

LEARNING AT EVERGREEN:  
An Assessment of Cognitive Development

Kirk Thompson

The Evergreen State College

Assessment Study Group

1991



LEARNING AT EVERGREEN:  
An Assessment of Cognitive Development

Kirk Thompson

The Evergreen State College  
Assessment Study Group

1991

---

(c) 1991 by Kirk Thompson

## **CONTENTS**

<b>MODEL AND METHODS</b>	<b>1</b>
The Perry model	1
Methods	5
<b>THE FRESHMAN YEAR</b>	<b>13</b>
The Washington Center study	13
Hypotheses	21
<b>"INDIGENOUS" GRADUATES, 1986-1988</b>	<b>24</b>
Description of the graduates	24
Class profiles	26
Gain at the individual level	44
Gain by demographic and academic groupings	57
<b>CONCLUSIONS AND APPLICATIONS</b>	<b>76</b>
Research agenda	76
Overall conclusions	77
Applications	81
<b>REFERENCES</b>	<b>85</b>
<b>APPENDICES</b>	<b>88</b>

## MODEL AND METHODS

In addition to assessing piecemeal the numerous skills and disciplines which are supposed to add up to a college education, can assessment directly address students' *cognitive development*, which is arguably the focal point of liberal arts and sciences education? This study explores students' cognitive development at The Evergreen State College, using the "Perry scheme" or "Perry model" for assessment (Perry, 1970, 1981). The Perry model and this study both assume that a central function of college education is to help students think in more complex ways.

The study shows that Evergreen students differ from students elsewhere in the complexity of cognitive structures which they employ at matriculation, in their development during the freshman year, and in their level of cognitive development at graduation. It provides a "value added" measure, gain in cognitive development rating from the freshman to the senior year; and it explores the magnitude and distribution of this kind of value added. It finds that most gain occurs in the freshman and senior years, and that freshmen who rank lower in cognitive development are likely to gain more than those who initially rank higher. It indicates that Evergreen tends to serve different demographic categories of students equally, and that all of Evergreen's modes of study contribute about equally to cognitive development. Initially, gain in complexity of thought is fostered by interdisciplinary coordinated studies programs; but in the third year of coordinated studies there is a point of slight diminution of returns, and more gain in cognitive development seems to accrue from advanced work in the individual contract mode (independent study and internships).

Overall, the study describes success: by this measure, the College is accomplishing its mission by providing an alternative mode of collaborative, student-centered learning that works as well as, and apparently even better than, traditional educational systems to foster cognitive development.

### The Perry model

Choice of outcome measures depends on assumptions about what is supposed to

be accomplished in college. The assumptions of this study center on the value of interdisciplinary liberal arts studies. William G. Perry, Jr.'s scheme or model of cognitive development provides "goodness of fit" to liberal arts and sciences education in general, and specifically to Evergreen, which heavily emphasizes full-time, interdisciplinary, team-taught programs. The present study assumes that liberal arts studies and particularly interdisciplinary studies (and the variants called collaborative learning environments, learning communities, and integrated or integrative studies) aim at educational outcomes less in the areas of specific skills and disciplines, and more in the area of cognitive development. In Dewey's phrase, "development is the aim of education." Overall, the aim of education is development of the whole person; but a college or university focuses particularly upon cognitive development, or places it in the lead. Vocational courses and academic skill courses develop capacities for specific tasks; academic disciplines work at understanding particular knowledge-domains; but interdisciplinary or coordinated liberal arts studies work primarily at development of the qualities, styles, and habits of thinking per se. The Perry model describes such cognitive structures and arranges them in order of increasing complexity. Within the model, cognitive perspectives which focus on skills and disciplines have a definite place and value; but the model also accounts for more complex modes of thought, and ultimately is sensitive to the full development of personality.

The model offers further advantages. To begin with, it is student-centered. Perry developed it from longitudinal interviews with students, and it rests on what they said about thinking and learning. It permits fairly easy, valid, and reliable assessment by using a performance or production instrument called the Measure of Intellectual Development (or MID), in conjunction with a standardized rating scheme for assigning student writing samples to positions in the model (Knefelkamp, 1974; Mentkowski, Moeser, & Strait, 1983; Moore, 1982/86, 1986, 1989). Enough data has been collected nationwide to allow rough comparisons. The rating system has proved adaptable to a special circumstance at Evergreen, where extensive portfolio material is available for rating. The model has potential limitations, particularly because of change in the gender, ethnicity, and values of college students since the original subjects were interviewed. It thus remains subject to revision (see Perry, 1981; Belenky, Clinchy, Goldberger, & Tarule, 1986; Thompson, 1988). The present research, in a later phase, will attempt to update the model. But such revisions will probably serve to incorporate contemporary learning styles and styles of expression (along the lines of Belenky et

### *Model and Methods: 3*

al.), without fundamentally revising Perry's concepts of the cognitive structures or "epistemologies" which students employ. The model seems to remain robust in its power to classify cognitive positions or perspectives in terms of their increasing complexity.

