance to a new position. Although development in one domain can be sparked by development in another, no pattern has been identified to date. Hence, weighting the ratings of domains was not necessary in deriving total protocol ratings. The total protocol rating represents a modal reasoning, but unlike the domain rating, may include more than one position.
These assumptions are based on the notion that one major perspective exists at any one time in the developmental scheme. Change begins to occur due to dissonance in one domain and as more domains change the individual goes through a transition to the next position. Thus the coding strategy used here allows for modal position and transition. The specific coding strategy is included in the preliminary and revised manuals (Appendices F & G).

Rating was done by at least two raters, working independently. For each sample rating was done by domain, that is raters completed one domain for all subjects before going on to the next domain. Discrepancies in rating by the two raters were discussed in order to determine unified or compromise ratings.

Construction of the rating manuals was based on the work of Loevinger and Wessler (1970) and Tamashiro (1981). Qualitative differences across the scheme were identified on the basis of the Perry theory. Pilot data were separated by position and reviewed to select the best examples for each position. The preliminary manual subsequently consisted of qualitative position descriptions accompanied by examples (see Appendix F). The manual was organized by domains such that a separate manual existed for each domain, including position descriptions and examples. This organization attempted to provide the degree of specificity recommended by Loevinger and Wessler (1970) in order to reduce rater inference in position assignment.

Sample

The sample was drawn from the population of students enrolled at The Ohio State University during Autumn Quarter 1982. The 1981-1982 Annual
Statistical Summary prepared by the Office of Registration Services at The Ohio State University provides a description of the student population. The undergraduate population was characterized by an approximate balance of men and women with a small percentage of students over age 25 or married. Undergraduate quarter point-hour ratios were reported to average 2.6 on a 4.0 scale. The Ohio State University has an open admissions policy for undergraduates. The majority of undergraduate students were reported to come from the state of Ohio. The Graduate School population was evenly balanced by sex with most students older than age 25 and approximately one third married. Students from the state of Ohio accounted for 75% of the enrollment. One eighth of the population came from foreign countries. The average point-hour ratio for this group was 3.5.

In order to elicit development across the Perry scheme, the groups in Table 2 were targeted for the sample. This combination of subjects allowed for analysis of group differences based on level of education and sex. Subjects in each group (except freshmen) were randomly selected from the student population at The Ohio State University. Freshmen were selected from students enrolled in an introductory psychology course that serves as one of a group of basic curricular requirements. The freshmen students enrolled in this course served as the population of freshmen. Interested students volunteered to participate through a sign-up process associated with the course. The remaining subjects were selected randomly through the Student Directory and contacted by mail. Solicitation materials are included in Appendix B. As noted in Table 2, response rate was low for the three groups.
### Table 2
Description of the Sample

<table>
<thead>
<tr>
<th>Level of Education</th>
<th>Subjects Solicited</th>
<th>Actual Respondents</th>
<th>Percentage of those Solicited</th>
<th>Sex</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>College Freshmen</td>
<td>73</td>
<td>50</td>
<td>68%</td>
<td>22</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>College Seniors</td>
<td>124</td>
<td>23</td>
<td>18.5%</td>
<td>16</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Graduate students under age 30</td>
<td>142</td>
<td>43</td>
<td>30.3%</td>
<td>24</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Graduate students over age 30</td>
<td>114</td>
<td>39</td>
<td>34.2%</td>
<td>17</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>453</td>
<td>155</td>
<td>34.2%</td>
<td>79</td>
<td>76</td>
<td></td>
</tr>
</tbody>
</table>
solicited through the mail. Non-respondents in these groups were evenly distributed by sex and across levels of education. Included in the non-respondent category were subjects who did not respond at all, subjects who indicated they did not wish to participate, and those who agreed to participate but did not. The number of seniors who agreed to participate but did not complete the instruments was higher than that of the two graduate groups. Nearly half of the non-respondents indicated on the response card that they were too busy to participate. The distribution of non-respondents revealed no information to suggest that these students differed from those who chose to participate.

Description of the sample on the basis of socioeconomic status appears in Table 3. Table 3 also contains the average Duncan Socioeconomic Index (SEI) score for each group. The Duncan scale was originally constructed on the basis of 1950 U. S. Census data (Mueller & Parcel 1981) but was updated to the 1970 Census data by Featherman and Hauser (1981). The average scores in Table 3 indicate that subjects in this sample are fairly homogeneous, with the majority in the first two occupation categories. Given this information, socioeconomic status was not considered to be a significant variable in the study.

The 155 subjects were randomly divided in half. One group (the derivation sample) was used in manual construction and revision. The remaining half was utilized for cross validation (Tamashiro, 1981). Reliability measures were computed for both parts of the sample, each requiring 100 units of data (Fox, 1969). Since each instrument contains six units of data, sufficient units of data were available for reliability assessment.
### Table 3

**Socioeconomic Status* by Level of Education**

<table>
<thead>
<tr>
<th>Major Occupation Categories</th>
<th>Level of Education</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freshmen</td>
<td>Seniors</td>
<td>Graduates under 30</td>
<td>Graduates over 30</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Profession &amp; Technical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managers &amp; Administrators</td>
<td>13</td>
<td>12</td>
<td>31</td>
<td>26</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Sales Workers</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Clerical Workers</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Craftsmen</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Operatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Except Transport Laborers</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Except Farm</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Farmers &amp; Farm Managers</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Farm Laborers &amp; Foremen</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Service Workers Exc. Private</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Duncan SEI Average Score</td>
<td>61.1</td>
<td>69.3</td>
<td>68.4</td>
<td>68.4</td>
<td>66.8</td>
<td></td>
</tr>
</tbody>
</table>

Administration of Instrumentation

Subjects completed the Measure of Epistemological Reflection and the Measure of Intellectual Development (for construct validity purposes). Administration of the instruments to freshmen subjects occurred in a group setting, while the remaining subjects received and returned the instrument package in the mail (see Appendix C). The time required for administration varied, dependent on the extent of each subject's responses. The range of completion time of the MER was one to two hours. The MID took approximately forty-five minutes to an hour for completion. The order of the instruments was altered for half the subjects in order to avoid any bias that might result from consistently taking one instrument last. In addition subjects completed a demographic sheet based on the Duncan scale (see Appendix D). This information was collected to determine whether socioeconomic status should be considered as a variable in developmental levels of subjects.

Revision of the Preliminary Manual

There were eighty-four subjects who served as the derivation sample. This group included 25 freshmen, 11 seniors, 25 graduate students under age 30, and 23 graduate students over age 30. The derivation sample was rated using the preliminary manual. The following steps guided rating of the derivation sample (Tamashiro, 1981).

a) Prior to rating the derivation sample, rater agreement was established on preliminary coding of Perry position descriptions. Revisions as necessary were made prior to administration of the MER to the sample.
b) During rating of the derivation sample new reasoning structures were identified and modifications to the position descriptions were made as necessary based on the data in the derivation sample.

c) Rater agreement on new reasoning structures and revised position descriptions was established.

d) Responses were assigned to agreed upon positions and recorded.

The above method is Tamashiro's version of the Loevinger and Wessler empirical verification process for developing cognitive developmental instruments and manuals. The process serves to continually refine the rating process and the theory if necessary on the basis of the empirical data. Thus the rating of the derivation sample served as the basis for revision of the preliminary manual. The revised manual was later used in rating the cross-validation sample.

Upon completion of rating the derivation sample as described, domain responses were separated on the basis of the position assigned. Data within each assigned position were reviewed to identify categories of response based on reasoning structure. For each position in each domain, reasoning structures were identified, along with appropriate examples, and both were used to revise the preliminary manual. Generally, this process resulted in maintaining the original position descriptions while adding specific reasoning structures and examples. The manual excerpts below illustrate the contrast between the preliminary and revised manual.

Preliminary Manual

Position Two
Learning is still guided by the search for the right answer but sometimes authorities want students to work to find the answer themselves. The role of the learner is to work hard at doing what the teacher wants. It is at this position also that the notion of some authorities being poor authorities appears. Those who do not seem to know the truth are labeled as dumb authorities. This is seen as an exercise that is good for the mind and is not a genuine search for the truth since authorities already know the truth.

2.2.1 I would tell them to pay good attention in class, take notes (a lot of notes), ask questions when something is not clear and go to class everyday.

2.2.2 The key to doing well in college courses is to stay current with school work and understand every why, what, where and how of everything that's said in a college course.

Revised Manual

Position Two

Learning is still guided by the search for the right answer and students prefer the most direct route to getting the answers. The role remains one of acquiring the truth. However, sometimes authorities want students to work to find the answer themselves. In this case the role of the learner is to work hard at doing what the teacher wants. This is seen as an exercise that is good for the mind and is not a genuine search for the truth since authorities already know the truth. It is at this position also that the student recognizes that some authorities do not seem to know the answers. They are simply labeled as "dumb" authorities.

Reasoning Structures

2.2.1 Preferred type of learning is the method that is clearly right and is the easiest way to obtain the answers.

Example: I learn best in classes which focus on factual information because I can remember things better as they are real. Concepts and ideas are hard for me to remember. The advantages are that I can look back to the information if I need to. There are no disadvantages.

2.2.2 Preferred type of learning is one that authorities suggest is good for you.

Example: Classes that deal with ideas and concepts. Because factual information bores me and when learning ideas and concepts you have to think. Learning ideas and concepts help stretch your mind. Factual information can be retained by a computer but ideas and concepts can only be understood by humans. Factual information is a necessity to everyday life and sometimes we need to understand it before we can understand ideas and concepts.
The cross-validation sample, consisting of the remaining 71 subjects in the total sample, was rated repeating the steps mentioned previously. The raters used the revised manual (Appendix G) to determine positions on the developmental scheme. Data from the derivation sample supported the notion that content and structure are not always indicative of the same position. Thus careful analysis is needed in developing manuals and the rating process. For example, different contents can have the same underlying structure. In the data for the role of peers question, some responses indicated a content preference for "a lot of talking" or participation by students while others noted a preference for "very little talking" by students. However, the reasons behind these opposite preferences were often the same. Those who preferred little student talking at Position Two commented on giving the teacher time to give the answers. Those who preferred student talking at Position Two described "talking" as asking questions in order to get the answers from the teacher.

In addition to different contents having the same thinking structure, the same content can have different underlying structures. Again looking at the role of peers question, some of the respondents who preferred "students talking in class" made meaning of participation in different ways. Respondents in earlier positions describe talking as asking questions or keeping the class from getting boring. Respondents in the later positions describe talking as sharing ideas to broaden or narrow perspectives. Thus a focus on and the identification of reasoning structures in the data of the study helped to more clearly
define the qualitative differences across positions and to provide examples for the revised rating manual.

**Raters**

There were three categories of raters included in this study. The two authors of the instrument served as highly trained raters of the data, due to their involvement in the instrument's development and manual construction. An expert in the area of Perry rating was also included in the study. After self-training the expert rater rated the derivation sample with the preliminary manual and the cross-validation sample with the revised manual. The inclusion of an expert allowed for comparison of ratings between manual constructors and an expert not involved in the manual construction. Finally, two additional raters, knowledgeable in the Perry scheme, were asked to engage in self-training through study of the revised manual prior to rating the cross-validation sample. The inclusion of self-trained raters was intended to examine the degree of consistency in coding that can be achieved on the basis of self study of the manual without additional training.

**Reliability**

Reliability of the Measure of Epistemological Reflection and coding method was approached through an evaluation of internal consistency of the MER and interrater agreement. Internal consistency focused on the extent to which domain ratings account for or vary from the modal position of the entire protocol. Interrater reliability was utilized to determine whether the coding strategy was used consistently across raters. Interrater reliability in this study included:
a) comparisons of percentage agreement on domain scores, treated as ordinal data, among all combinations of raters;
b) percentage agreement for total protocol ratings, treated as ordinal data among all combinations of raters; and
c) correlations of total protocol ratings, transformed to continuous data, for all combinations of raters.

Percentage agreement in the above cases was calculated for same position ratings and ratings one position apart.

Validity

Although extensive validity studies were beyond the scope of the present research, establishing preliminary validity was seen as essential to the introduction of the MER. It seemed reasonable to establish the face validity of the MER through comparing its domains with the current body of literature on the Perry scheme. The domains included in the MER were common to the literature, including Perry's original work. Although Perry's work was descriptive rather than prescriptive in nature, he addressed in his descriptions all of the domains included here. As noted in Chapter Two, Kurfiss included decision-making (moral), evaluation, role of instructor and view of knowledge in her work. Reference to the same constructs are found in the work of Clinchy and Zimmerman (1975). The work of Knefelkamp and Widick, and subsequent extensions of this work at the University of Maryland and Ohio State University, included these constructs and the role of peers and the learner in the learning process. Thus the inclusion of the six domains in the MER did not constitute new material, but was rather a new design that assists in separating responses by
domains. The degree to which these same domains had been explored in relation to the Perry scheme implied that the current instrument would measure development on the Perry scheme. Similarly the face validity of the preliminary coding manual was established on the basis of the theory and subsequent research. The literature provided descriptive data from which to determine distinctions between levels of reasoning. Rationale for distinctions and subsequently the content of the preliminary coding manual rested on the theory, while revisions to the manual were based on the empirical data in the derivation and cross-validation sample.

Concurrent validity was examined in two ways. It was expected that a relationship existed between class rank, or level of education, and reasoning level. The original Perry scheme indicated that increasing complexity occurs over the college years. This phenomenon has been observed in most cognitive developmental research on the college years. Research studies using the KneWi conducted at The Ohio State University (Rodgers, 1975-1979) indicated a relationship between class rank and developmental level. Rodgers reported that 83% to 89% of the first quarter freshmen in his studies were rated in dualism, while 1% to 2% were rated relativistic, with the remainder in transition between Position Three or Four. He reported that first year graduate students fell into the following categories: 15% dualism, 10% transition between Position Three and Four, 60% relativistic and 15% commitment. Heidke and Omaha (1981), also using the KneWi, found a significant difference between freshmen and seniors, with freshmen tending to be rated at Position Two and seniors in transition between Positions Three and Four. Thus the relationship of ratings on the current instrument with
class rank were examined to assist in establishing the validity of the MER. An analysis of variance and Scheffe's method of post hoc multiple comparisons were utilized to examine these relationships.

The second method of examining concurrent validity involved comparing the results on the MER with the results on other measures of development on the Perry scheme using the same subjects. A difficulty existed here in that established measures differ in two ways from the MER. The degree of specificity and structure of stimuli are greater in the MER than in those previously established. The rating of position on the MER is also focused solely on the rationale for thinking whereas others may be more susceptible to content to varying degrees. As a result of these differences, correlations among the MER and established instruments might be expected to be confounded. In spite of this difficulty, such correlations could assist in the establishment of validity of the MER. One of the most rigorous works in the area of instrumentation has been done by Knefelkamp and her colleagues. Thus the Measure of Intellectual Development (MID) served as the measure which was selected to compare with the MER. As previously mentioned the subjects were given both instruments at the same time. MID protocol rating was conducted by Knefelkamp's trained raters at the Center for Developmental Instruction at The University of Maryland. Correlations between ratings of the MID and ratings of the MER were explored to determine the extent to which the MER and MID represent subjects at the same position.

Limitations
Inherent limitations existed in the areas of solicitation of subjects, data collection and time constraints. Selection of freshmen subjects differed from that of the remaining groups. Although the freshmen sample was viewed as reasonably representative of the freshmen population, it was not a random sample of the entire population. A second difference between the freshmen and remaining groups was the fact that freshmen received class credit for participation in the study. This type of incentive did not exist for the remaining groups. The solicitation of seniors and graduate students through the mail yielded low return rates, and consequently smaller samples than anticipated. This was particularly true in the case of seniors. Time constraints of the study necessitated a limit to further solicitation. Finally, the instruments were group administered to the freshmen but mailed to the remaining groups. Thus control of environmental factors and opportunity for verbal instructions and response to questions was lost.

The rating process focused on Positions One through Five, thus limiting clarification of potential differences in higher reasoning levels. Lack of specificity of these higher levels in current research precluded the ability to define them in the same degree of specificity as the first five levels.
CHAPTER FOUR

Results

The Measure of Epistemological Reflection was created in order to try to provide a more reliable and valid method of data collection and evaluation. As a result interrater reliability was of primary interest in assessing whether the MER accomplishes its purpose. Interrater reliability in the study was explored in terms of domain and total protocol ratings as well as between various combinations of raters. Interrater reliability for the most part is reported by sample since rating was conducted in two phases utilizing the preliminary and then the revised manual.

Interrater Reliability: Domain Ratings

The most rigorous assessment of interrater reliability lies in comparison of domain ratings. Domain ratings consist of one position assigned per domain, the position being determined as the modal reasoning structure in the responses to the questions within each domain. As such, domain ratings constitute ordinal data and were compared through percentage agreement. Percentage agreement is reported for exact and within one position agreement between raters. Percentage agreement by domains for raters one and two, the highly trained raters, are reported in Table 4. Exact agreement for the derivation sample ranged from 41.7% to 64.3%. Domain one (decision-making) agreement was the lowest of the six and the only question showing agreement under
### Table 4
Percentage Agreement of Highly Trained Raters Domain Ratings by Sample

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Derivation Sample</strong></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>41.7</td>
</tr>
<tr>
<td>Within One Position</td>
<td>21.4</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>63.1</td>
</tr>
<tr>
<td><strong>Cross Validation Sample</strong></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>54.9</td>
</tr>
<tr>
<td>Within One Position</td>
<td>23.9</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>78.8</td>
</tr>
</tbody>
</table>
50%. Cumulative agreement, including those ratings discrepant by only one position, reveals agreement in the high .80's or above with the exception of domain one.

The range of exact agreement on the cross validation sample was similar to that of the derivation sample, with the exception that domain two (role of the learner) represented the lowest agreement. Exact agreement increased in the cross validation sample for domains one, four and five, and decreased for domains two, three and six. Cumulative agreement including the within one position discrepancies showed slightly higher agreement in general. Only domains three and four indicated slight decreases.

The highly trained raters compromised ratings upon completion of the rating of each sample. These compromise ratings were used to compare interrater reliability with rater three, the self-trained expert. Percentage agreement by domains for the compromise ratings and rater three are reported in Table 5. Exact agreement for the derivation sample ranged from 48.8% to 59.5% with domain four being the only domain below 50%. Cumulative agreement was generally in the high .80's or above with the exception of question one, as was the case in comparing the highly trained raters earlier.

The range of exact agreement on the cross validation sample was wider than that of the derivation sample. Domain one showed the least agreement with 46.5%. Domains three, five and six all indicated 52.1%, and domains two and four revealed agreement in the low .60's. Again exact agreements improved over the derivation sample in three domains: two, three and four. The remaining domains showed slight decreases.
### Table 5

Percentage Agreement of Highly Trained Raters' Compromise Domain Ratings and Self-Trained Expert Domain Ratings by Sample

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Domains</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td><strong>Derivation Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>50.0</td>
<td>50.0</td>
<td>51.2</td>
<td>48.8</td>
<td>59.5</td>
<td>59.5</td>
</tr>
<tr>
<td>Within One Position</td>
<td>27.4</td>
<td>41.7</td>
<td>38.1</td>
<td>38.1</td>
<td>27.4</td>
<td>31.0</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>77.4</td>
<td>91.7</td>
<td>89.3</td>
<td>86.9</td>
<td>86.9</td>
<td>90.5</td>
</tr>
<tr>
<td><strong>Cross Validation Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>46.5</td>
<td>62.0</td>
<td>52.1</td>
<td>63.4</td>
<td>52.1</td>
<td>52.1</td>
</tr>
<tr>
<td>Within One Position</td>
<td>26.8</td>
<td>29.6</td>
<td>39.4</td>
<td>25.4</td>
<td>35.2</td>
<td>38.0</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>73.3</td>
<td>91.6</td>
<td>91.5</td>
<td>88.8</td>
<td>87.3</td>
<td>90.1</td>
</tr>
</tbody>
</table>
Cumulative agreement for the cross-validation sample was quite similar to that of the derivation sample, with agreement decreasing noticeably only in domain one.

The percentage agreement of domain ratings between the self-trained expert and each highly trained rater is reported in Table 6 and Table 7. Examination of these tables reveals that both exact and cumulative agreement is similar between the self-trained expert and each highly trained rater, although rater two's agreement is slightly higher in general. Comparison of Tables 6 and 7 to Table 5 indicates that compromising the ratings of the highly trained raters only slightly altered the percentage agreements in all categories. Compromising increased the agreement in the derivation sample on domains one and five and in the cross-validation sample in domains two, four and six.

The self-trained raters rated only the cross validation sample since the preliminary manual was not perceived to be specific enough to test consistency achieved through self-training. The percentage agreement for raters four and five, the self-trained raters, is reported in Table 8. Exact agreement ranged from 53.5% to 64.8% with domains three and four indicating the lowest agreement. Cumulative agreement was 90% or above with the exception of domain one which showed 85.9%.

To determine the degree of consistency of each self-trained rater's performance each was compared to the compromise rating of the highly trained raters. These percentage agreements are shown in Table 9. Comparison of raters four and five indicates that rater five obtained less agreement with the highly trained raters across all domains on both exact and cumulative agreement. Rater four obtained exact and
<table>
<thead>
<tr>
<th>Agreement</th>
<th>Domains</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td><strong>Derivation Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>44.0</td>
<td>50.0</td>
<td>52.4</td>
<td>46.4</td>
<td>53.6</td>
<td>59.5</td>
</tr>
<tr>
<td>Within One Position</td>
<td>29.8</td>
<td>38.1</td>
<td>35.8</td>
<td>36.9</td>
<td>28.6</td>
<td>32.1</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>73.8</td>
<td>88.1</td>
<td>88.2</td>
<td>83.3</td>
<td>82.2</td>
<td>91.6</td>
</tr>
<tr>
<td><strong>Cross Validation Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>43.7</td>
<td>54.9</td>
<td>52.1</td>
<td>60.6</td>
<td>52.1</td>
<td>47.9</td>
</tr>
<tr>
<td>Within One Position</td>
<td>29.6</td>
<td>36.6</td>
<td>38.0</td>
<td>26.8</td>
<td>33.8</td>
<td>42.3</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>73.3</td>
<td>91.5</td>
<td>90.1</td>
<td>87.4</td>
<td>85.9</td>
<td>90.2</td>
</tr>
</tbody>
</table>
### Table 7
Percentage Agreement of Rater Two Domain Ratings and Self-Trained Expert Domain Ratings

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Derivation Sample</strong></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>45.2</td>
</tr>
<tr>
<td>Within One Position</td>
<td>22.6</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>67.8</td>
</tr>
<tr>
<td><strong>Cross Validation Sample</strong></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>46.5</td>
</tr>
<tr>
<td>Within One Position</td>
<td>28.2</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>74.7</td>
</tr>
</tbody>
</table>
### Table 8
Percentage Agreement of Self-Trained Raters Domain Ratings

<table>
<thead>
<tr>
<th>Agreement</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Validation Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>54.9</td>
<td>64.8</td>
<td>53.5</td>
<td>57.7</td>
<td>56.3</td>
<td>60.6</td>
</tr>
<tr>
<td>Within One Position</td>
<td>31.0</td>
<td>28.2</td>
<td>36.7</td>
<td>36.6</td>
<td>38.0</td>
<td>29.6</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>85.9</td>
<td>93.0</td>
<td>90.2</td>
<td>94.3</td>
<td>94.3</td>
<td>90.2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreement</td>
<td>Domains</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Rater Four</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>45.1</td>
<td>62.0</td>
<td>52.1</td>
<td>54.9</td>
<td>56.3</td>
<td>62.0</td>
</tr>
<tr>
<td>Within One Position</td>
<td>22.5</td>
<td>25.4</td>
<td>39.4</td>
<td>35.2</td>
<td>33.8</td>
<td>31.0</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>67.6</td>
<td>87.4</td>
<td>91.5</td>
<td>90.1</td>
<td>90.1</td>
<td>93.0</td>
</tr>
<tr>
<td>Rater Five</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>38.0</td>
<td>59.2</td>
<td>46.5</td>
<td>47.9</td>
<td>38.0</td>
<td>47.9</td>
</tr>
<tr>
<td>Within One Position</td>
<td>28.2</td>
<td>26.8</td>
<td>40.8</td>
<td>32.4</td>
<td>45.1</td>
<td>40.8</td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>66.2</td>
<td>86.0</td>
<td>87.3</td>
<td>80.3</td>
<td>83.1</td>
<td>88.7</td>
</tr>
</tbody>
</table>
cumulative agreement in the same range as did the self-trained expert with the highly trained raters.

**Interrater reliability: Total Protocol Ratings**

Total protocol ratings (TPR's) were compared in two ways. First an interval data TPR was determined in order to calculate Pearson product-moment correlation coefficients. These TPR's were determined by adding the domain ratings within a protocol and dividing by the number of domain ratings to yield an average score. Later these TPR's were rounded to the nearest whole number in order to calculate percentage agreement on modal overall position.

Comparison of TPR's as average scores is presented in Table 10. The correlation coefficients ranged from .75 to .84, all obtaining significance at p < .0001. Correlations were slightly higher for the cross-validation sample with the exception of the two highly trained raters. The highest correlation in the derivation sample was between the highly trained raters. The highest correlation in the cross-validation sample was between the highly trained raters' compromise TPR and the self-trained expert.

Comparisons with the self-trained raters are noted in Table 11. As previously mentioned the self-trained raters rated only the cross-validation sample. As was the case with domain ratings, rater four's correlations with the highly trained raters and the self-trained expert were higher than those of rater five. However, all correlations were significant at p < .0001. Rater four's correlation with the highly trained raters' compromise (.80) was similar to that achieved by the self-trained expert and the highly trained raters' compromise (.84) for the same sample.
Table 10
Pearson Product-Moment Correlation Coefficients for Total Protocol Ratings of Highly Trained Raters and Self-Trained Expert by Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th>Raters</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 and 2</td>
<td>1 and 3</td>
<td>2 and 3</td>
<td>1 &amp; 2 Compromise and 3</td>
</tr>
<tr>
<td>Derivation</td>
<td>.82</td>
<td>.77</td>
<td>.75</td>
<td>.79</td>
</tr>
<tr>
<td>Cross Validation</td>
<td>.78</td>
<td>.82</td>
<td>.81</td>
<td>.84</td>
</tr>
<tr>
<td>p &lt; .0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11
Correlations of Total Protocol Ratings of Self-Trained with Highly Trained and Self-Trained Expert Raters

<table>
<thead>
<tr>
<th>Raters</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>1 and 2 Compromise</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four</td>
<td>.77</td>
<td>.73</td>
<td>.73</td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td>Five</td>
<td>.65</td>
<td>.70</td>
<td>.70</td>
<td></td>
<td>.67</td>
</tr>
</tbody>
</table>

p < .0001
Rounded total protocol ratings were obtained in order to assess modal position agreement among raters. Percentage agreement among highly trained raters and the self-trained expert is reported in Table 12. Exact agreement between highly trained raters was 59.5% for the derivation sample and increased to 73.2% for the cross-validation sample. However, exact agreement between the highly trained raters' compromise and the self-trained expert remained approximately the same in the cross-validation sample. Cumulative agreement indicated that ratings discrepant by more than one position were rare. Table 13 contains percentage agreement on modal TPR's for the self-trained raters. Rater four showed higher agreement with both the highly trained raters and the self-trained expert. Cumulative agreement again showed very little discrepancy more than one position apart. In comparing Table 13 with Table 12 it should be noted that rater four achieved higher agreement (64.8%) with the highly trained raters' compromise TPR's than did the self-trained expert (63.4%).

**Internal Consistency**

Internal consistency of the Measure of Epistemological Reflection was assessed in three ways. All tests of internal consistency were based on the compromise ratings of the highly trained raters. Further, since internal consistency relates to the instrument rather than the rating process, internal consistency assessment was done for the entire sample rather than separating the derivation and cross-validation sample. First, Cronbach's alpha coefficient was computed to determine to what extent domain scores represented the overall scores for all subjects combined. Second, percentage agreement between domain and
Table 12
Percentage Agreement on Modal TPR's for Highly Trained Raters and Self-Trained Expert

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Raters</th>
<th>1 and 2</th>
<th>1 and 3</th>
<th>2 and 3</th>
<th>1 &amp; 2 Compromise and 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Derivation Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>59.5</td>
<td>59.5</td>
<td>51.2</td>
<td>61.9</td>
<td></td>
</tr>
<tr>
<td>Within One Position</td>
<td>36.9</td>
<td>39.3</td>
<td>47.6</td>
<td>38.1</td>
<td></td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>96.4</td>
<td>98.8</td>
<td>98.8</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cross Validation Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>73.2</td>
<td>59.2</td>
<td>74.7</td>
<td>63.4</td>
<td></td>
</tr>
<tr>
<td>Within One Position</td>
<td>26.6</td>
<td>38.0</td>
<td>23.5</td>
<td>35.2</td>
<td></td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>99.8</td>
<td>97.2</td>
<td>98.2</td>
<td>98.6</td>
<td></td>
</tr>
</tbody>
</table>
Table 13

Percentage Agreement on Modal TPR's for Self-Trained Raters with Highly Trained and Self-Trained Expert Raters

<table>
<thead>
<tr>
<th>Agreement</th>
<th>4 and 1 &amp; 2</th>
<th>4 and 3</th>
<th>5 and 1 &amp; 2</th>
<th>5 and 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Compromise</td>
<td></td>
<td>Compromise</td>
<td></td>
</tr>
<tr>
<td>Exact</td>
<td>64.8</td>
<td>62.0</td>
<td>43.7</td>
<td>54.9</td>
</tr>
<tr>
<td>Within One</td>
<td>33.8</td>
<td>38.0</td>
<td>52.1</td>
<td>42.3</td>
</tr>
<tr>
<td>Position</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Total</td>
<td>98.6</td>
<td>100.0</td>
<td>95.8</td>
<td>97.2</td>
</tr>
</tbody>
</table>
total protocol ratings was calculated to determine how scores in each domain compared to the total protocol scores by person. Correlations were not possible since domain ratings are ordinal data. Thus total protocols were rounded to form ordinal data for the purpose of calculating percentage agreement. Finally, means, standard deviations and variances overall and by level of education were computed.

Cronbach's alpha measures internal consistency on the basis of the variance of domain scores in relation to the variance of overall scores. The alpha coefficient for the entire sample was .76, indicating a moderately high degree of internal consistency. Alpha coefficients by level of education were also computed and are reported in Table 14. Since a smaller range of scores exists by level the alpha coefficient is smaller, perhaps underestimating the degree of consistency. Thus the overall alpha is the best estimate of internal consistency.

Table 14
Cronbach's Alpha Coefficients by Level of Education

<table>
<thead>
<tr>
<th>Level of Education</th>
<th>N</th>
<th>Alpha Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshmen</td>
<td>50</td>
<td>.44</td>
</tr>
<tr>
<td>Seniors</td>
<td>23</td>
<td>.61</td>
</tr>
<tr>
<td>Graduates under 30</td>
<td>43</td>
<td>.57</td>
</tr>
<tr>
<td>Graduates over 30</td>
<td>39</td>
<td>.74</td>
</tr>
<tr>
<td>Total Sample</td>
<td>155</td>
<td>.76</td>
</tr>
</tbody>
</table>
Percentage agreement between domain scores and total protocol scores are reported in Table 15. Domain one shows the least exact agreement and the least cumulative agreement within one position. This is the same domain that evidenced the lowest interrater reliability. Exact agreement for the remaining domains was mid .50's or higher with domain two having the highest exact agreement. The information in Table 14 would indicate that moderate levels of exact agreement exist between the domains and the overall scores.

Table 16 contains means, standard deviations and variance across domains for each of the four levels of education. As can be seen from the table, the range of means within levels was fairly small. The range of freshmen means was 2.06 to 2.46. For seniors the range was 3.00 to 3.43, although domain two was the only mean above 3.09. The range increased for graduate students under age 30, going from 2.95 to 3.93. A somewhat narrower range existed for graduate students over age 30, indicating 3.23 to 4.08. The ranges of means appear to increase in value with increased level of education with the exception of graduate students under age 30 whose range overlapped with the seniors and older graduate group. Deviation from the mean also appears to increase with level of education. Comparison of means, standard deviations and variance across domains however shows a moderate degree of consistency. Overall domain one appears to show the least consistency, particularly in the senior and over 30 graduate level. Looking at means by domain, a general increase in scores occurs with increased level of education.

Validity
Table 15
Percentage Agreement of Domain and Total Protocol Scores

<table>
<thead>
<tr>
<th>Domains</th>
<th>N</th>
<th>Exact Agreement</th>
<th>Within One Position Agreement</th>
<th>Cumulative Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>141</td>
<td>48.2</td>
<td>45.9</td>
<td>93.6</td>
</tr>
<tr>
<td>Two</td>
<td>149</td>
<td>63.8</td>
<td>34.9</td>
<td>98.7</td>
</tr>
<tr>
<td>Three</td>
<td>148</td>
<td>59.5</td>
<td>39.2</td>
<td>98.7</td>
</tr>
<tr>
<td>Four</td>
<td>141</td>
<td>56.7</td>
<td>39.7</td>
<td>96.4</td>
</tr>
<tr>
<td>Five</td>
<td>145</td>
<td>57.9</td>
<td>39.3</td>
<td>97.2</td>
</tr>
<tr>
<td>Six</td>
<td>153</td>
<td>56.2</td>
<td>41.8</td>
<td>98.0</td>
</tr>
</tbody>
</table>
Table 16
Means, Standard Deviations and Variance Across Domains by Level of Education

<table>
<thead>
<tr>
<th>Domains</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Freshmen</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>2.06</td>
<td>2.46</td>
<td>2.36</td>
<td>2.44</td>
<td>2.44</td>
<td>2.62</td>
</tr>
<tr>
<td>SD</td>
<td>.89</td>
<td>.79</td>
<td>.72</td>
<td>.91</td>
<td>.99</td>
<td>.97</td>
</tr>
<tr>
<td>Variance</td>
<td>.79</td>
<td>.62</td>
<td>.52</td>
<td>.82</td>
<td>:</td>
<td>.93</td>
</tr>
<tr>
<td><strong>Seniors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>3.09</td>
<td>3.43</td>
<td>3.00</td>
<td>3.04</td>
<td>3.09</td>
<td>3.00</td>
</tr>
<tr>
<td>SD</td>
<td>1.41</td>
<td>.66</td>
<td>.52</td>
<td>1.07</td>
<td>.95</td>
<td>1.04</td>
</tr>
<tr>
<td>Variance</td>
<td>1.99</td>
<td>.44</td>
<td>.27</td>
<td>1.13</td>
<td>.90</td>
<td>1.09</td>
</tr>
<tr>
<td><strong>Graduate Students under age 30</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>3.93</td>
<td>3.70</td>
<td>3.12</td>
<td>2.95</td>
<td>3.56</td>
<td>3.72</td>
</tr>
<tr>
<td>SD</td>
<td>1.30</td>
<td>1.12</td>
<td>1.24</td>
<td>1.45</td>
<td>.96</td>
<td>1.03</td>
</tr>
<tr>
<td>Variance</td>
<td>1.69</td>
<td>1.26</td>
<td>1.53</td>
<td>2.09</td>
<td>.92</td>
<td>1.06</td>
</tr>
<tr>
<td><strong>Graduate Students over age 30</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>4.08</td>
<td>3.59</td>
<td>3.56</td>
<td>3.23</td>
<td>3.44</td>
<td>4.08</td>
</tr>
<tr>
<td>SD</td>
<td>1.74</td>
<td>1.33</td>
<td>1.41</td>
<td>1.60</td>
<td>1.47</td>
<td>.77</td>
</tr>
<tr>
<td>Variance</td>
<td>3.02</td>
<td>1.77</td>
<td>1.99</td>
<td>2.55</td>
<td>2.15</td>
<td>.60</td>
</tr>
</tbody>
</table>
In order to determine whether the Measure of Epistemological Reflection assesses development on the Perry scheme, exploration of concurrent validity occurred in two ways. First, an analysis of variance was conducted on the MER data to determine whether differences existed by level of education as was hypothesized. Second, correlations were computed to examine the relationship of the MER to the Measure of Intellectual Development (Knefelkamp, 1974). The latter instrument has been used in various studies involving the Perry scheme, and initial reports on its validity are soon to be published (Moore, 1982).

The analysis of variance by level of education for the MER total protocol scores is shown in Table 17. As was the case with internal consistency tests, this analysis was based on the highly trained raters' compromise total protocol ratings for the entire sample. As was expected a difference exists between levels significant at \( p < .0001 \).

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Levels</td>
<td>7</td>
<td>8.0143</td>
<td>26.72</td>
<td>.0001</td>
</tr>
<tr>
<td>Error</td>
<td>147</td>
<td>.29999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>154</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 18 shows the analysis of variance by level, sex and interaction of
level and sex. Level of education accounted for the difference between
groups, being significant at $p < .0001$. Sex and interaction effects did
not reveal significant differences.

Table 18
Two Way Analysis of Variance
Level, Sex and Interaction Effects

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>3</td>
<td>54.7765</td>
<td>60.88</td>
<td>.0001</td>
</tr>
<tr>
<td>Sex</td>
<td>1</td>
<td>.9119</td>
<td>3.04</td>
<td>.0833</td>
</tr>
<tr>
<td>Level and Sex</td>
<td>2</td>
<td>.4121</td>
<td>0.46</td>
<td>.7159</td>
</tr>
</tbody>
</table>

In order to clarify the differences among levels of education, Scheffe's
method of post hoc multiple comparisons was utilized to compare levels
of education. Scheffe's method serves as the most conservative post hoc
test and is suitable when working with unequal n's (Kennedy, 1978).
Comparisons were conducted for all six combinations of levels of
education. All were significant ($p < .10$) with the exception of the
graduate students under age 30 compared to the graduate students over
age 30. The confidence interval was set at 90% due to the extremely
conservative nature of Scheffe's method in an attempt to get a more
accurate estimate of significance of the differences.

A second method of assessing validity of the MER involved computing
correlations with the Measure of Intellectual Development. Pearson
product-moment correlations between the two instruments are reported by sample in Table 19. Contrary to this study's hypothesis, the correlations were extremely low, being .11 for the derivation sample and .04 for the cross validation sample. Further exploration revealed very little variance in the mean MID scores by level as shown in Table 20. This narrow variance may account for the low correlations between the two measures. Comparing MID means with those of MER compromise total protocol ratings indicated a difference in variability of the measures. An analysis of variance on the MID scores, reported in Table 21, did not indicate a difference across levels of education.

Finally, rounded total protocol ratings for the MID and MER were compared for percentage agreement. Exact agreement was 41.29%, with an additional 44.52% with one position. Subsequently less than 15% of the total protocol ratings were discrepant by two or more positions.
Table 19
Correlations between the MER and MID by Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th>Derivation</th>
<th>Cross Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>84</td>
<td>71</td>
</tr>
<tr>
<td>Correlation Coefficients</td>
<td>.11</td>
<td>.05</td>
</tr>
<tr>
<td>P</td>
<td>.306</td>
<td>.705</td>
</tr>
</tbody>
</table>
Table 20

Means of MER and MID Scores Overall and by Level of Education

<table>
<thead>
<tr>
<th>Level of Education</th>
<th>Freshmen</th>
<th>Seniors</th>
<th>Graduates under 30</th>
<th>Graduates over 30</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>50</td>
<td>23</td>
<td>43</td>
<td>39</td>
<td>155</td>
</tr>
<tr>
<td>MID X</td>
<td>2.83</td>
<td>2.61</td>
<td>2.88</td>
<td>2.76</td>
<td>2.79</td>
</tr>
<tr>
<td>MER X</td>
<td>2.52</td>
<td>3.22</td>
<td>3.70</td>
<td>3.96</td>
<td>3.31</td>
</tr>
</tbody>
</table>
Table 21
Analysis of Variance of MID Scores

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Levels</td>
<td>7</td>
<td>.5557</td>
<td>.88</td>
<td>.527</td>
</tr>
<tr>
<td>Error</td>
<td>147</td>
<td>.6328</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>154</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Introduction

The purpose of this study has been to evaluate the Measure of Epistemological Reflection in order to determine its worth as a method of assessing development on the Perry scheme. To that end, eight research hypotheses were presented in Chapter Three. These hypotheses focused on three areas of test evaluation: interrater reliability, internal consistency and validity. The discussion of results in this chapter is organized around these hypotheses.

Review of the Results

Interrater Reliability

Hypothesis One: There will be a moderate to high correlation between the ratings of the two highly trained experts.

Comparison of both domain and total protocol ratings lend support to the predicted correlation of highly trained raters' ratings. Comparison of domain ratings constitutes the most rigorous method of evaluating reliability as these ratings are based on individual questions, or one sixth of the overall test. Exact agreement by domain is moderate, generally in the 50% and 60% range. Agreement of the highly trained raters within one stage was in the 80% and 90% range and therefore discrepancies of more than one position are minimal. Comparison of
total protocol ratings, reflective of the entire measure, show high correlation of the highly trained raters' ratings for both samples, .82 and .78, respectively. Both correlations are significant at $p < .0001$. Percentage agreement on modal total protocol ratings also shows high agreement between highly trained raters for both samples (59.5% and 73.2%, respectively). Consequently Hypothesis One is supported by the results. However, Tamashiro (1981) recommends that exact agreement of less than 70% warrants review of the manual. The self-trained expert rater viewed some of the reasoning structures in the revised manual to be discrepant by one stage and thus an additional revision of the manual will be conducted.

Hypothesis Two: There will be a moderate to high correlation between the ratings of the two highly trained raters and the self-trained expert rater.

In regard to Hypothesis Two, comparisons were conducted on the basis of self-trained expert ratings with ratings of each highly trained rater as well as with compromise ratings of the two highly trained raters. Again these comparisons support the predicted correlation of highly trained raters and the self-trained expert. Exact percentage agreement on compromise highly trained rater and self-trained expert domain ratings is slightly lower overall than was that of the two highly trained raters, although it remains in the 50% and 60% range. Within one position agreement is approximately 90% with the exception of domain one. Percentage agreement on domain ratings of each highly trained rater with the self-trained expert appear quite similar (50% range) to the results obtained utilizing compromise ratings. Pearson product-moment correlation coefficients based on total protocol ratings also
indicate similarity in comparison of highly trained raters individual and compromised ratings with the self-trained expert. Correlations for all combinations of these three raters are high (.75 to .84) and significant at \( p < .0001 \). Exact percentage of modal total protocol agreement among these same raters in generally in the 60% to 75% range. These comparisons overall indicate a high correlation between highly trained raters and the self-trained expert as was hypothesized. This would imply that the rating procedure of the new measure was learned and used by the self-trained expert in ways consistent with the highly trained raters.

Hypothesis Three: There will be a moderate correlation between the ratings of the two highly trained raters and the two self-trained raters.

As noted in Hypothesis Three, self-trained raters were expected to obtain only a moderate level of reliability with highly trained raters. This conservative prediction was based on the assumption that high interrater reliability would only be achieved after more comprehensive training with the rating manual and expert knowledge of the theory. The results, however, reveal that in the case of these self-trained raters fairly high correlations were obtained. The self-trained raters achieved exact percentage agreement on domain ratings in the 50% and 60% range, similar to the range achieved by highly trained raters. The self-trained raters within one position agreements were very close (high .80's and low .90's) to those achieved by highly trained raters. When compared with the highly trained raters on domain ratings, rater four achieved exact agreement in the 45% to 62% range. Rater five's agreement was somewhat lower ranging from 38% to 59%. 
Review of total protocol ratings indicates highly significant Pearson product-moment correlations for both self-trained raters with highly trained raters and the self-trained expert. Rater four correlated with these raters in the .70's, and .80 with the compromise of highly trained raters. Rater five's correlations were in the high .60's and low .70's. All correlations were significant at $p < .0001$, indicating that the self-trained raters achieved a degree of accuracy higher than was predicted. Both self-trained raters in this study hold doctorate degrees in Student Personnel Work and are practitioners in the field. Their level of education and experience may account for their success in rating. As noted previously rater five achieved less consistency with expert raters. Of the two self-trained raters, rater five has had more in-depth experience with other methods of rating Perry type data which may have affected the perspective with which the MER manual was approached.

Summarizing results pertinent to interrater reliability, the study reveals fairly high interrater reliability among the three categories of raters included in the study. These correlations were statistically significant both on the basis of comparison of the categories and individuals across and within categories. Exact percentage agreement is slightly below the 70% recommended by Tamashiro (1981). Therefore the planned revisions noted earlier may serve to raise exact agreement in future studies with the MER. Nevertheless, the findings of this study do support the proposition that the MER elicits ratable data and that the rating procedure and manual result in fairly high consistency among raters in domain and total protocol ratings.
Hypothesis Four: The interrater reliability for the cross-validation sample will be higher than for the derivation sample.