The full Perry model appears in Appendix A. It consists of nine "positions," or levels of complexity in cognitive development and in the relation of the knower to the known. MID ratings usually fall within Positions 2 through 5, which are outlined below. The average freshman starts college in transition from Position 2 to Position 3, and the average graduate ends up in Position 3, or in transition between Positions 3 and 4 (Moore, 1982/86).

*Position 2, dualism or received knowing.* The student's main concern is "what to learn." Knowledge is information from Authorities (teachers and texts), and can be classified basically as right or wrong, true or false. The student's job is to make sure she got the teacher's or the text's information straight, and to be able to repeat it.

*Position 3, early multiplicity, or separate, procedural knowing.* In some areas, the right answers are not yet known. The student's main concern is "how to learn," and she begins to see herself as a problem-solver in different disciplinary domains. The methods for problem-solving seem to exist separately and discretely, and the student experiences no concern, as yet, with relationships among methods and domains.

*Position 4, late multiplicity, or subjective and connected knowing.* The student's main concern is "how to think." She encounters diversity of informed opinions, wonders whether or not different points of view can be reconciled, and becomes concerned about forming her own opinion. Authority in the singular becomes authorities in the plural; and authorities are observed to differ because their methods and assumptions, and particularly their gender, class, and culture, afford competing points of view.

4a, Oppositional stance (critical and subjective; closer to the masculine stereotype). Everyone's entitled to his own opinion; you can't call anyone wrong or let them call you wrong; relativism is the only truth. Trust your own gut, and then use critical reasoning defensively to undo other people's assumptions. From this stance, students often generate "independent-like" thought: they have observed what clear conclusions would look like, and they offer likenesses of such assertions without yet knowing how to muster coherent supporting arguments.

4b, Adherent stance (cooperative, caring, connected; closer to the feminine stereotype). I'm trying to express my own opinion the way I see it - the way that's right for me. I grant the same personhood to others and try to appreciate all the differences among opinions, mainly in terms of the different determinants that make people think the way they do.

*Position 5, contextual relativism or constructivism.* The student undertakes "thinking about thinking" (or "metathinking"). While she can no longer believe in abso-



lute or certain knowledge, nevertheless she can inquire by constructing tentative theories and models, some of which are demonstrably more appropriate than others to the context at hand. (Perry, 1970, 1981, with modifications suggested by Belenky et al., 1986.)

Using this model, one can rate interview protocols or written material in terms of which position is "dominant" in them. Ratings indicate both the position in which most of the material belongs, and the direction of any apparent change. The idea of a "dominant" position does not imply that a student functions exclusively in one position at one time. Functioning is complex because from each position one has access to the positions or cognitive structures prior to it, whenever their use is appropriate. For example, a student capable of contextually relativistic thinking (Position 5) retains access to dualistic structures and to multiplicity (Positions 3-4). While she might rankle at the simplistic dichotomies of a "dualistic" true-false test, she can provide true-false answers when called upon to do so. In addition, cognitive functioning is complex because a student can move ahead in one area of inquiry while still lagging behind in another one (as in Piaget's *décalage*). Perry hypothesized that learners tend to recapitulate the whole sequence when entering a new field of inquiry (1981; see also Kegan, 1982).

In this model, the transitions are just as important as the stable positions or stages. This makes the model relevant to teaching and learning. Knefelkamp (1974) formulated the idea of "developmental instruction," that teaching and learning can be thought of as inducing transitions from one stage to another. Perry (1981) agreed that the model could be reversed to emphasize not the stages themselves, but the transitions between them. In emphasizing transition, change, and developmental instruction, Perry researchers are following Vygotsky's notion that every student has a "Zone of Proximal Development," a next floor of the structure of cognitive development which she is ready and able to construct, especially when teachers build bridges or scaffolds extending into this new territory. "The ZPD is the distance between the actual developmental level as determined by independent problem solving and the level of potential development under adult guidance or in collaboration with more capable peers" (Vygotsky, quoted in Bruner, 1986, p. 73). Perry researchers follow a "Plus One rule," adapted from Kohlberg's research, that students can recognize and appreciate cognitive structures one stage or position beyond what they can currently produce in their own oral and written work. Change occurs primarily at the threshold of the next accessible position. Probably all good teachers sense this; but research using the Perry model brings this tacit knowledge to consciousness.