Hypothesis Four proposed that higher interrater reliability would be achieved in the cross-validation sample. This hypothesis was based on the assumption that more specification of reasoning structures in the revised manual would increase interrater reliability. Recall that the derivation sample was rated with the preliminary manual, while the cross-validation sample was rated with the revised manual. Comparison of highly trained raters domain ratings by sample do not provide substantive clarification of potential advantages of the revised manual. While increases in exact agreement from 5% to 15% exist for three domains, decreases from 2% to 18% exist for the remaining three. Domain one exhibits the greatest increase from the derivation to the cross-validation sample. The question for this domain, which focuses on decision-making, yielded more diverse data than any other question in the instrument. The question did not specify the type of decision the respondent should describe and subsequently many respondents discussed decisions that had no direct relation to education or intellectual development. In this domain, increased clarification of reasoning structures in the revised manual appear to have assisted in rating these more diverse responses. It seems logical to expect that the diversity of response required more inference in the derivation sample. Degree of inference would be reduced by the increased specificity of the revised manual. Domains four and five, focusing on the role of peers and evaluation respectively, showed moderate increases (5% to 7%). The role of peers data did serve to further clarify respondents' thinking on the
value of peer involvement in the learning process. Perry's original work and subsequent research has focused on whether students see peers as having valid knowledge or expertise as a criterion for position assignment. It has been assumed that a preference for low peer involvement indicates lack of knowledge on the part of peers and high involvement indicates peers do have legitimate knowledge. Thus even though the derivation sample rating focused on the reasoning structure, this notion probably affected position assignments. However, the empirical verification process conducted on derivation sample data clearly revealed that the above assumption is not necessarily accurate. Reasoning structures related to degree of peer involvement emerged across positions indicating that involvement and the assignment of knowledge are not directly connected. For example, a Position Two structure described the role of peers as high in involvement but described involvement as asking questions to obtain answers for the class. A Position Four structure, in contrast, described high involvement as sharing of ideas, perspectives and opinions. A Position Five structure described high involvement as discussion of questions and solutions contributed by class members which were supported by rationale and evidence. These finer distinctions may account for increased agreement in the cross-validation sample for domain four. A similar situation existed for domain five on evaluation. Many responses advocated the same types of evaluation methods but for different reasons. Clarification and categorization of these reasons helped to increase rater agreement with the revised manual.
Of the three domains revealing decreases in agreement, domains two and three addressing the role of the learner and the instructor showed the most decrease. Although more specific reasoning structures emerged during empirical verification of derivation sample data, the distinctions across positions in these domains are less clear than for other domains. Highly trained raters experienced a fair amount of difficulty in these domains distinguishing between Position Two and Three reasoning structures. (Compromising of ratings upon completion of the cross-validation sample revealed that most discrepancies were in these areas.) An example of two reasoning structure emerging during empirical verification of domain two appears below:

Position Two: Preferred type of learning is the method that is clearly right and the easiest way to obtain the answers.

Position Three: Preferred type of learning is one that is easiest to learn or remember.

As can be seen from the example, if the respondent does not clearly elaborate on his/her reasoning structure, distinguishing between the positions above becomes difficult. A similar situation exists in domain three, the role of the instructor. A Position Two reasoning structure focuses on the instructor making learning entertaining while a Position Three structure focuses on rapport between the instructor and students. Both are aimed at making learning easier but finer distinctions are needed to reduce inference in rating. These issues do not appear to exist for domain six, as evidenced by only a 2% decrease in agreement. The lack of increase here would indicate that substantial improvements were not achieved in specificity of the revised manual for this domain.
Exploring Hypothesis Four from the vantage point of highly trained raters' compromise with the self-trained expert domain ratings adds information about particular domains. Although decreases in exact agreement again appeared in three domains, they are much smaller (4% to 7%). One of these decreases occurred in domain one, where highly trained rater agreement increased. As mentioned previously it is expected that highly trained rater agreement increased due to increased specificity of the distinctions across positions for decision making. It is possible that the self-trained expert viewed these data from a broader perspective and thus inferences on the part of the highly trained raters and self-trained expert may have varied. A similar possibility exists for domains five and six. A second possibility in these three domains relates to the use of language and complexity of verbal response. These domains were addressed in more general terms by most respondents than were two, three and four. These data varied in the extent and style with which subjects responded to the questions. The highly trained raters, having mutually explored reasoning structures at great length during empirical verification, should have been more readily able to identify structures despite differences in the complexity of the response. In some cases the highly trained raters consciously assigned higher positions due to inferred reasoning structures even if the response did not initially appear as complex. The self-trained expert, not having the same involvement in empirical verification, may have inferred differently in cases of lack of complexity in the style of responses. Finally, the three domains in which exact agreement did increase are those in which reasoning
structures emerged to enhance distinctions assumed by earlier research. The moderate increases in domains two and three are undoubtedly affected by the difficulty described earlier by the internal experts in these domains. Without this problem, increases probably would have been closer to that of domain four (15%) which represents the type of increase expected on the basis of finer distinctions in the rating manual.

In summary, results related to Hypothesis Four are confounded by a) the specificity of questions in the instrument, b) differing degrees to which reasoning structures in different domains are differentiated, c) unclear distinctions between position reasoning structures that remain after empirical verification, and d) potential differences in rater focus on structure. Despite these difficulties, exact percentage agreement on modal total protocol ratings shows a slight increase in the cross-validation sample. Correlation coefficients also show increases among combinations of highly trained raters and the self-trained expert, but not of the two highly trained raters. All of these results imply that the concept of the revised manual's increased specificity leads to better interrater reliability may be a sound one. However, the issues noted here call for resolution before this particular manual will clearly reveal the advantages of specificity.

Internal Consistency

Hypothesis 5: The Measure of Epistemological Reflection will contain a moderate degree of internal consistency across domains.

A moderate degree of internal consistency was expected on the basis of two notions that may appear initially to be in conflict. The first
notion related to what individual development on the Perry scheme may look like at a given point in time. Since change or transition occurs on the basis of encountering dissonance, an individual moves to a new way of thinking in a particular domain. This change does not necessarily affect the person's overall position on the scheme, particularly if the change is the first in a new way of thinking. The theoretical perspective of cognitive development would assert that subsequent changes serve to alter more domains until the person has acquired a new modal way of thinking. This process holds important implications for the assessment in internal consistency of the MER. Since the MER intentionally taps six different domains, it is not expected that one subject would be in the same position on all domains. Although most subjects would be expected to have a modal way of thinking, some domains may have not yet changed to the modal position or be in transition toward a new position. At the same time the domains included in the MER were expected to measure the same development due to previous Perry scheme research. Since total protocol ratings incorporate modal and transitional thinking (the latter being defined as one third of the instrument or two domains) internal consistency between domain scores and total scores was expected.

Cronbach's alpha coefficient of .76 for the entire sample indicates a fairly high degree of internal consistency between domain and overall scores. This suggests that in general scores on a given domain are representative of overall scores across domains. Comparison of agreement between domain and total protocol scores serve to enlighten the differences across domains. Domain one showed the least consistency
with total protocol scores, probably due to the lack of uniformity of the responses for this domain noted earlier. The remaining domains showed exact agreement ranging from 56.2% to 63.8%. These agreement levels do not exhibit much difference across domains two through six in terms of the degree to which they represent overall scores. The percentage of agreement within one position of domain and total protocol scores takes on a different meaning given the theoretical notion described earlier. A person in transition would be expected to have some domain scores that would be one position higher or lower (or both) than the modal total protocol score. The average agreement within one position across the six domains was approximately 40%. These results would imply less consistency than evidenced by Cronbach's alpha on the surface, but are affected by the fact that computation of percentage agreement necessitated use of modal scores rather than total protocol ratings that were adjusted to include all scores in the instrument.

Combining exact and within one position agreement for domain and total protocol ratings shows that in domains two through six only a small degree (2 to 4%) of discrepancy of more than one position exists.

Finally, means of domain scores within levels of education showed that average scores in a domain are as consistent as would be expected with average scores in other domains. Since difference scores would be predicted for different levels of education, looking at levels of education separately provides a clearer picture of consistency of means. For freshmen the means are generally in a range from 2.36 to 2.62. Only domain one, decision-making (2.06), falls outside that range. Standard deviations and variances are similar and suggest high
consistency across domains. The senior group indicates even higher consistency among domains with a range of 3.00 to 3.09. Only domain two, role of the learner (3.43), falls outside of that range. Standard deviations and variances are less similar, however, suggesting that domains two and three (role of instructor) have more variance than the rest. Both graduate student groups show a wider range of means and higher variance. The ranges are still within one position and indicate narrower variance in domains two through five. This could be due to the focus of these domains being in a classroom context whereas domains one (decision-making) and six (nature of knowledge) are more general.

In summary Hypothesis Five is supported indicating that domains of the MER do exhibit a moderate degree of internal consistency. The results support the prediction that the domains measure the same development consistently. Variances are consistent with theoretical expectations of persons in transition or persons who have a modal position and on some occasions use lower positions.

Validity

Hypothesis Six: MER ratings will indicate differences between levels of education in the direction of increasing complexity on the Perry scheme.

Hypothesis Seven: MER ratings will not indicate differences on the basis of sex.

Hypothesis Six is based on the theoretical perspective that complexity of thinking increases as individuals encounter and accommodate more complex experiences. The literature on the Perry scheme noted in Chapter Two supports this notion and shows that increased levels of education tends to result in more complex ways of viewing the world.
Hypothesis Seven is based on the assumption that sex differences do not have an impact on access to or processing of more complex experiences. Sex differences have not yet been found in other Perry research.

A two-way analysis of variance of the MER data provides strong support for Hypotheses Six and Seven. Highly statistically significant differences exist by level of education ($p < .0001$). Sex and interaction effects were not significant. Overall means showed a progression of increased complexity by level of education as did domain means discussed in relation to internal consistency. Further clarity of the differences among levels of education is provided through multiple comparisons. Use of the Scheffe method to compare all possible combinations revealed statistically significant differences ($p < .10$) exist between freshmen and seniors and between seniors and graduate students under age 30. Subsequently freshmen differ from both graduate groups, as do seniors. The only comparison not achieving statistical significance is that of the two graduate student groups even though the mean of the older group is higher. These differences are in line with the prediction of Hypothesis Six.

In summary Hypotheses Six and Seven are supported by the results. Progression on the Perry scheme is evident across levels of education and there are no apparent sex differences. The progression is statistically significant across all levels of education except the two levels of graduate students.

Hypothesis Eight: There will be a positive correlation between the MER and the MID.

Hypothesis Eight is based on the assumption that the MER measures the
same development as does the Measure of Intellectual Development (MID). As noted in Chapter Three, differences do exist in the instrument stimuli and rating procedures which may confound the relationship of the two instruments. However, both utilize similar content areas or domains, and were therefore expected to exhibit a positive correlation.

The results of this study do not support the hypothesized relationship between the MER and the MID, and subsequently Hypothesis Eight was rejected. Correlations were surprisingly low for both the derivation and cross-validation samples (.11 and .05 respectively) and are not significant. Further description of the MID data provides useful information in interpreting the meaning of these correlations. The MID means for all levels of education revealed a narrow range (2.61 to 2.88). This lack of variability would affect the correlation of the two instruments significantly. The progression of scores across levels of education differed from that of the MER as well. MID means showed seniors with the lowest mean followed by graduate students over age 30, freshmen and graduate students under age 30, in that order. An analysis of variance of the MID data did not reveal significant differences by level of education.

Since the use of the MID in the current study was intended to provide a measure of validity of the MER, the low correlation between the two could be interpreted as a sign that the MER does not measure development on the Perry scheme. Based on the additional information on the MID scores, however, that conclusion may be erroneous. As noted in Chapter Three, Rodgers (1975-1979) has conducted research using the KneWi which
indicates a relationship between level of education and Perry level. His data show the majority of freshmen in dualism, or Positions two and three (2.52 to 2.71), as do both the MER and MID scores in this study. (The MID mean was 2.83 while the MER mean is 2.52.) Rodgers' research on graduate students doing first year masters level work shows 75% of these students in relativism (Position Four or higher). The graduate group mean scores on the MID in this study were 2.76 and 2.88, indicating dualistic positions two and three. The MER mean scores of 3.70 and 3.96 for the same group places graduate students in higher positions (Four and Five). Thus the MER data is a closer match to Rodgers' research than is the MID. Other research reports differences between freshmen and seniors. Meyer (1977) reported significant differences between freshmen and seniors on Perry level with seniors scoring higher. Heidke and Omahan (1981) also reported significant differences between freshmen and seniors, with seniors scoring higher. The MID average score for seniors in this study is lower than that of freshmen. The research on the Reflective Judgment scheme (King, 1977) also indicates a difference in undergraduate and graduate students. Although this scheme measured reflective judgment which King describes as slightly different but based in part on the Perry scheme, it also found increasing complexity occurred with increased level of education. King reported significant differences between college (undergraduate) students and graduate students on reflective judgment scores.

In addition to the literature described above, Moore (1982) indicates that MID data from various colleges and universities contained in the
Center for Application of Developmental Instruction data bank place entering freshmen between Position Two and Three, while seniors are generally between Three and Four. This information would imply that MID data taken collectively do converge with other research on differences across levels of education. This variability Moore describes does not seem to appear in the Maryland raters' rating of the current data. This lack of variability probably would not be due to unique sample characteristics as the sample was randomly drawn within levels of education. Other variables confounding the correlations are most likely a) difference in stimuli, b) amount of data the respondent provides, and c) the specificity of the rating process. Due to the MID's open ended questions, which are not separated on the page, the respondent could provide a general response. The follow-up questions included in the MER may have prompted the respondent to elaborate further and subsequently reveal more complex reasoning. The amount of data provided was in some cases greater with the MER which may have led to a clearer picture of the respondent's reasoning with which to determine a rating. The rating processes for the MER and the MID also differ. The MER rating process, as described in this study, focuses on reasoning structure underlying the response and is done through a specific manual. The MID process focuses on specific cues which reflect underlying structures. Moore (1982) emphasizes that raters must be highly versed in the Perry theory and cues since the process is not a "cookbook" approach. This may imply that greater flexibility in assigning ratings is afforded MID raters than is the case with the MER. Regardless of the source of the lack of variability in MID scores, the data from the MER vary as expected in scope and level.
Implications

The development of the MER was intended to move measurement of intellectual development on the Perry scheme toward a more standardized process. As this movement has occurred in other areas of measurement, the major concern has been whether increased structure of stimuli would restrict the expression of respondents' frame of reference, thus giving a less accurate picture of their thinking structure. The results of this study indicate that the MER serves as an example of more structured stimuli that does not adversely affect position assignment. Subjects' position assignments in this study were as expected based on previous studies. This could indicate that the increased structure of the instrument still allowed for adequate expression of frame of reference or thinking structure. Therefore the advantages of this more standardized data collection can be enjoyed without risk of losing the essence of thinking required to elicit Perry position.

One major advantage of the increased structure of the MER is the quality of ratable data which results. Focusing respondents' thinking provided a context in which the respondent could speak to the domain specifically. The focus on reasoning structure helped clarify the basis for the respondents' thinking. Separating follow-up questions, each with its own space for response, assisted in respondents elaborating on their thoughts, resulting in most cases in a thorough response to each domain. Finally, the fact that the MER explores six distinct domains appears to provide an adequate picture of the respondents' thinking.

A second major advantage of the standard format of the MER is the resulting specificity that is then possible in a rating manual. The
specificity of the MER revised manual indicates that fairly high correlations can be acquired with self-trained raters. This may eventually lead to more practical use of the MER in instrument development research as well as practice.

Limitations
Although the preliminary results on the reliability and validity of the MER are very encouraging, a few cautions are in order. The sample of subjects, while adequate for initial study, is not sufficiently large to explore the characteristics of the MER in a comprehensive manner. The sample of seniors (N=23) was particularly small. A second caution related to the sample is that the subjects have relatively homogeneous socioeconomic backgrounds. Since the sample is random by level of education it is expected that the socioeconomic status of subjects included is similar to that of the general population at The Ohio State University. This information serves to emphasize that generalizability of these results is not possible beyond a university setting.

Initial review of validity shows that the MER places students at expected Perry levels. However, more in-depth study is needed to assure that the MER is a valid measure of development on the Perry scheme. Finally, the data and analysis of the data reveal that another revision of the rating manual is necessary prior to additional study.

Areas for Future Inquiry
The status of the MER as a result of data analysis warrants further work in the areas of the instrument stimuli, the rating manual and validity. Minor revisions in the instrument are underway on the basis of this study in an attempt to improve sections of the instrument. The
questions to elicit decision-making (domain one) need revision to maintain the educational context. The revised manual also needs revision on the basis of empirical data in the cross-validation sample. This second revision should include an attempt to clarify distinctions between positions in domains that may not be distinct enough in the revised manual. The revision should also include an analysis of congruency of position reasoning structures across the six domains and agreement on new reasoning structures evidenced in the cross-validation sample.

Following the above revisions, the MER needs to be tested with additional and preferably larger samples from a variety of colleges, universities, and secondary schools. For example, future samples might include high school students, college freshmen and seniors, graduate students under and over age 30, and faculty in departments such as philosophy. This broad sampling might help empirically describe the lower and upper levels of the scheme. Longitudinal research is needed to explore validity of the MER as well as use of the instrument in developmental interventions. At the same time further exploration of the degree of training acquired to achieve adequate interrater reliability is necessary.

Summary

The development of the Measure of Epistemological Reflection serves as a step forward in assessment of development on the Perry scheme. Although in-depth study of the instrument will be essential to its full development in the next few years, it holds promise for more standardized and practical measurement. In order to utilize Perry's
theory of intellectual development in providing meaningful experiences for college students, practical production measurement devices are necessary. Thus the MER is not only a step toward improved measurement but a helpful tool in practice. Increased standardization of data collection and analysis is expected to help further knowledge of the Perry scheme and its utility in numerous aspects of higher education.
APPENDIX A

MEASURE OF EPISTEMOLOGICAL REFLECTION

INSTRUCTIONS: The questionnaire that follows has to do with your perspective on a number of concerns related to college students. Each of the questions on the following pages asks for your opinion or choice on a given subject, and the REASONS why you have that particular perspective or opinion. We are interested in understanding your perspective as fully as possible. Please give as much detail as you can to describe how you feel about each question. Feel free to use the backs of pages if you need more space. Thank you.

NAME ____________________________
AGE ________ CLASS RANK ______
SEX (circle) MALE FEMALE
COLLEGE MAJOR _______________
FATHER'S JOB _________________
MOTHER'S JOB _________________
DATE _________________________
THINK ABOUT THE LAST TIME YOU HAD TO MAKE A MAJOR AND DIFFICULT DECISION IN WHICH YOU HAD A NUMBER OF ALTERNATIVES (E.G., WHICH COLLEGE TO ATTEND, COLLEGE MAJOR, CAREER CHOICE, ETC.). WHAT WAS THE NATURE OF THE DECISION?

WHAT ALTERNATIVES WERE AVAILABLE TO YOU?

HOW DID YOU FEEL ABOUT THESE ALTERNATIVES?

HOW DID YOU GO ABOUT CHOOSING FROM THE ALTERNATIVES?

WHAT THINGS WERE THE MOST IMPORTANT CONSIDERATIONS IN YOUR CHOICE? PLEASE GIVE DETAILS.
DO YOU LEARN BEST IN CLASSES WHICH FOCUS ON FACTUAL INFORMATION OR CLASSES WHICH FOCUS ON IDEAS AND CONCEPTS?

_________________________________________________________

WHY DO YOU LEARN BEST IN THE TYPE OF CLASS YOU CHOSE ABOVE?

_________________________________________________________

_________________________________________________________

_________________________________________________________

WHAT DO YOU SEE AS THE ADVANTAGES OF THE CHOICE YOU MADE ABOVE?

_________________________________________________________

_________________________________________________________

_________________________________________________________

WHAT DO YOU SEE AS THE DISADVANTAGES OF THE CHOICE YOU MADE ABOVE?

_________________________________________________________

_________________________________________________________

_________________________________________________________

IF YOU COULD GIVE ADVICE TO ANYONE ON HOW BEST TO SUCCEED IN COLLEGE COURSEWORK, WHAT KIND OF ADVICE WOULD YOU GIVE THEM? TALK ABOUT WHAT YOU BELIEVE IS THE KEY TO DOING WELL IN COLLEGE COURSES.

_________________________________________________________

_________________________________________________________

_________________________________________________________

_________________________________________________________

_________________________________________________________
DURING THE COURSE OF YOUR STUDIES, YOU HAVE PROBABLY HAD INSTRUCTORS WITH DIFFERENT TEACHING METHODS. AS YOU THINK BACK TO INSTRUCTORS YOU HAVE HAD, DESCRIBE THE METHOD OF INSTRUCTION WHICH HAD THE MOST BENEFICIAL EFFECT ON STUDENTS.

_________________________________________________________________________________

_________________________________________________________________________________

_________________________________________________________________________________

WHAT MADE THAT TEACHING METHOD BENEFICIAL? PLEASE BE SPECIFIC AND USE EXAMPLES.

_________________________________________________________________________________

_________________________________________________________________________________

_________________________________________________________________________________

WERE THERE ASPECTS OF THAT TEACHING METHOD WHICH WERE NOT BENEFICIAL? IF SO, PLEASE TALK ABOUT SOME OF THE ASPECTS AND WHY THEY WERE NOT BENEFICIAL.

_________________________________________________________________________________

_________________________________________________________________________________

_________________________________________________________________________________

WHAT ARE THE MOST IMPORTANT THINGS YOU LEARNED FROM THE INSTRUCTOR'S METHOD OF TEACHING?

_________________________________________________________________________________

_________________________________________________________________________________

_________________________________________________________________________________

PLEASE DESCRIBE THE TYPE OF RELATIONSHIP WITH AN INSTRUCTOR THAT WOULD HELP YOU TO LEARN BEST AND EXPLAIN WHY.

_________________________________________________________________________________

_________________________________________________________________________________

_________________________________________________________________________________
DO YOU PREFER CLASSES IN WHICH THE STUDENTS DO A LOT OF TALKING, OR WHERE STUDENTS DON'T TALK VERY MUCH?

________________________________________________________________________

WHY DO YOU PREFER THE DEGREE OF STUDENT INVOLVEMENT/PARTICIPATION THAT YOU CHOSE ABOVE?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

WHAT DO YOU SEE AS THE ADVANTAGES OF YOUR PREFERENCE ABOVE?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

WHAT DO YOU SEE AS THE DISADVANTAGES OF YOUR PREFERENCE?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

WHAT TYPE OF INTERACTIONS WOULD YOU LIKE TO SEE AMONG MEMBERS OF A CLASS IN ORDER TO ENHANCE YOUR OWN LEARNING?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
SOME PEOPLE THINK THAT HARD WORK AND EFFORT WILL RESULT IN HIGH GRADES IN SCHOOL. OTHERS THINK THAT HARD WORK AND EFFORT ARE NOT A BASIS FOR HIGH GRADES. WHICH OF THESE STATEMENTS IS MOST LIKE YOUR OWN OPINION?

____________________________________________________________________________________________________________________________________________________________________________________

IDEALLY, WHAT DO YOU THINK SHOULD BE USED AS A BASIS FOR EVALUATING YOUR WORK IN COLLEGE COURSES, AND WHO SHOULD BE INVOLVED IN THE EVALUATION?

____________________________________________________________________________________________________________________________________________________________________________________

PLEASE EXPLAIN WHY YOU THINK THE RESPONSE YOU SUGGESTED ABOVE IS THE BEST WAY FOR EVALUATING STUDENTS' WORK IN COLLEGE COURSES.

____________________________________________________________________________________________________________________________________________________________________________________
SOMETIMES DIFFERENT INSTRUCTORS GIVE DIFFERENT EXPLANATIONS FOR HISTORICAL EVENTS OR SCIENTIFIC PHENOMENA. WHEN TWO INSTRUCTORS EXPLAIN THE SAME THING DIFFERENTLY, CAN ONE BE MORE CORRECT THAN THE OTHER?

WHEN TWO EXPLANATIONS ARE GIVEN FOR THE SAME SITUATION, HOW WOULD YOU GO ABOUT DECIDING WHICH EXPLANATION TO BELIEVE? PLEASE GIVE DETAILS AND EXAMPLES.

CAN ONE EVER BE SURE OF WHICH EXPLANATION TO BELIEVE? IF SO, HOW?

IF ONE CAN'T BE SURE OF WHICH EXPLANATION TO BELIEVE, WHY NOT?
APPENDIX B

First Solicitation Letter

This letter is to ask you to participate in a research study this quarter. I am a doctoral student in the College of Education, conducting dissertation research on cognitive or intellectual reasoning. In order to complete my research I need a number of students to complete two essay questionnaires about learning, classroom instruction and academic decision-making. The time involved will be approximately two hours.

The purpose of my research is to develop a practical method of measuring how students think about learning and college coursework. Should this be possible, instructors could use the measure to gain information about what their students need in the learning process and increase student satisfaction with college coursework. A summary explanation of my study is included on the attached page.

Your name was chosen from the Student Directory. I would sincerely appreciate your willingness to participate. Unfortunately, I am unable to pay students who assist. I will, however, meet with any participants who wish to know how their questionnaires are interpreted and provide any other additional information about the study.

All information collected on the questionnaires will be entirely confidential. Should you have questions about participating please contact me at 876-8728, or contact my faculty supervisor Dr. Robert F. Rodgers at 422-7700. The enclosed postcard is for your convenience in letting me know whether you are able to participate. Thanks in advance for your anticipated cooperation!

Sincerely,

Marcia B. Taylor
The information below is intended to answer potential questions you may have about the study in which you are being asked to participate.

WHY STUDY INTELLECTUAL REASONING?

In 1970, William G. Perry studied and described students' perceptions of their experience with college. Perry indicated that different students viewed learning, knowledge and instruction differently. Many researchers subsequently attempted to organize learning experiences and formal instruction in ways that would correspond to various students' methods of reasoning. The current methods to measure students' reasoning, however, need a great deal of interpretation to be useful. Thus a new, more practical method would assist in considering students' reasoning in designing coursework.

WHAT DOES THIS STUDY INVOLVE?

In order to study a new method of measuring students' reasoning, a group of students is needed to complete the new questionnaire. The results would then be interpreted to see if the new questionnaire is able to measure the same variations in students' methods of reasoning. For comparison purposes, a current questionnaire to measure reasoning needs to be completed also. None of the questions on either instrument are personal. The time involved to complete the questionnaires would be approximately two hours.

HOW WERE PARTICIPANTS CHOSEN?

A random group of persons were selected from the Student Directory. It is important to obtain a random group of persons for a research study so that characteristics of persons selected represent students in general and not a special group which might be biased on the content of the study. Some participants were selected from students enrolled in Psychology 100, since that group is representative of OSU students in general.

WHAT WILL HAPPEN TO THE QUESTIONNAIRES?

Each questionnaire will be given a number at which time the name of the respondent will be removed from the questionnaire. The researchers will maintain a list of participants and their number in order to summarize biographical data such as age, sex and class rank, but no one else will be permitted to know who filled out the questionnaires. All information from the questionnaires will be reported anonymously.

WHY SHOULD I PARTICIPATE?

By participating you will have an opportunity to learn more about how you reason and think. By giving your time to help with this study you may also help to create a means by which instructional methods can be developed to meet students' needs more effectively.
Second Solicitation Letter

Thank you for your willingness to participate in my research! Enclosed you will find the consent form, the two questionnaires, a page requesting demographic information, and an envelope in which to return the materials. Please keep your copy of the consent form.

Both questionnaires ask for your perspective on a variety of topics. It is important that you express your own views and the reasons that make up your perspective. It would be helpful if you would complete the questionnaires in the order in which you find them in the envelope. If possible, please try to complete both questionnaires in one sitting. Please complete the page entitled "Additional Demographic Information" after you have finished the questionnaires.

You will notice that the return envelope shows my campus address. (If your address implied that you did not come to campus regularly, I enclosed a stamped envelope.) Please drop the completed materials in any campus mail slot. I would appreciate having the questionnaires back by the end of the quarter, but realize that timing may be a problem! Therefore if you need to wait until the close of the quarter to complete the questionnaires, please feel free to do so.

Should you have questions about the study or the questionnaires please contact me at 876-8728 (home) or 422-3930 (work), or contact my faculty advisor Dr. Rodgers at 422-7700. Thanks again for your assistance!

Sincerely,

Marcia B. Taylor
APPENDIX C

Instruction Letter to First Group of Participants

This letter is in reference to my dissertation research I wrote to you about in November. I realize that the timing of my initial letter was problematic since it was so close to the end of the quarter. I am writing now to ask you to reconsider participating in hopes that you may have more time this quarter.

I need to request your help for two reasons. First, the sample of students chosen is a random one. If it becomes necessary to solicit participants from a new sample, the value of having a random sample will be significantly decreased. Secondly, I need one hundred more participants in order to complete my research. I have enclosed the explanation sheet for the study for your use in deciding whether you are able to participate.

A self-addressed, stamped postcard is enclosed for your convenience in responding. Should you decide to participate, I will mail the questionnaires to you with appropriate postage for their return. The time commitment involved would be approximately two hours. All information will remain confidential.

Should you have questions about my research please feel free to contact me at 422-3930 (work) or 876-8728 (home), or contact my faculty advisor Dr. Robert F. Rodgers at 422-7700. I would sincerely appreciate your assistance and hope to hear from you soon! Thanks for your consideration of this request!

Sincerely,

Marcia B. Taylor
Instruction Letter to Second Group of Participants

Thank you for your willingness to participate in my research! Enclosed you will find the consent form, the two questionnaires, a page requesting demographic information, and an envelope in which to return the materials. Please keep your copy of the consent form.

Both questionnaires ask for your perspective on a variety of topics. It is important that you express your own views and the reasons that make up your perspective. It would be helpful if you would complete the questionnaires in the order in which you find them in the envelope. If possible, please try to complete both questionnaires in one sitting. Please complete the page entitled "Additional Demographic Information" after you have finished the questionnaires.

You will notice that the return envelope shows my campus address. (If your address implied that you did not come to campus regularly, I enclosed a stamped envelope.) Please drop the completed materials in any campus mail slot. I would appreciate having the questionnaires back in two weeks from the date you receive them. If this is problematic please give me a call.

Should you have questions about the study or the questionnaires please contact me at 876-8728 (home) or 422-3930 (work), or contact my faculty advisor Dr. Rodgers at 422-7700. Thanks again for your assistance!

Sincerely,

Marcia B. Taylor
APPENDIX D

(Graduate Students)

ADDITIONAL DEMOGRAPHIC INFORMATION

In a study of college students' perspectives on learning it is important to gather information regarding the students' family background. For that reason you are requested to fill out the questions below. The information will be used to define socioeconomic categories of persons who completed the questionnaires. This is done to determine if perspectives differ according to socioeconomic status. This information will be coded and used anonymously. Again thanks for your help!

Please answer the questions as they relate to your job if applicable. If you have not yet established yourself in an employment field, please answer the questions as they relate to your father. If you are not currently employed but have been established in an employment field before, use your previous job to answer the questions.

What kind of work do you do?

(Examples: electrical engineer, stock clerk, farmer, homemaker)

What are your most important activities or duties?

(Examples: kept account books, filed, sold cars, finished concrete)

What kind of business or industry is this?

(Examples: TV and radio mfg., retail shoe store, Government, farm)

Are you: CIRCLE ONE
an employee of a PRIVATE company, business or individual for wages, salary or commission? .................. PR
a GOVERNMENT employee (federal, state, local)........ GOV
self-employed in OWN business, professional practice or farm........................................... OWN
own business incorporated........................................ INC
working WITHOUT PAY in a family business or farm..... WP
ADDITIONAL DEMOGRAPHIC INFORMATION

In a study of college students' perspectives on learning it is important to gather information regarding the students' family background. For that reason you are requested to fill out the questions below. The information will be used to define socioeconomic categories of persons who completed the questionnaires. This is done to determine if perspectives differ according to socioeconomic status. This information will be coded and used anonymously. Again, thanks for your help!

Please answer the questions as they relate to both of your parents jobs if applicable.

What kind of work does your father do? What kind of work does your mother do?

(Examples: electrical engineer, stock clerk, farmer, homemaker)

What are your fathers most important What are your mothers most important activities or duties? activities or duties?

(Examples: kept account books, filed, sold cars, finished concrete etc.)

What kind of business or industry is this? What kind of business or industry is this?

(Examples: TV and radio mfg., retail shoe store, Government, farm.)

Is your father: (Mark One) Is your mother: (Mark One)
an employee of a PRIVATE an employee of a PRIVATE
comp any, business or company, business or
individual for wages, individual for wages,
salary or commission?.. salary or commission?...
PR PR
a GOVERNMENT employee a GOVERNMENT employee
(federal, state, local). (federal, state, local).
GOV GOV
self-employed in OWN self-employed in OWN
business, professional business, professional
practice, or farm...... practice, or farm......
OWN OWN
incorporated.......... incorporated............
INC INC
working WITHOUT PAY working WITHOUT PAY
in a family in a family
business or farm..... business or farm.....
WP WP
APPENDIX E

THE OHIO STATE UNIVERSITY

CONSENT FOR PARTICIPATION IN
SOCIAL AND BEHAVIORAL RESEARCH

I consent to participating in (or my child's participation in) research entitled:

The Development of the Measure of Epistemological Reflection

Dr. Robert F. Rodgers or his/her authorized representative has
(Principal Investigator)
explained the purpose of the study, the procedures to be followed, and the expected duration of my (my child's) participation. Possible benefits of the study have been described as have alternative procedures, if such procedures are applicable and available.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Further, I understand that I am (my child is) free to withdraw consent at any time and to discontinue participation in the study without prejudice to me (my child). The information obtained from me (my child) will remain confidential unless I specifically agree otherwise by placing my initials here _____________.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

Date: ______________ Signed: _______________________

Signed: ____________________ Signed: ____________________

(Principal Investigator or his/her Authorized Representative) (Person Authorized to Consent for Participant - If Required)

Witness: ____________________

RS-027 (Rev. 12/81) -- To be used only in connection with social and behavioral research.
APPENDIX F

PRELIMINARY MANUAL
MEASURE OF EPISTEMOLOGICAL REFLECTION

January 2, 1983
Taylor & Porterfield © 1983

114
The Measure of Epistemological Reflection (MER) is a method of assessing development on the Perry scheme. Although the assessment is based on Perry's original scheme (1970), the instrument and rating process are different from other Perry measures currently in use. Thus it is essential that raters thoroughly understand the design of the instrument and the specific rating procedures.

The MER is organized such that each series of questions focuses on one domain relevant to the Perry scheme. The domains (decision-making, learning, instructors, peers, evaluation, and truth) and their corresponding question number are listed across the top of the rating sheet. Each series of questions begins with an attempt to focus the respondent's thinking in the domain, followed by specific probe questions to elicit the respondent's justification for his/her thinking. The rating manual is organized by question to match the instrument. The degree of specificity of the manual and rating procedures is intended to heighten the accuracy of rating. The manual thus contains a specific coding strategy and, for each question, descriptions of broad categories of the Perry scheme, each position within each category, and example responses collected from pilot data.

Protocols are coded to eliminate demographic and identification information. The rating sheets are coded accordingly. Rating of the MER is to be done by question to avoid bias in rating a complete instrument. Thus the rater should separate all protocols to groupings by question. Ratings of Question One for all protocols should be completed prior to beginning Question Two, etcetera. Care must be taken to keep protocol and rating sheets in the same order to insure recording of each rating on the appropriate rating sheet.
CODING STRATEGY

Unit of Analysis

The focus of rating is on the respondent's reasoning structure or explanation of why he/she thinks what he/she expresses. Reasoning structure includes:

a) reasons for preferences chosen or ideas expressed
b) evidence/support/opinions provided for or against a preference chosen or idea expressed.

The content of responses for each question provides clues to Perry position and can be utilized as starting points in position assessment. However, the reasoning structure must serve as the sole basis for final assignment to a position. Reasoning structure may be concentrated in one section of a question, or spread throughout. In each question the initial question of the series is designed to focus the respondent and is not to be used in rating since the content of the response does not necessarily distinguish between positions. The follow-up questions in each series are the questions most likely to reveal reasoning structure.

Coding Categories

Categories for coding are the following:

| Position One  | 1 |
| Position Two  | 2 |
| Position Three| 3 |
| Position Four - Adhering | 4-A |
| Position Four - Oppositional | 4-O |
| Position Five  | 5 |

Categories correspond to the original Perry scheme. The positions beyond Five are included under the category of relativism since this rating scheme is based on the belief that no structural change occurs beyond Position Five.

Assigning Responses to Categories

1. Read the entire response for each domain. Identify on the basis of section one above the reasoning structure contained within the response.
2. Using the manual, identify the general area of the scheme in which the reasoning structure falls, i.e. Dualism, Transition, Relativism. Proceed to read the position descriptions within that broad category, including the examples. Match the response to the position and example it fits most closely.
3. Record the example number on the rating sheet under the appropriate question, across from the position assigned. If the protocol matches a position but not an example, list the question, position and a zero. (Example: For Question 1, Position 4, list 1.4.0.)
4. In the case of a lack of "fit" into categories:
   A. If steps 1 and 2 do not result in finding a match in one of the positions, review the domain protocol again for the reasoning structure within the response. Inferences regarding reasoning structure are limited to guessing what the respondent meant, a meaning the respondent would recognize as his/her intended meaning rather than inference about underlying motives.
   B. If the reasoning structure can be identified but does not seem to fit existing categories, it may be representative of transition. See scoring algorithm for instructions. If necessary define a new category based on the reasoning structure. Explore existing categories to determine where the new category fits in the sequence. Note the new category and definition on the rating sheet.
   C. If insufficient information exists to identify the reasoning structure using a reasonable level of inference, the response is not ratable. Note this on the rating sheet by indicating 0.0.0.
   D. Omissions, indecipherable responses, and the like do not necessarily result in an unratable domain protocol. As long as sufficient response is present to identify the reasoning structure, the domain protocol is ratable.

Scoring Algorithm

1. Each domain protocol receives one number which represents the position the modal reasoning structure matches. This number is the middle number of the three digit number listed by each example. If in the case the response does not match the examples or position descriptions in the manual, the response may be representative of a transition. If at least one third of the reasoning structure appears to be in transition, indicate this by a plus or minus after the position number. (Example: modal four with one fourth toward five = 4+). Note: It is reasonable to expect that the same respondent could be at one level in one domain and at another in another domain. However, this scoring algorithm assumes that within a domain the subject is generally in one position, possibly opening toward the next. Modal reasoning structure would remain in the earlier position. If the "opening" is extensive enough to indicate modal reasoning in the next position, that position is assigned. If one fourth of the reasoning appears to remain at the earlier level, the minus sign is used along with the modal number. In the case that more than one level appears in equal strength, list both levels in parentheses.

2. In order to determine the total protocol rating, find the dominant rating of the six domain protocols. List any other rating that occurs in two domain ratings in parentheses. (Example: Of the six domain ratings, if four are Position Two and two are Position Three, the TPR = 2(3).) If the ratings are evenly split between two stages, list the ratings separated by a dash. (Example: With three Position Three ratings and three Position Four ratings, the TPR = 3-4.) Domain ratings that appear only once are not included in the total protocol.
| POSITION 1 |       |       |       |       |       |       |
| POSITION 2 |       |       |       |       |       |       |
| POSITION 3 |       |       |       |       |       |       |
| POSITION 4A |   |       |       |       |       |       |
| POSITION 40 |   |       |       |       |       |       |
| POSITION 5 |   |       |       |       |       |       |

UNSCORED DATA
(Note reason as per manual)

NEW CATEGORY
(Please explain)
QUESTION ONE

Question one focuses on the construct of choosing or decision making. Reasoning structures regarding making choices or decisions vary across positions in the Perry scheme. General and position changes are described below.

DUALISM

In Positions One and Two the world view is absolute, characterized by either-or, right-wrong categories. Since this is coupled with the notion that truth is known, choosing becomes a matter of finding the true or right choice. Since authority figures are believed to know the truth, they have a strong influence on the decision making process. Alternatives are not viewed as real options to consider. The process of choosing is generally described as one of finding out which option is the right one, utilizing simple criteria such as happiness and satisfaction. There is no introspection on decisions, and external circumstances are responsible for what happens.

Position One

Position One is an absolute perspective in its purest sense. The person in this structure cannot imagine that more than one alternative could exist or that a question could have more than one answer. The automatic assignment of alternatives to right and wrong results in a conflict-free world. Choices do not exist in the sense of thinking about options since there is no detachment or introspection. One just does the right thing, that which authority espouses.

1.1.1 My major alternatives to making a choice about school were my high school guidance counselors and now my academic advisor.

1.1.2 My last decision was where to go to school. My parents said I should come to school here, and so there weren't any alternatives for me to think about.

Position Two

Alternatives are recognized in Position Two even though they are not yet real options. The availability of alternatives is seen as confusing in the attempt to determine the right one. Alternatives exist to help the student in learning to find the right answers. Authorities still have the answers and external circumstances are still responsible for what happens.

1.2.1 The last time I made a major decision was the beginning of this quarter and that was to set a specific goal academically for Spring quarter. Alternatives are comforting only until depression sets in because you know that only one alternative is right for you. I chose the one that was best for me. I knew that if I chose it I would be happy. I would have an opportunity to gain peace with
myself academically.

1.2.2 It was a carefully planned decision based on facts and advice from others. Either work for the rest of my life when I got out of school or go to college. I felt I was not ready to work for the rest of my life. But I had heard that college was a good learning experience and it is lots of fun too. I took a great deal of advice from elders and then based my decision on that. A college education. A better job if I went through college. More money after getting a college degree. My parents are paying for my college education. It seemed the right thing to do.

1.2.3 Which major to take was probably my last major decision. I felt like I had to make the decision on which one before I entered by college...then during the summer I got a job and by the end of the summer I was the night manager. So I decided management was right.

TRANSITION

In Positions Three and Four a transition occurs in the world view and thus the reasoning sstructure regarding choices. The acceptance of uncertainty as legitimate characterizes these two positions. In Position Three uncertainty is acknowledged but temporary. Since answers will be known in the future, dependence on authority remains. As uncertainty arises more and more, it seems to become commonplace and thus the structure is altered to "there are no answers". There are some areas in which certainty still exists but for the most part things are uncertain. The person at Three utilizes a process to insure that things will work out and at Four says all choices are equal in value so there is no right choice.

Position Three

In Position Three choosing is still a matter of finding the right answer but is complicated by the notion that all truth is not known yet. Authorities still know the truth that is known, and thus maintain much of their influence. Alternatives are still confusing since it is even harder to find the right one. The Position Three person generally describes choosing as engaging in the right process or formula for deciding, with the notion that if the process is right it will work out in the end.

1.3.1 I chose the one which had the components included in the major, or could be included. (1. satisfaction that I am helping, 2. possibility of coaching in high school, and 3. money). Important things were 1. To be happy in life, and to be able to help someone who may have a problem with their speech, 2. That I could be hired as a speech specialist professional by a school district and be eligible to coach high school baseball, 3. That I could be financially comfortable.
1.3.2 My last major decision was what to major in. I was very confused about what major to take since I have taken so many courses here. My advisor was very helpful. He told me to list the pros and cons of each major. That way I could find which major had the most pros, and choose the one that was right for me.

Position Four

Position Four brings with it the notion that all truth is not known. Uncertainty exists as legitimate for the first time. The initial response is that there are no answers, no criteria for evaluating alternatives. Authorities opinions are no better than anyone else's, reducing their influence on choices significantly. Alternatives are now viewed as fun since they represent freedom to explore. This is not an uncomfortable feeling since all alternatives are equal; none are better than the others, making choosing less of a risk. Introspection becomes possible in the sense of thinking about one's own opinion. The actual choosing, however, may be described as frustrating since there are no criteria on which to base a choice. The Position Four person wants to expand rather than narrow by making firm choices. The oppositional four is more pessimistic about the lack of criteria for making choices and is likely to express a cynicism when talking about making choices.

1.4.1 I wanted to study in an area that I know could benefit and help others. Anthropology (cultural) studied the interrelations of various cultures and I thought that if I gained a more in-depth perspective in inter-cultural interactions then maybe I could help people in some way. Also I'd like to be interested in an area that allowed me to keep on learning about the world.

1.4.2 I liked these alternatives. If in a year or so I decided that I didn't like dentistry I could look at one of my other alternatives.

RELATIVISM

The relativist's perspective on choosing is one of evaluating alternatives and supporting one's choice of alternatives with evidence. Truth is relative, that is it exists in a context. Authority rests on the basis of experience and expertise in that context. The relativist recognizes that even though diversity exists, better or worse choices can be made. Since detachment and introspection are possible for the first time, a more serious exploration of alternatives occurs. The description of the process of choosing is an account of options, reasons to support or not support each one, and why this together supports the choice made (or the choice prefered if not yet made).

Position Five

Alternatives are no longer right or wrong. Choices are a routine part of life, although individuals may not be ready to face the implication of such a relativistic world for themselves. Choices are based on one's own articulated values or reasons instead of external circumstances or
authority figures. The person at Position Five explores alternatives and makes decisions on the basis of evidence and can articulate support for his/her choice.

1.5.1 Whether to take a leave from my job and return to school, or to continue on the job and try to pick up night classes. There were quite a number of alternatives and consequences for each of my major choices. I tried to consider all of them. However, soon you realize that no matter which decision you make there will be any number of consequences both good and bad. The main thing then becomes making a choice that you can feel good about. I made the choice to come back to school because I knew that I wanted to finish my degree. The other choices and consequences were still there, but you have to make a decision to do one thing or another sometimes.

1.5.2 I felt good about having so many options, but frustrated that each carried with it several drawbacks. C'est la vie. None of the alternatives were perfect but a choice had to be made. You reach a point when you must take action.
QUESTION TWO

Question two focuses on the role of the learner. As reasoning structures change across Perry positions the role of the learner is perceived differently.

DUALISM

In Positions One and Two truth is known and authorities serve as the holders of knowledge. The role of the learner is to obtain this knowledge. As a result the learner sees himself in a rather passive role of recipient. The instructor or textbook gives the knowledge and the learner must remember and be able to repeat the right answers. The dualist will be rather confused or frustrated with abstract learning experiences where the right answers are not clear. Advice on how to learn will be in the form of paying close attention and getting accurate notes on what authorities say so that the right answers are not missed.