## **Methods**

This study is not an elegant experiment in developmental psychology. It belongs instead to the *bricoleur* or Rube Goldberg tradition of applied social science, the kind that cobbles together the best available data for post hoc analysis. Until assessment research becomes an established tradition, it will often have to proceed this way; and because student bodies and educational concerns are changing so rapidly, especially in the direction of multicultural inclusiveness, timeliness can compete with elegance as a value in research design. In particular, conventional longitudinal studies, if begun now, cannot yield results for several years. The purpose here is to capitalize now on a wealth of available information: first, research data on cognitive development in freshman programs, gathered at Evergreen by the Washington Center for Improving Undergraduate Education; and second, data in the form of student self-evaluations available in the Evergreen registrar's files. The latter constitute portfolio material which could be examined for evidence of cognitive development. The methods used here to rate and interpret cognitive development are not impeccable, but they represent a best effort to begin this kind of assessment enterprise.

(The methods are spelled out in the remainder of this section. The reader whose interest is primarily in the results and discussion, and not in the method, may wish to skip to p. 13.)

### *Washington Center study*

In 1986-87, the Washington Center administered the standard Measure of Intellectual Development (MID) pretest (version A,  $N = 299$ ) and posttest (version AP,  $N = 179$ ) to students in four Evergreen Core programs, in which the students are almost all freshmen (MacGregor, 1987). These MIDs were rated by William S. Moore of the Center for the Study of Intellectual Development (CSID), according to the method subsequently published by Moore (1989; for information on validity and reliability, see Moore, 1982/86). A single rater was considered sufficient because the rater is the principal trainer of other raters, which minimizes the problem of interrater reliability. The Washington Center did not pair the pre- and post-tests, and so the data are between-subjects data. They merit comparison with aggregate data from other institutions which have used the MID (Appendix B). They indicate the number and percentage of students in each position in the pre- and post-tests, and the mean gain; but they do not address gain or "value added" at the individual, within-subjects level.

*Assessing portfolio materials*

Evidence of considerable increase in MID ratings during the freshman year suggested a cross-sectional study of the sophomore, junior, and senior years, or a longitudinal study of one class through time. The latter was the preferable design, and a promising variant was available: a "retrospective longitudinal study," beginning with data collected over a period of time from students who had graduated recently. Such a study could use portfolio materials - specifically, some of the thousands of student narrative self-evaluations on file at Evergreen.

Assessment researchers are increasingly interested in portfolio materials because these possess what has been called "ecological validity" (Neisser, 1976) and are rooted in the educational process itself (Hutchings, Amiran, Halonen, & Scholnick, 1989). Researchers have become disenchanted with standardized tests which focus on narrow skills but then do not deliver the goods in terms of validity and utility. Specifically, such tests do not discriminate well among skills, and they add little to what can be predicted from grades and SATs (Council of Presidents, 1989). In contrast, portfolio materials represent performance in actual learning processes and show how students work on substantive academic problems, rather than in artificial testing situations. Furthermore, portfolio evaluation can be developmental rather than distracting or intrusive, because preparing portfolio documents can be an important learning experience in itself (Thompson, 1989b).

Evergreen possesses whole vaults of portfolio materials in the form of narrative student self-evaluations (hereafter, SSEs). Faculty and student narrative self-evaluations function at Evergreen as transcript material in lieu of grades. At the end of every program, almost every student writes and files a document called "The Student's Own Evaluation of Personal Achievement," which addresses virtually the same question as Perry's interviews (1970, p. 19): "Why don't you start with whatever stands out for you about the year?" The documents are written under complex constraints, because the student knows she will be discussing them in an evaluation conference with her faculty member and will ultimately file them in her transcript. But these constraints, though they no doubt result in edited presentations of the self, can do little to alter the *cognitive structure* or form of thinking used in the document; and it is on the basis of the cognitive structure of the document, rather than on its specific content, that Perry position ratings are based.