Position One

Knowledge or things learned are those things which the external authority dictates as the truth. The Position One person cannot imagine there being more than one answer to a question. There is no question that authorities are right.

2.1.1 I prefer classes that focus on facts. The best way to learn is by the teacher presenting the facts so you know what they are and what you need to know.

2.1.2 Listen carefully to the teacher and do exactly what they say to do. You have to do what they tell you to do to learn the subject.

Position Two

Learning is still guided by the search for the right answer but sometimes authorities want students to work to find the answer themselves. The role of the learner is to work hard at doing what the teacher wants. It is at this position also that the notion of some authorities being poor authorities appears. Those who do not seem to know the truth are labeled as dumb authorities. This is seen as an exercise that is good for the mind and is not a genuine search for the truth since authorities already know the truth.

2.2.1 I would tell them to pay good attention in class, take notes (a lot of notes), ask questions when something is not clear and go to class everyday.

2.2.2 The key to doing well in college courses is to stay current with school work and understand every why, what, where, and how of everything that's said in a college course.

2.2.3 To learn means to gain some type of knowledge in a field that is unfamiliar or something that is being taught and you receive that
subject. When I understand what I am doing and am graded high when being evaluated by someone. I would tell them to go to class and do the homework. Try to do whatever they tell you so you can get it right. The key to doing well is to try hard and don't let someone influence you to do the wrong thing.

TRANSITION

Learning becomes increasingly confusing in Positions Three and Four. This confusion is due to the ever increasing body of knowledge in which uncertainty exists. The realization that all truth is not known is tempered initially by the belief that it will be known in the future. As more areas are revealed as uncertain, uncertainty is accepted as the "way it is". The role of the learner shifts from one of a quantitative interest to an interest in the process. The role of the learner initially is the same as in earlier positions for concrete areas of knowledge. For the uncertain areas the learner focuses on the processes for finding the answer. When uncertainty becomes common, the focus remains on the process but not necessarily to find the right answer. The process becomes one of thinking the way teachers want students to think since the right answers no longer exist.

Position Three

In Position Three the notion that some of the truth is unknown at present even to authorities complicates finding the right answers. Abstract learning contexts like philosophy are likely to be viewed as areas where the truth is yet to be discovered. The learners role remains one of receiving the answers in concrete contexts, and shifts to one of learning along with the authority how to find the answer in uncertain contexts. Advice about the latter is in the form of going through the right steps to insure that the truth will be discovered in the future.

2.3.1 Try to attend class regularly. If you have any problems or questions talk to the instructor. Never fall behind in a class. Don't cram the night before a test. Take courses that you like and will enjoy.

2.3.2 The key word in course work is to put your effort into it - go in open minded, prepared for anything. If you get discouraged, which everyone does, don't give up. Realize that college courses are hard. This is higher education and you get out of it what you put in.

2.3.3 First of all, don't hesitate to speak up and ask questions of professors and advisors. Second, talk to other students and try to find who the good profs and TA's are. Third, spend much time studying. I think the key to success is learning how to budget your time efficiently. Also learning good study habits.
Position Four

Uncertainty arrives full force in Position Four, altering the role of the learner significantly. The world view incorporates uncertainty as the way the world is, a sort of "there are no right answers" notion. Learning is no longer the search for answers but rather a process of thinking. The focus is on how to learn the method of thinking the teacher wants. Oppositional four's are likely to advise playing the game and assert independence by argument since teachers are no longer a source of right answers. Adhering four's are likely to advise trying to learn the method and trying to think independently.

2.4.1 It doesn't matter to me. There's not much point in focusing on facts since no one really knows what the facts really are. Even when concepts are discussed there really isn't any right answer. Everyone has their own ideas about facts and concepts.

2.4.2 I enjoy concepts. People's ideas are really interesting to me. Everyone looks at things with their own concepts in mind. Hearing different people's ideas about a subject can help you see other ways to look at things. Talk a lot in class and take advantage of the opportunity to hear other's ideas that may be different than yours.

RELATIVISM

The relativistic view of truth is regarded as existing in a context. The acquisition of knowledge is based on experience, expertise, and competence in exploring a given context. Authority is based on this same process and authorities who are perceived as experts can help in clarifying ideas for the student. However, the learner's role is to engage in a genuine search for evidence to support ideas. Advice is likely to focus on having a genuine interest in the course topic and critical examination of materials and opinions.

Position Five

The learner's role is to actively engage in genuine inquiry, taking responsibility for one's own learning. The teacher is seen as a source of guidance who can assist with clarification of ideas but who may not be an authority on the topic under exploration.

2.5.1 Getting involved with other students in your class is good also because you can set up study groups and help each other. Also realize that even though some areas of knowledge might seem boring or incomprehensible there really is a lot of relevance in any area of study and nothing is impossible as long as you are motivated. If one can develop a high degree of motivation and sustain it for awhile then that is a great achievement in itself.

2.5.2 Classes that focus on ideas and concepts challenge a person to explore the possibilities and integrate the knowledge into the person. I'm not sure there is a way to best succeed. I feel so much
depends on what the individual sees succeeding to be.

2.5.3 I prefer concepts because there is more opportunity for "give and take" and the consideration of opposing views.
QUESTION THREE

Question Three focuses on the role of the instructor in the learning process. As the learner's view of truth changes his/her expectations of the instructor change accordingly.

DUALISM

In Positions One and Two the truth is assumed to be known, and known to authorities. Instructors are automatically regarded as authorities whose role it is to pass on knowledge and truth to the learner. Since the dualistic perspective is an absolute one, concrete structured teaching methods are preferred. Dualistic learners prefer presentations by the instructor which clearly spell out the right answers. Clear and specific assignments are preferred. They tend to see class discussion as less useful, since their peers are not authorities. The relationship expected is one of approval when the learner demonstrates acquisition of the right answers.

Position One

The individual at this position does not recognize uncertainty or alternate possibilities. Only one right answer exists for each question and that answer is whatever authority dictates. What authority says is innocently accepted without question. The role of the instructor is to give the right answers in as clear a method as possible.

3.1.1 I like the freedom to ask questions anytime I am confused.

3.1.2 I like for teacher to talk. We're there to listen to them, not interrupt them. They will tell you what you need to know.

Position Two

Authorities still know the right answers. Although it is preferable that they simply present them, students in Position Two accept that it is good for them to look for the answers themselves. If the instructor wants this to occur he/she should be specific as to how to proceed and provide the right answers eventually.

3.2.1 Getting the class(student) involved in every lecture. Class participation keeps a class alert and is fun - which is the easiest way to learn. Nothing beats learning and learning in a fun way. Class participation is beneficial because students can speak out loud and by doing this anything unclear can be cleared using more everyday related events. The type of relationship with an instructor that would help me learn best is one in which the teacher comes down to the student level (in a sense) themselves. If they talk and act in a "college type" way, students would feel closer to the instructor and would become more interested in what the teacher's about - the subject they're teaching.
3.2.3 Because he talked a lot and used plenty of visual aids and even had extra sessions for students who needed help.

TRANSITION

The expanding role of uncertainty alters the role of the instructor as well as the student expectations of instructor-student relationships. Authorities are no longer regarded as having all the answers. Initially authorities role is to help the student learn the way to find the answers that are unknown at present. Developing a relationship with the instructor is perceived as helpful to learning the process. This is even more important when uncertainty prevails and learning is simply learning a way to think. At this point authorities are relieved of responsibility of having any answers since there are no right answers.

Position Three

In Position Three instructors maintain a great deal of their authority. However, the learner has discovered that some truth is not known at present. As the learner's focus turns to finding a process to use to discover the unknowns, the instructor's role alters from giving the answers to helping the students learn the process of finding them. The learner is more open to abstract teaching methods than before but is still likely to see student participation as having minimal significance to learning. Preferred teaching methods are those that assist in learning the process for thinking and searching. The instructor-student relationship is beginning to expand since the student sees getting to know the instructor as part of this learning process.

3.3.1 TA's seemed to care a lot more than profs do, they get to know you on a more personal basis maybe because they've gone through the same things we have more recently.

3.3.2 I like an instructor who wants to get to know you - and puts forth effort - he or she has to want you to get something out of a class. This provides for a good atmosphere for learning.

3.3.3 If you liked the teacher you would pay more attention and therefore learn more. If he cracked a couple jokes you were more relaxed etc. Friendly and easy to talk to and understand. May crack a few jokes but get through the material at the same time.

Position Four

In Position Four the learner's expectations begin to focus on the relationship. Since there are no answers now, the instructor is no longer an expert. The learner views his task to be learning a method of thinking, which the instructor models. However, greater interaction is expected and preferred both with the instructor and peers. Exploring ideas and sharing opinions is regarded as important and should occur since the teacher no longer has the answers. The oppositional four does not accept the importance of learning a method of thinking but rather
views learning as a game to be played. The instructor is perceived as having no more knowledge than the students and oppositional students rebel against the instructor and what they perceive as the game of learning.

3.4.1 Dr. Smith is an instructor who is an expert in his field, he was a very effective instructor in my book, but most of all he had a sense of humor. I learned that a person should be open-minded to other people's beliefs, especially when it comes to the evolution of man. A relationship that is give and take. You can learn from the instructor and the instructor can learn about you and from you.

3.4.2 The ones who through their approach directed their message and modified their message to the needs of the students. Visual aids, examples that are relevant to the audience, policy and assignment structure all were examples of what I considered good teaching methods of that particular teacher. How everybody is different and satisfaction to all is a constant challenge that engulfs the precision of specialization, and all corners of hard work. The best relationship would be one in which the instructor could monitor my growth. Communications either verbally or by other means could enhance some of the finer points of education.

RELATIVISM

With the relativistic viewpoint, the possibility of truth existing returns. Truth is viewed as existing in a context with holders of the truth identified as those who have experience, expertise, and competence in searching for the truth in a given context. Learning is viewed as a genuine search for the truth, looking for evidence to support opinions. The role of the instructor is to engage in the search mutually with students. Abstract teaching methods are preferred, as well as mutual interaction with instructors.

Position Five

Student and instructor become partners in a mutual search for truth. Instructors are viewed as competent in given contexts in which they can provide evidence for thinking. Instructors are expected to be encouraging and accepting of critical questions and alternate viewpoints if supported by evidence.

3.5.1 I like classes where the teacher and the students are both able to talk about a subject and exchange ideas. I have enjoyed the classes where the instructor has had some expertise in a particular field and has brought the expertise to the classroom as a means of facilitating discussion among the class members. It is a good opportunity for students to add their own relevant expertise to the thoughts and opinions of the teacher. This makes for a well rounded discussion and sharing of ideas among a group of learners. The teacher has an institutional obligation to bring in information relevant to a given topic, but all share in the obligation to learn
from one another and share ideas.

3.5.2 I prefer there to be an atmosphere of mutual respect. I prefer to be viewed as a person with something to offer rather than a simple recipient of someone else's knowledge. I prefer discussions or at the very least lecturers who identify the issues and then recommend answers based on evidence, perspective and interpretation.
QUESTION FOUR

Question four focuses on the role of peers in the learning process. Their role is dependent on the learner's view of knowledge and the learning process. Changes occur across positions as described below.

DUALISM

In Positions One and Two authorities are viewed as holders of the truth. Since learning is a matter of acquiring the truth, student participation is not viewed as useful. The dualist wants to hear the truth from the instructor, and not be confused by discussion. Although students asking for clarification is seen as positive, discussion is seen as confusing, distracting and often a waste of time. The dualist sees learning directly from the instructor as more beneficial. Relationships among members of the class are not a focus for the dualist.

Position One

Involvement consists of listening to authorities and getting the right answers. Questions should be direct attempts to get the right answers. Additional discussion is seen as a waste of time.

4.1.1 I don't think we should be talking a lot in class. When others begin to talk and ask questions it is really hard to remember what the teacher has said was the correct way to do something. If other students would just listen to the teacher, they could get the right answers instead of wasting time distracting the teacher.

4.1.2 I prefer a class where the teacher talks a lot. If people ask a lot of questions you may not get through the teacher's information, and then you have missed part of the subject.

Position Two

Class participation can be useful if it is part of the exercise of finding the right answers for oneself. The instructor should, however, not allow the discussion to confuse the right answers. This is a danger in class discussion since no one but the instructor knows the answers.

4.2.1 Students do a lot of talking because sometimes you may not understand something but be afraid to ask. Then when someone asks the same question then you both will benefit (advantage). This brings out more questions or in depth study. Disadvantages are that the other person wastes time or confuses you. By questions being asked you will learn more about a subject but someone who doesn't know what you are talking about can confuse you by getting you mixed up.

4.2.2 Very little. Class participation sometimes confuses me. Because I think the instructor can get something across much better than a fellow student.
4.2.3 The more questions one student asks the more people have a chance to find out answers. If the student couldn't become involved, then a lot of questions will go unanswered and not much could be gained from that class.

TRANSITION

Student involvement becomes increasingly acceptable as it relates to learning a process and relationships. The focus on process that accompanies the rise of uncertainty allows for class discussion to be a part of the process. During the transition the process changes from one that insures the truth will be found to one that focuses on learning how to think. Relationships gain increasing importance as authorities are no longer the source of all knowledge.

Position Three

Student involvement is a bit more acceptable in Position Three since even the instructor may not know all the answers. Peer interaction is seen as legitimate as it relates to discovering the process of finding the truth such as in small group discussions.

4.3.1 Where students do a lot of talking. The advantages is that it makes the class a lot more interesting and the disadvantages is that students tend to clown around more. The more degree of student involvement, the more I usually learn. A friendly relationship where everyone is acquainted with each other.

4.3.2 I like classes where I can talk a lot. That way the teacher can tell what I know and also know how to help us to get the right answers.

Position Four

Peers are as legitimate a source of knowledge as the instructor since everything is uncertain. Everyone's opinion is equally valid and the learner prefers high student involvement to share opinions. In addition the Position Four person focuses as much on class relationships as on learning. The oppositional four is likely to be more outspoken regarding their notion that students' opinions are as valid as the instructor's since they are rebelling against the instructor and the game of learning.

4.4.1 It depends on how complex the subject matter. Again its a two way street. By class participation the instructor can get an idea how much subject matter is being comprehended by the class with class participation. You can learn from fellow students viewpoints. Exchange of information. I have been helped out by fellow students before in a class that I found difficult.

4.4.2 I like both depending on the situation. If I'm interested in the
I like to talk, if not I don't. Because I like to talk about what I'm interested in. I'm a fairly social person. The more people talking the more viewpoints of the material are known. Of a subject I don't particularly like I might need to get more viewpoints or information to ever get to like it. Free discussion with open people.

4.4.3 I enjoy classes where we can talk a lot. You can learn just as much from some people as you can a teacher. All teachers do is present ideas, and everybody has ideas.

RELATIVISM

All knowledge is relative (contextual) so authority rests on expertise acquired through searching for the truth. Since learning is a genuine search for truth and peers are a legitimate source of knowledge, class involvement is seen as positive. The learner prefers sharing ideas in a mutual sense with peers and instructors.

Position Five

Classes are a forum for seeking out knowledge and soliciting new perspectives. Participation is an opportunity to critically assess ideas and gather information to further explore ideas. Detachment to explore ideas objectively occurs.

4.5.1 Both, depending on the material and the class. If the material is such that mutual benefit could be gained then class participation is positive. Disadvantages could be that student growth could be hampered by slowing down the show. One in which progress is made either by more communication or by more specialized teaching procedures.

4.5.2 An atmosphere of genuine cooperation rather than competition. Where students could share insights, problems and understanding without a fear of being ridiculed or shut-out.

4.5.3 The sharing of informed opinions. It is not enough to say what one thinks. I want to hear a rationale and evidence.
QUESTION FIVE

Question five focuses on how learning is evaluated. The view of knowledge and roles of instructor and learner affect the learner's notion of how evaluation should be accomplished.

DUALISM

In Positions One and Two the instructor is seen as the authority who knows what is right and wrong. The instructor is also awarded the right to evaluate the learner's progress since the instructor is the best judge of truth. The learner in these positions believes that good grades should automatically follow demonstration of having acquired the right answers.

Position One

The student accepts without question the grade the instructor assigns. Grading is a simple process of the instructor counting the right answers that the student provides.

5.1.1 Teachers are the only people who can evaluate you because they know whether you're doing the right things. Teachers grade you on what you do right.

5.1.2 You have to work hard and put effort into your classes so the teacher can know whether you're doing everything to learn. Teachers are the graders because they know what you learn.

Position Two

Grading is still generally based on the student providing the right answers. In cases where the instructor has emphasized working to find the answers on one's own, hard work and effort should be rewarded. Even though this is not as easy, teachers still know what students deserve. In the case of dumb authorities, grades may not be perceived as right.

5.2.1 I believe that hard work and effort will result in high grades. I think they should be determined by a combination of how much effort was put in (or how much the student cares) and the actual grades and participation in class. The instructors should know the students better and give them the grades. I think people who really try but just can't grasp some concepts should be given some credit. I think instructors should put it upon themselves to see how hard an individual is trying.

5.2.2 Grades are definitely out. I feel you should be graded on improvements, what you now know that you didn't before. And the instructors should evaluate us...College is hard enough without grade pressure. Some student can't study or have good grades, because they are worried about good grades! Grades are for high school and therefore should be left there!
TRANSITION

Grading becomes more confusing as the degree to which authorities know the truth dwindles. Since the focus of learning has shifted to processes so should evaluation of student's learning. Initially, the student is uncertain how the instructor can evaluate work other than looking at quantity, but sees authority as the best way to judge progress. Eventually the student sees evaluation as impossible in a genuine sense.

Position Three

In Position Three issues of grading become confusing as a result of some knowledge being unknown. Since the final authority rests with the instructor, that person is still responsible for assigning grades. Since this can no longer be done on the basis of providing the right answers, the focus shifts to quantity of work and how the student expresses him/herself. The instructor should know the student as a person in order to better understand and evaluate them.

5.3.1 I think hard work and effort is a basis for the result of high marks. The more you put into something, the more you will get out. I think there should be more standardization for classes with lecture and recitation. Some TA's graded quizzes and experiments more strictly than others. Therefore some people got better grades for the same quality work and answers. The prof and TA's should evaluate. Who else would grade us? They are the ones who have our work in front of them.

5.3.2 The first statement to a tee! I believe that students should receive a grade on how much participation and hard work they put into the class. But many classes are set up differently, where only the tests count, not on how much work the students put in. Teaching associates should evaluate students. There are many students who know the material for a test but for some reason don't perform as well as they could. That is why it would be more fair to a student if he or she could also be graded on how much participation and hard work they put into the course.

5.3.3 Hard work and effort (most often). Objective, essay and oral recitation. They should judge by performance and general ability to apply matter. Because those who can't put their ideas into words could do it orally, and those who don't like to face people could do it in writing.

Position Four

Grading becomes totally confusing in Position Four since all knowledge is uncertain. Now no basis exists for grades; the instructor has no more knowledge than the students do. The Position Four person does not see grades as legitimate since no criteria exist to evaluate, and no one
person's opinion is superior to another. Since the focus in this position is on learning how to think, the learner may describe determining grades regarding independent thought on the basis of how well the learner demonstrates thinking the way the teacher wants. Oppositional fours view evaluation as useless since authorities have no more knowledge than students. At best they describe the evaluation process as a game one plays to get grades.

5.4.1 Neither of the above. It doesn't matter what you do some teachers aren't going to grade you on effort. Grades are useless anyway, because there is no standard way to grade. What is the need of having grades anyway?

5.4.2 Hard work and effort is essential with high results in anything. In this competitive world, one has to work hard to be ahead of the next guy. Plus, results is what you're judged by for almost everything today. I think a student should be evaluated, not by exams, because very few of what's on exams will one ever encounter in the real world, but through examples (ie what is your solution to the above problem and how would you justify your solution). There are so many ways to just solve one problem. The student should be involved in the evaluation also - just to hear how they actually feel about the work, etc. Because more "real world" problems could be injected into a student and that's what's needed more of today. Solving the problem outside the classrooms and textbook, not in them. College should be a place to expand one's own mind to develop new and improved ideas, not just the ideas of the books' author, etc. That's just one of 500 million people's ideas or answer to the problems.

RELATIVISM

The relativistic person acknowledges uncertainty, but views knowledge as existing in a context. Anyone who has competitively engaged in the search for truth in a given context can serve as an authority and thus judge others learning. This could include instructors, peers or self. Grades might also be viewed as unnecessary since the search for evidence and ability to support one's opinions is sufficient for learning.

Position Five

If grades are to be associated with learning they ought to be based on whether one's thoughts are adequately supported by evidence. Grades ought to represent an appreciation of the value of one's ideas.

5.5.1 Ideally, I'd like to see more joint evaluation between the student and instructor. This could allow for an evaluation of the students' learning of the course material as well as examining the integration of that material into the student's personal development process.

5.5.2 A basis for evaluation might be on how well one is able to use or
understand in a personal way the knowledge presented in the class. How has the challenge of new ideas made an impact on the student seems to be more important than a regurgitation of facts. Has the knowledge gained been synthesized and personalized to be shared and appreciated by the student will ultimately mean more than a grade or degree.

5.5.3 The extent to which critical thought has gone into my work should be the primary basis upon which I am evaluated. Ideally, the teacher would allow me to offer input into my grade, however, this is not often practical given student/teacher ratios.
QUESTION SIX

Question six is intended to explore the view of knowledge, truth or reality. The degree to which one believes truth exists and how one arrives at the truth are the focus of the changes as described below.

DUALISM

The dualistic person believes that truth exists. Only one perspective on any subject is the true one. Therefore conflicting notions are dealt with as though one of them is not there or is not true. The process of dealing with conflicting versions of the truth is a simple one of deciding which one is right.

Position One

The categorization of conflicting notions into right and wrong is a simple process. One idea is simply wrong or an explanation that just seems different. The one assigned to the right category is likely to be closest to what other authorities in the student's life have espoused. No reason need be given for why one authority has the wrong answer.

6.1.1 No. There is only one explanation for an historical event or scientific phenomena. They are probably explaining different parts of an event which makes it seem different. I would believe the explanation that is right. Once you know which one is right you can be sure that is the one to believe.

Position Two

The categorization of conflicting notions into right and wrong remains a simple process. The only difference is that the explanation for there being two perspectives is that one of the authorities is dumb and doesn't know the right answer.

6.2.1 Yes. One is wrong. Some instructors don't know their material as well as they should. Poor instructors give out wrong information. I would believe the explanation that the instructor who knows his field gives. For example, some instructors say man evolved from apes. But everyone knows that God created man. I would believe the right explanation.

TRANSITION

In Positions Three and Four the notion that truth definitely exists in all areas begins to change. The person initially discovers that although some truth is known, other areas seem to not be known yet. As more and more areas fall into the latter category, the thinking shifts to the notion that everything must be uncertain. The Position Four person concludes that truth is not really known, everyone just decides for themselves.
Position Three

In Position Three it is possible for the truth to be disputed since some things are yet to be known for sure. It becomes difficult however to define how the disagreement can be resolved. Should one expert appear to have a better process for finding the truth, he/she may be presumed to be right, at least until the truth is really discovered. The Position Three person believes that the truth will be known for sure in the future.

6.3.1 Maybe. It could be that the real explanation is not known yet and they are just guessing. Or one may know the real explanation. If neither instructor really knows the real explanation, you don't know which one to believe. You just have to wait and see. No, you can't be sure until someone finds out which one is right.

Position Four

In Position Four it is even more possible for the truth to be disputed. Since no one knows the truth there can be an infinite number of explanations for any notion. Discrepancies cannot be decided since no criteria exist to decide. Everyone has a right to his/her own opinion and no one has the right to call another person's opinion wrong.

6.4.1 They could be, but there could be lots of explanations for the same event. Everyone has a different opinion of what happened. Since all the explanations are someone's opinion, you can believe any of them. No one really knows for sure, so what's the point of choosing one? No - you can't ever be sure because there aren't any ground to choose one or the other - nobody really knows what to believe.

RELATIVISM

In relativism the truth can be known again but on a basis different from that of the earlier positions. Truth is now based on experience and expertise in exploring a particular area of knowledge. Although truth may not be known in an absolute sense, it can be decided on the basis of the evidence at hand.

Position Five

Persons in Position Five will analyze both views in terms of the evidence and exhibit a more complex understanding of the issues involved. Judgements of better and worse arguments exist and decisions are based on evidence rather than authority, norms, or personal opinion.

6.5.1 Yes, it's possible that one has more sound evidence than the other. Work through both explanations, looking at the degree to which they explain the situation and potential discrepancies in each explanation. It could be that both are correct and are explanations from different viewpoints. I would choose to believe the
explanation that in my judgement was the more adequate of the two. It depends on what being "sure" means. I can choose one and be sure it is the one I think is most adequate. Inherent in making my choice is a decision to believe that one. I guess one could always wonder which explanation to believe. Even after choosing one, I might contemplate the advantages of the other.
The Measure of Epistemological Reflection (MER) is a method of assessing development on the Perry scheme. Although the assessment is based on Perry's original scheme (1970), the instrument and rating process are different from other Perry measures currently in use. Thus it is essential that raters thoroughly understand the design of the instrument and the specific rating procedures.

The MER is organized such that each series of questions focuses on one domain relevant to the Perry scheme. The domains (decision-making, learning, instructors, peers, evaluation, and truth) and their corresponding question number are listed across the top of the rating sheet. Each series of questions begins with an attempt to focus the respondent's thinking in the domain, followed by specific probe questions to elicit the respondent's justification for his/her thinking. The rating manual is organized by question to match the instrument. The degree of specificity of the manual and rating procedures is intended to heighten the accuracy of rating. The manual thus contains a specific coding strategy and, for each question, descriptions of broad categories of the Perry scheme, each position within each category, reasoning structures within each position, and example responses collected from data collected in the initial study of the MER.

Protocols are coded to eliminate demographic and identification information. The rating sheets are coded accordingly. Rating of the MER is to be done by question to avoid bias in rating a complete instrument. Thus the rater should separate all protocols to groupings by question. Ratings of Question One for all protocols should be completed prior to beginning Question Two, etcetera. Care must be taken to keep protocol and rating sheets in the same order to insure recording of each rating on the appropriate rating sheet.
CODING STRATEGY

Unit of Analysis

The focus of rating is on the respondent's reasoning structure or explanation of why s/he thinks what s/he expresses. Reasoning structure includes:

a) reasons for preferences chosen or ideas expressed
b) evidence/support/opinions provided for or against a preference chosen or idea expressed.

The content of responses for each question provides clues to Perry position and can be utilized as starting points in position assessment. However, the reasoning structure must serve as the sole basis for final assignment to a position. Reasoning structure may be concentrated in one section of a question, or spread throughout. In each domain the initial question of the series is designed to focus the respondent and is not to be used in rating since the content of the response does not necessarily distinguish between positions. The follow-up questions in each series are the questions most likely to reveal reasoning structure.

Coding Categories

Categories for coding are the following:

Position One 1
Position Two 2
Position Three 3
Position Four - Adhering 4-A
Position Four - Oppositional 4-0
Position Five 5

Categories correspond to the original Perry scheme. The positions beyond Five are included under the category of relativism since this rating scheme is based on the belief that no structural change occurs beyond Position Five.

Assigning Responses to Categories

1. Read the entire response for each domain. Identify on the basis of section one above the reasoning structure contained within the response.
2. Using the manual, identify the general area of the scheme in which the reasoning structure falls, i.e. Dualism, Transition, Relativism. Proceed to read the position descriptions within that broad category, including the reasoning structures and examples. Match the response to the position and reasoning structure it fits most closely.
3. Record the reasoning structure number on the rating sheet under the appropriate question, across from the position assigned. If the protocol matches a position but not a reasoning structure, list the question, position and a zero. (Example: For Question 1, Position 4, list 1.4.0.)
4. In the case of a lack of "fit" into categories:
   A. If steps 1 and 2 do not result in finding a match in one of the positions, review the domain protocol again for the reasoning structure within the response. Inferences regarding reasoning structure are limited to guessing what the respondent meant, a meaning the respondent would recognize as his/her intended meaning rather than inference about underlying motives.
   B. If the reasoning structure can be identified but does not seem to fit existing categories, it may be representative of transition. See scoring algorithm for instructions. If necessary define a new category based on the reasoning structure. Explore existing categories to determine where the new category fits in the sequence. Note the new category and definition on the rating sheet.
   C. If insufficient information exists to identify the reasoning structure using a reasonable level of inference, the response is not ratable. Note this on the rating sheet by indicating 0.0.0.
   D. Omissions, indecipherable responses, and the like do not necessarily result in an unratable domain protocol. As long as sufficient response is present to identify the reasoning structure, the domain protocol is ratable.

**Scoring Algorithm**

1. Each domain protocol receives one number which represents the position the modal reasoning structure matches. This number is the middle number of the three digit number listed by each reasoning structure. If in the case the response does not match the position descriptions or reasoning structures in the manual, the response may be representative of a transition. If at least one fourth of the reasoning structure appears to be in transition, indicate this by a plus or minus after the position number. (Example: modal four with one fourth toward five = 4+). Note: It is reasonable to expect that the same respondent could be at one level in one domain and at another in another domain. However, this scoring algorithm assumes that within a domain the subject is generally in one position, possibly opening toward the next. Modal reasoning structure would remain in the earlier position. If the "opening" is extensive enough to indicate modal reasoning in the next position, that position is assigned. If one fourth of the reasoning appears to remain at the earlier level, the minus sign is used along with the modal number. In the case that more than one level appears in equal strength, list both levels in parentheses.

2. In order to determine the total protocol rating, find the dominant rating of the six domain protocols. List any other rating that occurs in two domain ratings in parentheses. (Example: Of the six domain ratings, if four are Position Two and two are Position Three, the TPR = 2(3).) If the ratings are evenly split between two stages, list the ratings separated by a dash. (Example: With three Position Three ratings and three Position Four ratings, the TPR = 3-4.) Domain ratings that appear only once are not included in the total protocol.
DOMAINE ONE

Domain one focuses on the construct of choosing or decision making. Reasoning structures regarding making choices or decisions vary across positions in the Perry scheme. General and position changes are described below.

DUALISM

In Positions One and Two the world view is absolute, characterized by either-or, right-wrong categories. Since this is coupled with the notion that truth is known, choosing becomes a matter of finding the true or right choice. Since authority figures are believed to know the truth, they have a strong influence on the decision making process. Alternatives are not viewed as real options to consider. The process of choosing is generally described as one of finding out which option is the right one, utilizing simple criteria such as happiness and satisfaction. There is no introspection on decisions, and external circumstances are responsible for what happens.

Position One

Position One is an absolute perspective in its purest sense. The person in this structure cannot imagine that more than one alternative could exist or that a question could have more than one answer. The automatic assignment of alternatives to right and wrong results in a conflict-free world. Choices do not exist in the sense of thinking about options since there is no detachment or introspection. One just does the right thing, that which authority espouses.

Reasoning Structures

1.1.1 Choice made on the basis of only one choice being perceived as truly available.

Example: I decided to come to OSU basically because my sister goes here. I had Bowling Green, Miami(Oxford), and OSU in mind for college. They were all relatively good schools. I signed up late for Bowling Green, Miami's standards were too high for me, and my sister goes to OSU, that was the big reason.

Position Two

Alternatives are recognized in Position Two even though they are not yet real options. A right choice exists and the choice is permanent. The choice is usually the one that is safest, that has the greatest degree of certainty. Criteria discussed are usually what one likes and what one can and should do. Pro's and cons if discussed, are described in a superficial manner and focus on external circumstances. The availability of alternatives may be seen as positive in terms of not feeling restricted or can be viewed as confusing in the attempt to determine the right one. Alternatives sometimes exist to help the
student in learning to find the right answers. Authorities still have the answers and external circumstances are still responsible for what happens.

**Reasoning Structures**

1.2.1 **Basis for choice is that it is right.**

*Example:* Which major to take was probably my last major decision. I felt like I had to make the decision on which one before I entered my college. Then during the summer I got a job and by the end of the summer I was the night manager. So I decided that management was right.

1.2.2 **Basis for choice is what one likes.**

*Example:* I had to make a decision of going to school for culinary arts or physical therapy. I chose physical therapy. I had many other alternatives to think about as a major. I felt that all of the alternatives I thought about were very interesting and something I would enjoy doing. I chose physical therapy because I like working with the handicapped. Also I like exercising and using my body such as lifting weights, running and various kinds of sports.

1.2.3 **Basis for choice is what one should do (socialized "what is right"). This sometimes takes precedence over what one likes.**

*Example:* Whether to keep my job or go to college. If I had kept my job I would have had enough money to pay for six months rent in a new apartment. But if I went to college I could get a pay increase or a better job. I felt kind of stuck in the middle. I asked myself if I thought I could make it or would like to stay at this job for the rest of my life. I knew that I couldn't support anyone with my salary.

*Example:* Going to college versus playing in a rock band. Attending OSU or not attending anywhere and trying to make a living as a musician. Being a musician seemed much more rewarding at least at first. But college would be more important for my long term future. I chose the one that would be most beneficial to me in the long run. Considerations were my future well-being, money, happiness, etc. Music would give me happiness now and maybe some money, but college would guarantee money in the long run.

1.2.4 **Basis for choice is that authority said to do it.**

*Example:* I had to decide if I wanted to come to OSU or stay in Texas with my family. I liked Texas. I didn't want to start school yet. I'm not sure about my major. I believe I would rather have stayed in Texas, but my parents suggested that I go for a year and see if I would like it. I chose to come to OSU.
TRANSITION

In Positions Three and Four a transition occurs in the world view and thus the reasoning structure regarding choices. The acceptance of uncertainty as legitimate characterizes these two positions. In Position Three uncertainty is acknowledged but temporary. Since answers will be known in the future, dependence on authority remains. As uncertainty arises more and more, it seems to become commonplace and thus the structure is altered to "there are no answers". There are some areas in which certainty still exists but for the most part things are uncertain. The person at Three utilizes a process to insure that things will work out and at Four says all choices are equal in value so there is no right choice.

Position Three

In Position Three choosing is still a matter of finding the right answer but is complicated by the notion that all truth is not known yet. Authorities still know the truth that is known, and thus maintain much of their influence. Alternatives are still confusing since it is even harder to find the right one. The Position Three person generally describes choosing as engaging in the right process or formula for deciding, with the notion that if the process is right it will work out in the end. Criteria discussed with regard to choosing are usually future happiness, success and fulfillment of goals.

Reasoning Structures

1.3.1 Basis for choice is what will work out best in the future in relation to what one likes, is capable of and will be happy with.

Example: To go back to post graduate school or continue working as a general dentist. Alternatives were to possibly buy in with my employer in his practice and work where I want to live, or return to school and thereby lose money and time previously spent developing a practice and leaving where I want to live. As others, I had many second thoughts as to what would be a better decision for the future. Also I was fortunate to have been accepted into my graduate program. I thought - what would be my best choice for the future, not just two years from now. Also I felt I would be happier specializing in what I enjoyed doing the most.

1.3.2 Basis for choice is the one that has the greatest number of positive aspects or important factors.

Example: Whether or not to leave the state for college and get a scholarship or stay in state and have my parents pay for my education. I felt as though I was in a dilemma. It was such a hard decision to make and I wanted to make the right one. I asked myself the good things involved in each alternative and the one with the greatest number along with the most important factors won out.
1.3.3 Choice made on the basis of a process to determine which choice is best for the future.

Example: Whether or not to attend OSU because of the scholarship opportunities and the academic qualities. I had many opportunities to attend numerous colleges on an athletic scholarship and also from my high school grades. Since I had so many choices it was very hard to decide on which college to attend. Many variables had to be considered, and since all the colleges were very good academically that even made it harder. I narrowed down the previous list of 20 to 10 colleges and tried to list the good and bad points of each. After that was done I cut it down to the top five and proceeded to do the same thing except for in more depth. Then I really did some heavy thinking on all the pros and cons again and tried to decide which college could fulfill my goals for the future in academics and swimming.

1.3.4 Choice occurs because of external circumstances even if person really does not prefer that choice.

Example: To decide whether to transfer schools and switch majors. Stay in engineering at an out-of-state school or change majors and transfer to OSU. I was reluctant to leave my old school because I really liked it there. However, I knew that in the long run, it would be better for me to transfer and also to change my major. I reached the last straw with engineering and had been closed out of enough classes to require staying in school an extra semester. I couldn't afford to do that. So my decision was made partly by finances and partly by my changing vocational preference.

Position Four

Position Four brings with it the notion that all truth is not known. Uncertainty exists as legitimate for the first time. The initial response is that there are no answers, no criteria for evaluating alternatives. Authorities' opinions are no better than anyone else's, reducing their influence on choices significantly. Alternatives are now viewed as fun since they represent freedom to explore. This is not an uncomfortable feeling since all alternatives are equal; none are better than the others, making choosing less of a risk. Introspection becomes possible in the sense of thinking about one's own opinion. The actual choosing, however, may be described as frustrating since there are no criteria on which to base a choice. If criteria exist they are hard to prioritize. Information discussed in relation to choosing includes one's own feelings or preferences and some practical information. The Position Four person wants to expand rather than narrow by making firm choices. Thus a choice may not be expressed. The oppositional four is more pessimistic about the lack of criteria for making choices and is likely to express a cynicism when talking about making choices.
Reasoning Structures

1.4.1 All choices are perceived as having equal value so choice becomes a matter of picking the one that is most convenient, workable or valuable at the time.

Example: Career choice: continuing education or accepting an attractive position in the field. Alternatives were a secure, good paying job with a state agency, or a graduate teaching associateship and beginning a new program at a new school. They were both good enough situations that I couldn't lose either way. Weighed pros and cons, chose what I thought was most valuable to my career goals. Considerations were career objective - degree vs. semi-valuable experience, finances - good paying job vs. stipend, and time - accomplish career objective now or delay objective.

Example: Career choice. To remain in pursuit of a full-time career as a musician or find something else to do part or full time. I loved playing the violin, so I have for 15 years, and it was difficult to not feel guilty about considering other choices although my experience since receiving my BM in 1980 showed me that it would be good to diversify my work. I did a lot of reading - books on choosing a career - a lot of talking with friends, some who knew me well and some who hardly knew me - a lot of writing and soul searching and simply day-dreaming! Considerations were my best guess as to what sort of occupation would make me the happiest, a career which would provide me with a variety of challenges and an adequate income.

1.4.2 All choices are perceived as workable but some have negative factors and thus are eliminated.

Example: The decision was to return to graduate school or return to teaching. Really no alternatives, it was either one or the other. There probably were some other career alternatives but not attractive career-wise for me to consider at that time. Well, I didn't want to return to teaching because of a bad experience with my administrators. So my alternative was really to find some other career (journalism) which made my interest expand into media law. I usually (when making a decision) list mentally the positive and negative aspects of the alternatives. Whatever comes up more positive, I usually do.

1.4.3 Choice constitutes a delay of choice due to indecision and a feeling that one can change one's mind later.

Example: What to do after high school. I didn't feel like rushing into college but did want to eventually. I didn't want a full time job but I did need the money. I didn't mind living at home, but didn't want to remain dependent on my parents. I knew I could go to college any time I wanted so I could always try some other things first. I decided a one year delay while living at home and try
finding a job which I could enjoy. Living at home took the pressure off me to earn a living at any job I could dig up. I wanted to maintain my carefree lifestyle as long as possible. I did not want to become an indentured servant like so many workers - ideally I would be self-employed. I needed time to decide what to study at college, when I did start.

1.4.4 All choices are perceived as equal so choice is a matter of picking the one that is best for now, putting the rest off until later.

Example: I decided to postpone going to graduate school to become an anesthesiologist for about five years. Alternatives were remain an RN without additional school, go to graduate school to get a masters in nursing, go to anesthesiologist school after college, or wait for a while to go to anesthesiologist school. I would not mind any of them. I think I would be happy doing any of the alternatives. I chose alternative number four because it was the most feasible and most congenial with my choice of lifestyle. I chose the alternative that most suited my personality and the choice I thought would make me the happiest.

1.4.5 All choices have equal value and a decision is not expressed.

Example: I had to decide between spring quarter in France or at OSU. Spring quarter in France was my first choice because I felt I needed a break in routine. Staying at OSU spring quarter was my second choice. I would learn a lot here which would be beneficial. I've been in school all my life; I need a breath of fresh air. First I thought about the emotional benefits. I feel like I'm going through a routine instead of digging in and learning. Secondly I had to deal with the cost of each alternative. Then I had to find out if I could graduate at the expected time and if I could get credit for going to France.

1.4.6 Choice just happens while alternatives are being debated.

Example: When I graduated from undergrad I had to decide what I should do before I started graduate school the following Fall. Work full or part time, if so where? Take graduate courses that would go towards a degree, if so, what in? I was confused. I didn't want to work in anything other than my field, but you can't get a decent job with a BA in sociology/psychology. I looked at the idea of taking a few courses but since you can only apply ten hours to a degree, I would only go to Winter quarter, then what would I do until Fall? It all seemed a waste of money in the long run, particularly since I didn't know what I was doing, so I decided to find a job. A job just fell into my lap so I felt I should take it. I was lucky when hundreds were trying to find a job in the city.
RELATIVISM

The relativist's perspective on choosing is one of evaluating alternatives and supporting one's choice of alternatives with evidence. Truth is relative, that is it exists in a context. Authority rests on the basis of experience and expertise in that context. The relativist recognizes that even though diversity exists, better or worse choices can be made. Since detachment and introspection are possible for the first time, a more serious exploration of alternatives occurs. The description of the process of choosing is an account of options, reasons to support or not support each one, and why this together supports the choice made (or the choice preferred if not yet made).

Position Five

Alternatives are no longer right or wrong. Choices are a routine part of life, although individuals may not be ready to face the implication of such a relativistic world for themselves. Choices are based on one's own articulated values or reasons instead of external circumstances or authority figures. Criteria for choosing are usually complex, including consideration of one's own needs, one's family, future goals, self-fulfillment, etc. Information is gathered through introspection, soliciting feelings of significant others and observing or acquiring experience in external circumstances. External experts may be consulted but the Position Five person evaluates the experts' advice. The person at Position Five genuinely explores alternatives and chooses the better choice. He/she can articulate support for the choice.

Reasoning Structures

1.5.1 Choice is described as the "better" one on the basis of one's criteria.

Example: I had to make the decision about starting a full time teaching position at a state juvenile institution or staying at an office position that would soon change. I did not like the situation for the teaching job but it would provide full time employment and earn hours toward an 8 year teaching certificate in the field I had obtained education degrees in. Present job was fine but the office was being closed down and I didn't know where I would actually be moved in the company. Pay was low for the amount of work that I did per week. I chose to teach at the institution because I needed to break into the education field in some way and this was the first full time opportunity to come up for me.

1.5.2 Choice is described as "better" one after criteria are established.

Example: To choose between anthropology or psychology for graduate school or to stay home. I was very ambivalent between the two majors but I was very sure I wanted to go on in graduate school and finish my studies. In order to take care of my family (I live 60
miles from OSU) I decided to go part time to study. I chose the subject (psychology) which would allow me to practice where I live, near my family. Considerations were not to neglect my family (husband and children) for my studies, to continue studying no matter what would be the conditions (time, fatigue, money). My priorities in life are to be a good mother, to be a good wife, and to try to fulfill my own personal aspirations, that is to say, to have a career of my own.

1.5.3 Choice is described as "better" one after a description of information collected for use in making the decision.

Example: Selection of my Ph.D. general examination committee members. More than 20 faculty members within the Department plus at least two from each of three other departments could be considered. I first made up a list of the most probable candidates (about 14). I talked personally to all my most likely selections and visited their labs, offices, classes and seminars in which one or more of these were participants. Then after consultation with my advisor I personally discussed the degree, my research and their potential participation in my committee with each of the last six or seven. Finally I selected and asked each of these in preferred order until my committee had been chosen.

1.5.4 Choices are described as "better" one after gathering needed information from experts.

Example: How to invest money for my parents whose health was failing and were forced to give up their business. Alternatives were stocks, bonds, banks, treasury bills, etc. In view of the many alternatives available today for investment, and in view of the ever changing economy, I was needless to say confused, but confident that a secure investment was possible and need not be permanent. Talking to a representative from each of the above mentioned alternatives I was able to determine the best possible course of action at the time.

1.5.5 Choice is described as "better" one in light of calculated risks and balancing priorities.

Example: Quitting my job (and having my husband quit his) so I could return to school. Alternatives were to keep jobs and forget school, keep jobs and commute to a nearby school, quit jobs and attend the program of my choice. The first alternative was comfortable but not entirely satisfying. Time seemed right for a change, but secure in present situation (friends, location). The second was unattractive because of half-way approach to school. The third was frightening but not implausible, also exciting. I looked at "worst case" that could result from any of the choices - decided which "worst case" I (we) couldn't live with. Then chose the alternative that was most attractive to us and had a worst case we could live with.
1.5.6 Choice is described after individual took initiative to create new alternatives.

Example: Continue working as a professional physicist or return to fulltime graduate study in physics towards a PhD. My philosophy toward decisions is that the alternatives are limited only by myself and that the alternatives I already have are shaped to a large degree by my ingenuity. I like my particular alternatives - continue doing work in laser and optical physics which I enjoyed very much and was being paid $31,000 a year to have fun while taking graduate classes part time, or return to my beloved academia and make $7,000 per year while both accelerating my education and providing me with longer vacations to go mountain climbing and sky diving. I was not satisfied with the alternatives so I modified them so that I could sort of have both. I obtained a several year leave of absense from my current position.

1.5.7 Choice is described as "better" one with implied weighing of alternatives.

Example: To take a teaching position or work in industry. I chose the teaching position over the industry job. I like the benefits of the teaching job over those of industry. I weighed the advantages and disadvantages of both positions. Considerations were teaching experience, reduced work load (I had time to finish ongoing research), location - the teaching job was on the coast in a small town as compared to Houston, Texas, and money - the teaching job paid much less but the location and reduced work load compensated for this.
Domain two focuses on the role of the learner. As reasoning structures change across Perry positions the role of the learner is perceived differently.

**DUALISM**

In Positions One and Two truth is known and authorities serve as the holders of knowledge. The role of the learner is to obtain this knowledge. As a result the learner sees himself in a rather passive role of recipient. The instructor or textbook gives the knowledge and the learner must remember and be able to repeat the right answers. The dualist will be rather confused or frustrated with abstract learning experiences where the right answers are not clear. Advice on how to learn will be in the form of paying close attention and getting accurate notes on what authorities say so that the right answers are not missed.

**Position One**

Knowledge or things learned are those things which the external authority dictates as the truth. The Position One person cannot imagine there being more than one answer to a question. There is no question that authorities are right. The Position One person describes the best learning situation to be one in which the truth is obtained and one that they like.

**Reasoning Structures**

2.1.1 Facts and concepts are both absolute/right. The preferred mode of learning is the one the student likes.

**Example:** I like to learn more about facts. Just the specific details. It seems in ideas and concepts you have so many to remember. Your learning of the actual facts of a subject, the most important details is an advantage. A disadvantage is that maybe you are missing the small details of a subject.

**Example:** I learn more with facts because I like history and it deals with a lot of dates. Historical facts stay with you forever, but they are boring and confusing.

**Position Two**

Learning is still guided by the search for the right answer and students prefer the most direct route to getting the answers. The role remains one of acquiring the truth. However, sometimes authorities want students to work to find the answer themselves. In this case the role of the learner is to work hard at doing what the teacher wants. This is seen as an exercise that is good for the mind and is not a genuine search for the truth since authorities already know the truth. It is at
this position also that the student recognizes that some authorities do not seem to know the answers. They are simply labeled as "dumb" authorities.

**Reasoning Structures**

2.2.1 Preferred type of learning is the method that is clearly right and is the easiest way to obtain the answers.

**Example:** I learn best in classes which focus on factual information because I can remember things better as they are real. Concepts and ideas are hard for me to remember. The advantages are that I can look back to the information if I need to. There are no disadvantages.

**Example:** Factual information. The information is easy to understand and relate to. It is easier to study because it isn't abstract. It is uniform and precise, the answers are right or wrong, there is nothing in-between. A disadvantage is that the answers are right or wrong, you don't get any partial credit.

2.2.2 Preferred type of learning is one that authorities suggest is good for you.

**Example:** Classes that deal with ideas and concepts. Because factual information bores me and when learning ideas and concepts you have to think. Learning ideas and concepts help stretch your mind. Factual information can be retained by a computer but ideas and concepts can only be understood by humans. Factual information is a necessity to everyday life and sometimes we need to understand it before we can understand ideas and concepts.

**TRANSITION**

Learning becomes increasingly confusing in **Positions Three and Four**. This confusion is due to the ever increasing body of knowledge in which uncertainty exists. The realization that all truth is not known is tempered initially by the belief that it will be known in the future. As more areas are revealed as uncertain, uncertainty is accepted as the "way it is". The role of the learner shifts from one of a quantitative interest to an interest in the process. The role of the learner initially is the same as in earlier positions for concrete areas of knowledge. For the uncertain areas the learner focuses on the processes for finding the answer. When uncertainty becomes common, the focus remains on the process but not necessarily to find the right answer. The process becomes one of thinking the way teachers want students to think since the right answers no longer exist.

**Position Three**

In Position Three the notion that some of the truth is unknown at present even to authorities complicates finding the right answers.
Abstract learning contexts like philosophy are likely to be viewed as areas where the truth is yet to be discovered. The learners role remains one of receiving the answers in concrete contexts, and shifts to one of learning along with the authority how to find the answer in uncertain contexts. Advice about the latter is in the form of going through the right steps to insure that the truth will be discovered in the future. The student generally draws a distinction between learning facts and concepts. Facts are learned by memorizing whereas concepts require understanding.

Reasoning Structures

2.3.1 Preferred type of learning is one that is easiest to learn or remember.

Example: I can think about the material and relate it with other ideas so that I can remember it better. It is often more interesting than straight facts. If it is factual information I am just being told. If it's conceptual information I am being asked to think. I can design my own thoughts and theories from discussed material. Also if it is factual it can be proven correct or incorrect. If it is conceptual this is not necessarily the case. The disadvantage is that I don't learn concrete facts well and they need to be learned.

Example: Ideas and concepts are easier for me because I don't get so bored with ideas as I do with straight facts. It is easier to learn. No memorization. You don't forget it as easily. At times it can be hard to understand, but once you understand it you don't forget it.

2.3.2 Preferred type of learning is one that results in practical information that is more useful at present and in the future.

Example: Ideas and concepts can be utilized and practiced. Facts cannot be used and thus I forget them shortly after committing them to memory for an exam. Ideas are more practical and interesting too. They can be disputed, make for good conversation.

2.3.3 Preferred type of learning is one that provides a basis for exploring unknown areas in the attempt to discover the truth.

Example: Because of the level of education I have attained, I am aware of factual information in my field and like opportunities to apply concepts and ideas to past learning. The fact that learning concepts and ideas involves students in the learning process is an advantage. Learning can be active discovery, not rote memorization. However sometimes the major point is missed or faulty assumptions can be made due to lack of factual information. Improper conclusions may be made.

2.3.4 Preferred type of learning is one that provides information
and/or a process to be used in solving other problems or answering other questions.

**Example:** Ideas and concepts. You learn procedures for analysis and problem solving. You can apply the solution procedures for any general problem. A disadvantage is that it can often be just memorization and not understanding what you're doing or why.

**Position Four**

Uncertainty arrives full force in Position Four, altering the role of the learner significantly. The world view incorporates uncertainty as the way the world is, a sort of "there are no right answers" notion. Learning is no longer the search for answers but rather a process of thinking. The focus is on how to learn the method of thinking the teacher wants. Oppositional four's are likely to advise playing the game and assert independence by argument since teachers are no longer a source of right answers. Adhering four's are likely to advise trying to learn the method and trying to think independently. Various relationships between facts and concepts are described, all aimed at expanding the possibilities and applying knowledge to a variety of situations.

**Reasoning Structures**

2.4.1 Preferred type of learning is one that focuses on a way to think.

**Example:** It doesn't matter as long as it is clearly expressed and organized so it can follow the thinking. I like ideas and concepts though. I need both. I need the change. It seems easier to go by myself from ideas and concepts to fact than vice versa. If you grasp well the ideas it is like an open door to a greater world of possibilities. You can find the possibilities yourself and explore them. You should not forget to come down to the facts from time to time. There is a danger to live in a world of ideas and to be away from reality.

2.4.2 Preferred type of learning is one that focuses on learning to think for oneself.

**Example:** I learn best in classes which focus on ideas and concepts. Because I am an English major and I need to learn to think for myself. I need to be able to read something and form my own ideas from the work. I've never been a stickler for factual information. I don't feel it is as important as thinking for yourself and forming your own ideas.

2.4.3 Preferred type of learning is one that allows for sharing of information among peers in order to expand one's knowledge of possible perspectives.

**Example:** I'd have to say I learn best from a class that combines
both aspects - applying ideas to facts or vice versa. If I had to choose one, however, it would be ideas and concepts. I find it more challenging to be allowed to voice my own ideas plus learn the ideas of others, teachers as well as students. This makes the class less boring as well. Ideas and concepts are subjective and no one can be wrong about their own ideas. Others ideas are interesting. A class that centers around facts requires a lot of memorizing etc. and does not allow one to form their own opinions or apply what they learn to the present or to themselves.

2.4.4 Preferred type of learning is one that expands the possibilities for how to use information. This is discussed in a broadening sense rather than using information to narrow one's perspective.

Example: I think it depends on what course you are taking and what you are trying to learn. I enjoy classes more when the focus is on ideas and concepts because I like the intellectual stimulation that ideas and concepts can have. The advantage is that you may be able to apply the ideas in more than one area or discipline.

Example: Definitely classes which focus on ideas and concepts. Very often factual information leaves me empty. One could spend a lifetime collecting facts and never understand anything about life. Ideas and concepts link life together while facts divide it into pieces. I can, as a person, learn and grow and understand myself and the world through ideas. One does need a certain number of facts to stay in touch with reality and to complement the ideas. Too much emphasis on ideas could lead to having one's head in the clouds.

2.4.5 Preferred type of learning is whatever is called for on tests, even if that information is of no use later. Learning is defined as a game to acquire grades although it is implied that this is not really learning.

Example: Factual information needs only to be memorized and I'm good at remembering facts. Tests are usually multiple choice, true-false, yes-no etc. and all of these focus on factual information rather than ideas. You remember the facts for the tests but they are rarely of any use after than.

RELATIVISM

The relativistic view of truth is regarded as existing in a context. The acquisition of knowledge is based on experience, expertise, and competence in exploring a given context. Authority is based on this same process and authorities who are perceived as experts can help in clarifying ideas for the student. However, the learner's role is to engage in a genuine search for evidence to support ideas. Thus the learner focuses on evaluation and refinement of thoughts and comparison of ideas for the purpose of choosing the better one(s). There is a quality requirement for the exchanging of information and application is
aimed at better uses in particular contexts. Advice is likely to focus on having a genuine interest in the course topic and critical examination of materials and opinions.

Position Five

The learner's role is to actively engage in genuine inquiry, taking responsibility for one's own learning. The teacher is seen as a source of guidance who can assist with clarification of ideas but who may not be an authority on the topic under exploration. The focus of exploration is to narrow to better choices through evaluation of information.

Reasoning Structures

2.5.1 Preferred type of learning is one that assists in evaluating one's current perspective as new information is obtained.

Example: Difficult question, may depend on circumstance, but overall factual information seems best. I seem to be more comfortable in taking factual information and forming my own ideas and concepts from that information than vice versa. You know your basis for decision is based on fact and the ideas and concepts can be modified as the need arises by new conflicting facts.

Example: Factual information allows evaluation of most ideas, deductively or inductively, and consequently leads to formation and refinement of useful concepts. Whereas more subjective educational efforts tend (I think) to be more often exercise in recognition use and interpretation of imagery.

2.5.2 Preferred type of learning involves exchange of ideas to further one's understanding through comparison of ideas.

Example: Ideas and concepts. I don't like memorizing facts that I could look up in books if I happen to need them at a future date. Nothing is absolute, everyone can look at the available data and draw their own conclusions from it. Different people's interpretations may vary, and comparison of them may help two or more to envision an even better interpretation.

2.5.3 Preferred type of learning involves conceptualization and thinking through problems in order to solve them. Better or worse solutions exist.

Example: I am most interested in how systems function and how systems interrelate. I think this ties into the way my brain functions best. I am particularly adept at seeing patterns and relationships and working intuitively with data. I have always found it more useful to deduce interrelationships and systems in practical, applied research, teaching and actual practice. If you have to solve problems, this is a better mindset.
2.5.4 Preferred type of learning involves integration of knowledge in order to apply in a context.

Example: I prefer classes which focus on ideas and concepts since I dislike memorizing facts without general principles, but I do well in both types of classes. Once I understand a basic principle and how it relates to other principles, I can use it logically and retain it seemingly forever. Some instructors are not creative enough to write tests that measure ability to apply rather than parrot knowledge.
Domain Three focuses on the role of the instructor in the learning process. As the learner's view of truth changes his/her expectations of the instructor change accordingly.

DUALISM

In Positions One and Two the truth is assumed to be known, and known to authorities. Instructors are automatically regarded as authorities whose role it is to pass on knowledge and truth to the learner. Since the dualistic perspective is an absolute one, concrete structured teaching methods are preferred. Dualistic learners prefer presentations by the instructor which clearly spell out the right answers. Clear and specific assignments are preferred. They tend to see class discussion as less useful, since their peers are not authorities. The relationship expected is one of approval when the learner demonstrates acquisition of the right answers.

Position One

The individual at this position does not recognize uncertainty or alternate possibilities. Only one right answer exists for each question and that answer is whatever authority dictates. What authority says is innocently accepted without question. The role of the instructor is to give the right answers in as clear a method as possible. The learner takes in everything from the instructor so that it can be accumulated and therefore learned.

Reasoning Structures

3.1.1 The instructor gives the information the student needs to learn and explains it.

Example: The most beneficial teaching method to me is when the teacher lectures and the class takes notes. I can understand it better when I write it down and the teacher explains it to us. It is not beneficial when the teachers go fast and you don't get everything down. The most important thing to learn is how to pay attention and to catch the important things.

Position Two

Authorities still know the right answers. Although it is preferable that they simply present them, students in Position Two accept that it is good for them to look for the answers themselves. If the instructor wants this to occur he/she should be specific as to how to proceed and provide the right answers eventually. The role if the instructor is to make sure students understand. They may use a variety of methods as long as it doesn't get out of hand.
Reasoning Structures

3.2.1 Preferred teaching methods involve providing details and clear facts. The instructor goes over the information, uses visual aids etc. to help students understand.

Example: The best method has to be the review method which is done every couple of days. This method was excellent because the subject was always kept fresh in your mind and you just didn't learn it for the day, once it's reviewed enough it's hard to forget.

3.2.2 Preferred teaching methods involve clarifying information and answering questions.

Example: Lecture method with allowance for questions. The instructor is able to explain the subject matter. Normally, the lecture method can be an effective method for information dissemination. The instructor should be open to questions during class and available for conferences with the student.

3.2.3 Preferred teaching methods involve making learning entertaining, since it is easier to listen and thus learn better.

Example: I think an instructor that doesn't make not one person bored all the time is a good instructor. He has to be on the level of the students, he has to know what is interesting to the students and what is boring. If it is boring, he has to make that subject interesting. If a student is interested in the subject he will listen and learn more. I learned if you have a good instructor you always keep your mind and thought on him. There is also less temptation to skip class.

3.2.4 Preferred teaching method is one that is "good for you", that lets the students do something and lets them know what they did wrong.

Example: I think the most beneficial effect of different teaching would be participation. Actually letting the student do a certain project on his own. Because the student could do something on his own and learn from his mistakes. You wouldn't know if you made a mistake unless the teacher told you, so a one to one relationship would be good so you know when you were wrong.

TRANSITION

The expanding role of uncertainty alters the role of the instructor as well as the student expectations of instructor-student relationships. Authorities are no longer regarded as having all the answers. Initially authorities role is to help the student learn the way to find the answers that are unknown at present. Developing a relationship with the instructor is perceived as helpful to learning the process. This is even more important when uncertainty prevails and learning is simply
learning a way to think. At this point authorities are relieved of responsibility of having any answers since there are no right answers.

**Position Three**

In Position Three instructors maintain a great deal of their authority. However, the learner has discovered that some truth is not known at present. As the learner's focus turns to finding a process to use to discover the unknowns, the instructor's role alters from giving the answers to helping the students learn the process of finding them. The teacher's role becomes one of guiding the learner to understanding information. The learner is more open to abstract teaching methods than before but is still likely to see student participation as having minimal significance to learning. Preferred teaching methods are those that assist in learning the process for thinking and searching. The instructor-student relationship is beginning to expand since the student sees getting to know the instructor as part of this learning process. A positive attitude on the part of the instructor towards students is essential to helping them learn better.

**Reasoning Structures**

3.3.1 Preferred teaching methods focus on rapport between the student and instructor.

*Example:* A relaxed atmosphere with some restrictions such as required attendance. It seemed as if the teaching was concerned with the students both as human beings with ideas and students that should learn something. Open conversation promotes learning and encourages more students to become involved. I prefer a friendly relationship where I would feel free to consult the instructor and not feel extremely stupid about doing so.

3.3.2 Preferred teaching methods include letting students become involved in class so they can contribute.

*Example:* Those who had the students experience the ideas that they had or themselves through role-playing or other actual hands on experiences. The students were able to see the effects of their ideas on the other people to see if their method works. Disadvantage would be that unless the students used the exact sample of people they were influencing the sample may show different effects. The best way to find out if your ideas work is to try them.

*Example:* Lecture combined with laboratory work covering similar material. It provided concrete application of the material currently being studied. Often abstract concepts difficult to grasp became understandable after experimenting with one's own hands, with plenty of time to investigate. The instructor's method involved patience, carefulness and precision. We learned what can result
from careful application of a scientific method.

3.3.3 Preferred methods focus on students effort as opposed to memory.

Example: One which treated students as intelligent beings with a craving to learn. The material was not simple recipes which were memorized but rather demanded periods of concentrated effort on behalf of the student in order to be appreciated. The method results in comprehension of the material presented.

Example: Combining the methods of two instructors I have had - allow the student to drop a test or two thereby alleviating pressure of every test, and giving all the time a student wants for taking a test. Pressure was taken off of testing and placed more on a students effort. This follows from my belief that any student can do homework if they try, and that those who do the homework will then do well on tests.

3.3.4 Preferred methods constitute an organized process for learning.

Example: The instructors who had the best effects on students were always well organized. They usually had good rapport with the students and were able to motivate them to learn or instill some excitement in them concerning the subject matter. It is much easier to learn when a teacher is organized and does not jump from one subject to another. A logical, concise and organized presentation is essential if learning is to occur.

Position Four

In Position Four the learner no longer automatically regards the instructor as an expert since truth is not known. The learner's focus shifts to a method of thinking, which the instructor models. Greater interaction is expected and preferred both with the instructor and peers. Exploring ideas and sharing opinions is regarded as important and the teacher should respect students ideas. The instructor should strive to make the classroom a forum for exchange of ideas. The oppositional four does not accept the importance of learning a method of thinking but rather views learning as a game to be played. The instructor is perceived as having no more knowledge than the students and oppositional students rebel against the instructor and what they perceive as the game of learning.

Reasoning Structures

3.4.1 Preferred methods foster opportunity for students to share their opinions.

Examples: The ones who let me think out my own solution. Let me gather and make up my own opinions. It taught me how to think about other people better, and hear their opinions. When someone gives their opinion I hear it with an open mind. This gives everyone a
chance to express themselves.

3.4.2 Preferred methods include alternatives to instructor's perspectives and the instructor should respect other opinions.

Example: Method which used a firm basis to which to refer to study. A class with a teacher secure of himself and knowledgeable. Lecture followed by questions and open discussions. A lecture with questions by teacher or students: it becomes a trampoline from which you get interested and jump to other classes or books to find answers. If the teacher is not secure, there are so many questions that it becomes a market of questions with no direction. I learned that you should consider a subject from many points of view, not only one. You should be interested by what you do and be open to others opinions. Students should respect the teacher's opinions and he should respect mine.

3.4.3 Preferred methods allow for students to define their own learning goals.

Example: I enjoyed instructors who were well prepared for lecture and who presented material in an understandable fashion. I also enjoy contracting for grades, but this requires high approachability on the instructor's part. Contracting for grades made the learning experience very real. Goals were set according to individual needs and objectives.

3.4.4 Preferred methods teach students how to learn and think.

Example: I think the teachers who teach more than just content and expect students to be responsible and creative are the best. Teachers who teach how to learn, how to approach school and schoolwork more effectively. Content is quickly forgotten but attitude and approach are not.

RELATIVISM

With the relativistic viewpoint, the possibility of truth existing returns. Truth is viewed as existing in a context with holders of the truth identified as those who have experience, expertise, and competence in searching for the truth in a given context. Learning is viewed as a genuine search for the truth, looking for evidence to support opinions. The role of the instructor is to engage in the search mutually with students. Abstract teaching methods are preferred, as well as mutual interaction with instructors.

Position Five

Student and instructor become partners in a mutual search for truth. Instructors are viewed as competent in given contexts in which they can provide evidence for thinking. Instructors are expected to be encouraging and accepting of critical questions and alternate viewpoints.
if supported by evidence.

Reasoning Structures

3.5.1 Preferred methods focus on problem solving or applying knowledge to a context.

Example: Those professors that were structured and expected a lot from their students. Professors that allowed thinking and formulation of concepts in an application sense. It allows me to absorb information, process it, give it some thought and be able to apply it (problem solving).

3.5.2 Preferred methods involve student and instructor discussion of perspectives in an evaluative manner.

Example: I have always been inspired by teachers who generate a learning atmosphere, i.e. relate their subjects to real life, make it possible for students to ask questions, allow discussion to generate around question-answer debate and provide substantial and decisive feedback to the student. Relating subject to real life equals learning through association. Asking questions equals formulating associations and ideas that can be communicated. Discussion puts individual associations in a social context. Substantial feedback equals substantiation of ideas held and development of new ideas.

3.5.3 Preferred methods allow for students teaching and critiquing each other.

Example: I like classes where the teacher and the students are both able to talk about a subject and exchange ideas. I have enjoyed the classes where the instructor has had some expertise in a particular field and has brought the expertise to the classroom as a means of facilitating discussion among the class members. It is a good opportunity for students to add their own relevant expertise to the thoughts and opinions of the teacher. This makes for a well rounded discussion and sharing if ideas among a group of learners. The teacher has an institutional obligation to bring in information relevant to a given topic, but all share in the obligation to learn from one another and share ideas.
Domain four focuses on the role of peers in the learning process. Their role is dependent on the learner's view of knowledge and the learning process. Changes occur across positions as described below.

DUALISM

In Positions One and Two authorities are viewed as holders of the truth. Since learning is a matter of acquiring the truth, student participation is not viewed as useful. The dualist wants to hear the truth from the instructor, and not be confused by discussion. Although students asking for clarification is seen as positive, discussion is seen as confusing, distracting and often a waste of time. The dualist sees learning directly from the instructor as more beneficial. Relationships among members of the class are not a focus for the dualist.

Position One

Involvement consists of listening to authorities and getting the right answers. Questions should be direct attempts to get the right answers. Additional discussion is seen as a waste of time.

Reasoning Structures

4.1.1 Peers do not have a role in learning.

Example: I prefer classes where students do not talk. I want as much of the instructor's knowledge as possible. Many students ask questions because they come to class unprepared and so take time away from obtaining information from the instructor. Perhaps a student might have considered a very important aspect which has not been clarified.

Position Two

Peers can have a role in the learning process at Position Two, although it is limited to asking questions for clarification or discussion to find answers for oneself. The instructor is still responsible for maintaining control of the class so that student participation doesn't get confusing.

Reasoning Structures

4.2.1 Role of peers is to ask questions to clarify information or bring up questions other do not think of.

Example: I prefer classes in which the students do a lot of talking. Because someone might bring up something I wasn't sure about or didn't think of asking myself.
4.2.2 Role of peers is to ask questions and give answers to make class less boring.

Example: I like classes when students get involved. Because it is more interesting and doesn't get boring. You can say what you feel about the topic of discussion. The disadvantage is that class can get off the subject.

4.2.3 Role of peers should be minimal so as not to confuse what the teacher is trying to teach.

Example: I prefer classes where students don't talk much or talk a lot if they know what they are talking about. I prefer the not talking because I can concentrate more without students confusing me. I like the students participating if they aren't going to turn things around and confuse the other students.

4.2.4 Role of peers is to study together because some students pick up knowledge that others do not and can explain it in a simpler way than the instructor.

Example: I prefer small study groups. Sometimes it is easier to learn from a student who understands the material, and is able to explain it to others.

TRANSITION

Student involvement becomes increasingly acceptable as it relates to learning a process and relationships. The focus on process that accompanies the rise of uncertainty allows for class discussion to be a part of the process. During the transition the process changes from one that insures the truth will be found to one that focuses on learning how to think. Relationships gain increasing importance as authorities are no longer the source of all knowledge.

Position Three

Student involvement is a bit more acceptable in Position Three since even the instructor may not know all the answers. Peer interaction is seen as legitimate as it relates to discovering the process of finding the truth such as in small group discussions. Interaction has a number of advantages as long as the instructor keeps the class under control and has the knowledge that he or she should to share with the class.

Reasoning Structures

4.3.1 Involvement on the part of students helps one to become involved, class is more interesting and therefore learn better.

Example: I prefer classes where students do a lot of talking because it gives the student the chance to be personally involved with the lecture and his classmates. If he talks about the subject
then he is learning that much better. Because the student feels a part of the lesson it becomes more important. The class can go off on a tangent. One student can mislead or confuse the others.

Example: Some talking. A better understanding of material that is being taught is possible when there is a feeling of enjoying learning. Less information may be passed along though.

4.3.2 Peer involvement is a process for learning.

Example: I prefer high participation as long as it is controlled. I feel that student questions help the instructor to fine tune his/her teaching to meet the needs of this particular class and help the more shy students to clarify their understanding. It keeps instructors thinking and students practiced in inquiry and critical thinking.

4.3.3 Involvement is a method of providing exposure to new ideas.

Example: I prefer talking because when other students are involved in the beginning more student will eventually tell their opinion. You get a chance to hear what others are wondering about. If your thinking was wrong and someone else says the same thing you learn from your mistakes. Or if a person says something that you never thought of.

4.3.4 Involvement can be a way of helping students think.

Example: Student do equal amount of listening and talking. Because professors have a lot of valuable knowledge, but students should not be stifled. They should be able to probe deeper and ask questions in order to formulate ideas and concepts.

Position Four

Peers are as legitimate a source of knowledge as the instructor since everything is uncertain. Everyone's opinion is equally valid and the learner prefers high student involvement to share opinions. In addition the Position Four person focuses as much on class relationships as on learning. The oppositional four is likely to be more outspoken regarding their notion that student's opinions are as valid as the instructor's since they are rebelling against the instructor and the game of learning.

Reasoning Structures

4.4.1 Involvement is seen as an opportunity to share viewpoints in the process of gaining a variety of diverse perspectives.

Example: I prefer class participation because student have fresh new ideas and can open up new avenues of insight that don't occur to the instructor. More indepth learning and understanding by
experiencing not only your viewpoint and the viewpoint of the instructor but the viewpoints of everyone else. A disadvantage is that you can get some moron in the class that just aggravates everyone and holds back the learning process.

**Example:** Talking is a part of my learning strategy. Learning is not only a result of the teacher-student relationship but the student-student relationship as well. Students often learn as much from each other as they learn from the teacher. Shared experiences lead to more different possibilities for subject related interpretation. Exercise in idea formulation and communication skills, immediate rewards for effective thinking.

4.4.2 Involvement is seen as positive since teachers and students are equal in terms of their knowledge.

**Example:** Lots of discussion. It is an active learning where everyone has the chance to learn and grow including the teacher. It is a non-elitist form of learning. People come from many points of view and they deserve respect. It respects the individual intelligence and opinions of each person.

**Example:** Some talking is useful. The students feel that their questions and comments are important and realize that learning is a two-way, give and take situation.

4.4.3 Involvement can be destructive if it is not controlled.

**Example:** Controlled talk, controlled by instructor decision of appropriateness. Better understanding of problems, concepts etc. would result if instructors had time to formulate and bring forth their knowledge or expertise on a subject. It does limit class participation which could open new ways of thinking.

**Example:** I prefer talking be limited to relevant, succinct questions. It allows more time for the instructor to present his views on the material. It is always fun to attack such views. It does not easily allow a class to degenerate or get off the track. I also think it is a more efficient use of class time.

4.4.4 Involvement is a way to learn to think independently.

**Example:** Lots of student talk. Advantages are that it is more stimulating, more active learning (how can you learn by osmosis?). Learning what you learn not what the teacher tells you. Independent thinking.

**RELATIVISM**

All knowledge is relative (contextual) so authority rests on expertise acquired through searching for the truth. Since learning is a genuine search for truth and peers are a legitimate source of knowledge, class
involvement can be very positive. The learner prefers sharing ideas in a mutual sense with peers and instructors.

**Position Five**

Classes are a forum for seeking out knowledge and soliciting new perspectives. Participation is an opportunity to critically assess ideas and gather information to further explore ideas. Detachment to explore ideas objectively occurs. If participation does not meet these criteria, however, it may be seen as negative.

**Reasoning Structures**

4.5.1 Involvement is useful because it draws on the expertise of the class as well as that of the instructor.

**Example:** Students have a wide experience base usually. Their ideas offer a new perspective/depth to the content. It recognizes individuals, draws on the expertise of the group, allows for the growth of a class.

4.5.2 Involvement may lead to enhancing the learning process and the content of the course.

**Example:** I like a mixture depending on the course material. When the instructor is an expert in his field I appreciate his participation. Although the students participation is also welcome. It is enhanced by the instructor's guidance and comments. Student participation is needed to help discuss and digest material presented. This is extremely important when discussing concepts and ideas.
DOMAIN FIVE

Domain five focuses on **how learning is evaluated**. The view of knowledge and roles of instructor and learner affect the learner's notion of how evaluation should be accomplished.

**DUALISM**

In **Positions One and Two** the instructor is seen as the authority who knows what is right and wrong. The instructor is also awarded the right to evaluate the learner's progress since the instructor is the best judge of truth. The learner in these positions believes that good grades should automatically follow demonstration of having acquired the right answers.

**Position One**

The student accepts without question the grade the instructor assigns. Grading is a simple process of the instructor counting the right answers that the student provides.

**Reasoning Structures**

5.1.1 Evaluation should be based on whether students know the right answers.

*Example:* Evaluation should be on whether or not you know the material because either you know it or you don't. The instructors should be the people who evaluate your work because they are the ones giving you the work.

5.1.2 Instructors evaluate because they know what to expect and whether students know the answers.

*Example:* No one but the instructor should evaluate and he should weigh it against the competition and what his experience tells him he should expect. The instructor is the most logical person to render the decision.

**Position Two**

In Position Two the student still sees the teacher as an authority capable of evaluating students. The instructors methods of measuring knowledge are not questioned. The student's role is to provide right answers on tests and do their work. Knowledge is quantifiable.

5.2.1 Evaluation should be based on one's progress, improvement or how much one learned in a class and done by the instructor.

*Example:* The basis for evaluating the work done in college courses should be how much learning took place. For some students this will mean a lot of work and for others it will take less. The teacher
should be the one evaluating how much learning took place and his/her decision should not be based on personal biases or feelings about the student.

5.2.2 Evaluation should be based on the degree to which one knows the material. The instructor is the one to determine this.

Example: The teacher should be the only one involved in the evaluation. Test scores usually are an accurate indication of how you are doing.

5.2.3 Evaluation should be based on completion of assignments and how hard one tries (i.e. did the work). Instructor is the one who determines this.

Example: The basis for grades should be the completion of assignments, papers and exams by the instructor only! The instructor is the person responsible for the course and the student's grades in that course.

Example: Grades should be based on past work done by former students. Also how much effort is put into the class i.e. attendance, punctuality, quality of assignments and test grades.

5.2.4 Evaluation should be based on whether one keeps up with other students.

Example: Effort, I think, has the most bearing on whether or not one learns as much as he can. If a student is keeping up with the other students, he is taking in the information presented to him. A student needs to know if he is lacking in certain skills or areas so he can try to improve them.

5.2.5 Evaluation should take into account whether the course is teaching what it is supposed to.

Example: Evaluation should be based on how well the students are doing, because if the teacher is a poor one, the students can do proportionally well. The teachers (TA and professor) should be involved in the evaluation. The TA is the one who works more closely with the student but the professor should know approximately how much the course should be teaching. The evaluations should be combined.

5.2.6 Student involvement in evaluation can be useful if it is a chance for the teacher to correct the student and know what material needs more attention.

Example: I think the student and teacher should be involved in evaluation and the basis we have now should be the same except that students should have more influence. If you have some say in your grade, you can better understand what you are doing wrong and are
able to correct the problem.

Example: Evaluation should be based on a series of small quizzes involving the material covered that day. All student should be involved. Because it lets the instructor know who has grasped the material, if any, and help to focus on material that raises difficulty.

TRANSITION

Grading becomes more confusing as the degree to which authorities know the truth dwindles. Since the focus of learning has shifted to processes so should evaluation of student's learning. Initially, the student is uncertain how the instructor can evaluate work other than looking at quantity, but sees authority as the best way to judge progress. Eventually the student sees evaluation as impossible in a genuine sense.

Position Three

In Position Three evaluation is no longer absolute and quantifiable. Issues arise with regard to objectivity, individual learning styles, practicality of material and confusion over alternate methods. It is still the responsibility of the instructor to take these into account when grading. The instructor is responsible for finding an objective way to accomplish this.

Reasoning Structures

5.3.1 Evaluation should be objective and fair.

Example: The teacher should be involved in the evaluation and should check to see if students' grades are the same on the tests as the homework. If someone does super on homework and quizzes and gets a D on a midterm they should not get punished for one bad grade. Everything should have equal value. Some students choke on tests.

5.3.2 Evaluation should account for individual differences in learning and appreciation of the student as an individual.

Example: I think effort should be involved in grading. If the teacher knows someone worked hard and gets nowhere they deserve a little credit. At least more than someone who shows no effort whatsoever. It takes in a personal characteristic of the student as well as his overall knowledge of the subject.

5.3.3 Evaluation should focus on practicality of material for future use.

Example: I think it should stress more labs, projects or other practical things that will be used in the future. There should be
less stress on theory or other subjects which will not be used after graduation.

5.3.4 Evaluation is seen as inadequate but new alternatives are not perceived.

Example: I'm not really sure there is an ideal way to evaluate students. To me just giving tests or just papers or just presentations is not fair. Students may choke on tests or do poorly because of nerves. It seems to me it is actually impossible to accurately judge what a student has learned in a class and to me this is what should be graded. To me grades are not really fair, but if we must have them, they should be based on diverse criteria and perhaps students should have a little input somehow.

Position Four

Evaluation shifts away from knowing the material since knowledge is now uncertain. With this basis for grades gone and the increasing value of students' opinions, numerous possibilities arise for evaluation. Students focus on their involvement in the evaluation process. At Position Four the instructor is no longer responsible for determining the most objective method to evaluate. Grades are generally seen as less useful than a conversation in which the instructor expresses an opinion about the student's learning while the student comments on the value of the course. Seemingly as long as both have their input, any number of methods can be used for evaluation. Some fully acknowledge the inadequacy of evaluation methods but perceive no criteria upon which to base evaluation.

Reasoning Structures

5.4.1 Evaluation is a mutual process between students and teacher in which they agree on an evaluation system.

Example: Ideally a type of written evaluation would be used. At the beginning of a course, the teacher and student or teacher and whole class would form a kind of contract outlining what they wanted to see accomplished. The teacher could explain what is advisable to learn and why but the decision would be mutual. At the end of the course the teacher could evaluate how much was accomplished relative to the goals. But even if the goals weren't met, the short written evaluation should stress what was accomplished.

5.4.2 Evaluation is an exchange of feedback with the instructor giving feedback on student learning and the student giving feedback on the value of the course.

Example: I feel the instructor should be involved in the evaluation as well as the student himself. I like the idea of feedback from the instructor but I also feel the student should be allowed to express his opinion of the instructor's advice.
5.4.3 Evaluation is based on becoming an independent thinker, thinking for oneself, in such a way as to open up new possibilities.

Example: Ideally, a student's ability to use the fundamental information he/she obtained to solve problems or develop ideas should be the basis. An oral discussion by a competent person (instructor?) would be advantageous.

5.4.4 Evaluation is problematic but no other alternative is perceived.

Example: This I have debated for many years. I'm totally against standardized testing but the two alternatives are this: a strict personal evaluation of the individual by the instructor, or a testing of students as a whole basically on a curve system. I can't personally make a decision.

RELATIVISM

The relativistic person acknowledges uncertainty, but views knowledge as existing in a context. Anyone who has competently engaged in the search for truth in a given context can serve as an authority and thus judge others learning. This could include instructors, peers or self. Grades might also be viewed as unnecessary since the search for evidence and ability to support one's opinions is sufficient for learning.

Position Five

Evaluation is based on the student's competency as related to the particular topic at hand. Competency includes understanding, ability to apply knowledge in a context and ability to support one's thoughts. Position Five persons note that a competent person should be involved in the evaluation.

Reasoning Structures

5.5.1 Evaluation should be conducted in such a way as to accurately measure competency in a subject matter.

Example: I would like to see the grading system thrown out. I would prefer a pass-fail system with the professor and student being involved. Exams do not really reflect a person's understanding of a subject thus the grade is not a true basis of evaluating competency in a subject matter. I would like to see perhaps papers or oral examinations be given.

5.5.2 Evaluation is a form of mutual exchange between instructor and student to maximize growth.

Example: Evaluation should be a cooperative exchange between teacher and students. Feedback is important, but students often know where they are if asked. If the goal is for growth then
student ideas must be accepted.

5.5.3 Written exams and grades can be acceptable method of evaluating competency if they are the best alternative in the context.

Example: In practice there is no practical alternative to grades, but they should be supplemented by leaving exams and theses or practicum to better demonstrate the student's true master of a subject of skill.

Example: Even though written exams aren't the best means of student evaluation, they sometimes are the only alternative. Ideally, an oral exam or two could show what, if anything, the student has been able to derive from the course. At a college the size of OSU I think that written exams are the only possible alternative. A single instructor can't possibly give an oral exam individually to each of the students in the class.

5.5.4 Evaluation is a combination of student and instructor working toward a goal together and measuring their progress.

Example: The basis would be a competency model. Before the course begins an outline should be available of what the student should be able to do as a result of the course. This provides a goal toward which instructors and students can work together and also a standard against which to measure performance.
DOMAIN SIX

Domain six is intended to explore the view of knowledge, truth or reality. The degree to which one believes truth exists and how one arrives at the truth are the focus of the changes as described below.

DUALISM

The dualistic person believes that truth exists. Only one perspective on any subject is the true one. Therefore conflicting notions are dealt with as though one of them is not there or is not true. The process of dealing with conflicting versions of the truth is a simple one of deciding which one is right.

Position One

The categorization of conflicting notions into right and wrong is a simple process. Differing explanations just seem different. The one assigned to the right category is likely to be the one most easily understood.

Reasoning Structures

6.1.1 Knowledge is certain but one might view explanations of knowledge differently due to lack of understanding.

Example: No, I don't think one can be more correct than the other because they should know what they are talking about. I would believe the one that I understand the most. I'm not really sure because we usually believe what the teacher tells us.

Position Two

The categorization of conflicting notions into right and wrong remains a simple process. The only difference is that the explanation for there being two perspectives is that one of the authorities is dumb and doesn't know the right answer.

Reasoning Structures

6.2.1 Knowledge is certain and discrepancies are due to one explanation being wrong.

Example: It depends on the degree to which the instructor has researched the topic. Taking in all the facts and finding out which theory could possibly be disproven. Yes, they would have to check themselves and either form their own opinion, or disprove another theory.

6.2.2 Knowledge is certain and different explanations are due to differing degrees of detail.
Example: It's possible. One might go into more detail than another one might. Go directly to the "horse's mouth". Ask what the explanations really mean and try to determine from there. I'm just not sure whether one can ever be sure. One can't be sure if they didn't go to the source and find out what the problem was.

Example: I don't think that one instructor can be too far off of another instructor's ideas, concepts and facts. They all know the material to their best but just explain it in a different way. One would have to ask the teacher if confused or better yet ask another teacher in the same field what he/she thought was right. More than likely he/she will agree with both of them. I really don't think one can be sure unless one explanation is so far fetched that even I know it is incorrect.

TRANSITION

In Positions Three and Four the notion that truth definitely exists in all areas begins to change. The person initially discovers that although some truth is known, other areas seem not to be known yet. As more and more areas fall into the latter category, the thinking shifts to the notion that everything must be uncertain. The Position Four person concludes that truth is not really known, everyone just decides for themselves.

Position Three

In Position Three it is possible for the truth to be disputed since some things are yet to be known for sure. It becomes difficult however to define how the disagreement can be resolved. Should one expert appear to have a better process for finding the truth, he/she may be presumed to be right, at least until the truth is really discovered. The Position Three person believes that the truth will be known for sure in the future.

Reasoning Structures

6.3.1 Knowledge falls into categories; that which is known and that which is not. Different explanations for the first category are due to difference in details, viewpoints and authorities. For the latter category differences are due to the knowledge being unknown.

Example: Yes they could. I would take the one that seemed to be more logical and seemed appropriate to the situation. Also I would check if there were supporting facts or not for the explanation. One can be sure by the facts supporting the explanation. Who's going to believe something that can't be supported? One can't be sure if neither explanation has support. The explanations are probably far out opinions, or a concept not based on any real facts.

6.3.2 Assuming the same two categories of knowledge, if one is discussing abstract areas, one cannot be sure what to believe but
can choose on the basis of logic, acceptance, what one believes and the most recent information.

**Example:** Yes. Historical events are facts. You cannot change their interpretation. Scientific phenomena may be explained differently. I would accept the one which sounds more logical to me and which was more widely accepted. One can't be sure unless they personally investigate this, they will not know whether it was explained by the person's opinions or by strictly factual occurrences.

**Position Four**

In Position Four it is even more possible for the truth to be disputed. Since no one knows the truth there can be an infinite number of explanations for any notion. Discrepancies cannot be decided since no criteria exist to decide. Everyone has a right to his/her own opinion and no one has the right to call another person's opinion wrong.

**Reasoning Structures**

6.4.1 Differences are due to everyone having their own set of beliefs. One just chooses what one believes is important or what is reality.

**Example:** Not necessarily, one instructor may view the event differently from others. I would adapt the explanation in which I personally believed in. One can't be sure because everyone is an individual and thinks differently about certain situations. They have their own set of concepts and beliefs.

6.4.2 Differences are due to interpretation or bias. One doesn't know what to believe so thinks it out on one's own.

**Example:** Sometimes. History is especially subjective, and even in science, research results will, at times, contradict each other. However, different explanations seem to be often affected by ideology, religion, and prejudices in general. One example that comes to mind is Freudian psychology. One instructor said it was worthless and another implied it was ultimate truth. I took several things into account when deciding which to believe. I listened to the instructor, considered his credibility, depended on outside information from reading and talking to people, and looked at the ideas I agreed and disagreed with. The only way to be sure about anything is to get as much information from as many varied sources as possible and do some thinking on one's own. With issues that involve morality and values, many explanations are equally valid simply because they are based on values and not facts. Also since scientific evidence can contradict itself at times, one can become confused and unsure.

6.4.3 Differences of opinion exist and one can choose parts of both or
decide there are not enough criteria to make a decision.

Example: "Can he be?" - yes. "Is he undoubtedly?" - no. There is frequently support for several different theories. Educational backgrounds and biases may cause one scientist to select one theory above another. I doubt I'd find a need to believe in either one of them above the other. I would find a need to be able to speak about both ideas intelligently however. If the "facts" leave only one answer to you then believe it. But hang on to the other ideas and be open to reevaluating your mind set.

6.4.4 Different opinions can be correct or incorrect. One accepts both in an attempt to stay open minded.

Example: No - just a matter of the person relating the information - his/her interpretation. Certainly no one answer is right, the world is not that simple. I would go on what I feel is logical. But I certainly would not discount the other; I would store it in my memory and hopefully recall it at a time when I needed to. One can never be sure. A lot of people have given credibility for example, to Freud's work - now it seems that psychologists are drifting from him. I think there are periods we go through when we give credibility to one idea and then we shift back, always changing. Two explanations can be logical and worthwhile, so I take the information given for that period of time for what it is.

6.4.5 One establishes one's own criteria for choosing unless in a course in which case one uses the instructor's criteria.

Example: Yes. One may be teaching on a more detailed level or the event/phenomenon may not be precisely enough investigated at this time to preclude alternate equally possible explanations. The explanation proposed and maybe the one you believe may depend upon your reason for investigating the topic in the first place. You must base such a decision on your own context, or in courses, upon the expressed goals and emphasis of the course/instructor.

RELATIVISM

In relativism the truth can be known again but on a basis different from that of the earlier positions. Truth is now based on experience and expertise in exploring a particular area of knowledge. Although truth may not be known in an absolute sense, it can be decided on the basis of the evidence at hand.

Position Five

Persons in Position Five will analyze both views in terms of the evidence and exhibit a more complex understanding of the issues involved. Judgements of better and worse arguments exist and decisions are based on evidence rather than authority, norms, or personal opinion.
Reasoning Structures

6.5.1 Knowledge is uncertain but one can determine the most appropriate viewpoint based on how well it fits the context in question.

Example: A theory must fit the facts. A theory is not "correct", it is merely the best explanation of the observed facts which are incomplete. One theory may fit the facts better but the facts are incomplete. We can't have total knowledge, so our predictions are based on incomplete data. New data may be predicted by theory but seldom perfectly accurately. Our most current physical theories tell us you can't determine anything absolutely, only high probabilities.

6.5.2 Knowledge is uncertain but judgement can be made after evaluating the evidence. (Evaluation of evidence may be implied or explicit.)

Example: Being a scientist I would say yes. Usually, but not always; one explanation is more plausible than others. I base my decision on which explanation has the most supporting evidence and which person has drawn the most logical conclusion from the evidence. By relying on information you have obtained and your own experience, you may be able to judge the merits of an explanation. If you have little knowledge and/or evidence in the area, judgements become difficult.

Example: No, as different theoretical perspectives may be given. Hopefully both are knowledgeable of the topic. I choose that which best fits my own philosophy which has been formulated from further readings in the literature and research journals. One can be sure through further investigation, reading and asking questions.
REFERENCE NOTES


21. Tamashiro, R. Seminar on Coding Methodology. The Ohio State University, 1981.
Bibliography


King, P. M. "The Development of Reflective Judgement and Formal Operational Thinking in Adolescents and Young Adults." Doctoral dissertation, University of Minnesota, 1977.