### *Model and Methods: 7*

Initial attempts suggested that SSEs could be read and rated by the same method as MIDs, and that they seemed to provide richer data for potentially more accurate ratings; but this left open the question of whether or not ratings derived from one document should be compared to ratings of the other. In a small feasibility study, 47 freshmen (from one academic program) and 26 seniors (from another program) wrote an MID and an SSE during the same month, and 21 of the seniors also turned in an academic paper. These were rated by William S. Moore, who called in a second rater to reach consensus on all problematic ratings. The results indicated that the mean rating, with freshmen and seniors weighted equally ( $N = 73$ ), ran slightly higher (.095) for SSEs. This was not quite a statistically significant difference (1-tailed  $p = .09$ ); but to make sure that error would fall on the conservative side, in all comparisons which follow between MID and SSE ratings, the former have been adjusted upward to equalize for the .095 difference. In addition, the MID and SSE distributions were significantly different ( $p < .0005$ ), with SSEs tending to be rated less often in stable positions and more often in transitions, particularly the 3-4 Transition. This is reasonable enough: the rater noticed that the SSEs contain "lots of transitional material" (Moore, personal communication), which would follow from the natural tendency to describe personal achievement as movement from an earlier State A to a later State B. Consequently, the graphs and figures below which compare MIDs and SSEs include a correction of the MID. (Specifically, the corrections are Position 2 x 1; Transition 2-3 x .99; Position 3 x .888; Transition 3-4 x 1.14; Position 4 x .98; Transition 4-5 x 1; Position 5 x 1.02).

This comparison between MID and SSE ratings generated an unanticipated finding with important implications for future use of the Perry model in assessing advanced students. In the small feasibility study, the mean freshman MID rating was 3.32, and the mean senior MID rating was 3.80. The mean SSE ratings were 3.48 for freshmen, and 3.82 for seniors. The mean rating for seniors was significantly higher than for freshmen on both MIDs ( $p = .001$ ) and SSEs ( $p = .02$ ), and so one could say with reasonable confidence from this sample that seniors as a group rate higher in Perry position than freshmen as a group. But while individual freshmen tended to receive about the same rating on both instruments, the seniors who rated high on one instrument tended less often to rate high on the other. For all subjects, the correlation between MID rating and SSE rating was an acceptable .52, higher than in other research correlating ratings of MIDs with ratings of other kinds of essays (Mentkowski, Moeser,

& Strait, 1983, vol. 1, p. 166). The MID-SSE correlation for freshmen was relatively high (.61), considering that the correlation between MID ratings and lengthy interviews is only about .70. The sample was not large enough to yield a significant MID-SSE correlation for seniors considered separately; but it clearly would have been lower, with the additional ratings of senior academic essays correlating not at all with the MID and SSE ratings.

In line with these findings from the feasibility study, subsequent study of graduates showed that variance tended to increase by academic year (the standard deviations were freshman year, .57; sophomore, .57; junior, .60; senior, .77). Together, these indications suggest that the rating system is less reliable for advanced students in higher positions, or that senior subjects diverge more in actual performance.

Regarding reliability, first of all, relatively few students have reached Position 4 and above in other studies (Appendix B; Moore, 1982/86; Mentkowski, Moeser, & Strait, 1983). Evergreen seniors tend to move into more advanced positions, where raters have not judged as many instances, the criteria are less clear, and thus the ratings may be less reliable.

Whether or not this is the case, the feasibility study suggested that the art of rating might profit from refinement in rating advanced students and higher positions. While aggregate data seem trustworthy, the overall variance and its increase by academic year allow for skepticism about findings pertaining to individuals. This qualification must be kept in mind when reading the section on "Gain at the individual level," which appears below. In particular, the ratings are certainly not accurate enough to allow assessment, advising, or counseling of students as individuals on the basis of their cognitive development rating (and participants in a recent Perry research forum rejected such use of the model [Thompson, 1988]).

Another possibility is that the rating system is working reliably, but students in higher positions are able to choose and switch their approaches to different tasks. This makes some sense from the standpoint of Perry's theory. MIDs, SSEs, and other portfolio materials differ from Perry's original research in that no probing interviewer is present to elicit the subject's most complex cognitive capacities. Without an interviewer, one might expect lower-division students to perform more consistently on different writing tasks because they have access to a smaller repertoire of ways of knowing and are more likely to address all tasks from the same position. Upper division students who have attained higher positions have a broader repertoire and perhaps de-

cide what cognitive structure - dualistic, multiplistic, or relativistic - will suffice for the task at hand. Individual seniors seemed to "fan out" into the different positions available to them when they addressed different stimuli. This is compatible with Perry's assumption that when a more complex structure is attained, the less complex ones remain available for appropriate tasks. It is also compatible with a rise over four years in mean cognitive development rating without a high correlation in the senior year between one writing sample and another. This fanning out would presumably occur less in an interview situation, in which an interviewer could more consistently stimulate the subject's most complex thought.

These lines of reasoning suggest there will be some unreliability in the findings for upper-division students in this and other Perry research, and they also point toward the conclusion that any future fine-grained studies of advanced students using the Perry model should be based neither on MID essays nor on portfolio materials, but on interviews.

*"Indigenous" graduates, 1986-88*

The subjects for the retrospective longitudinal study, using SSEs to determine Perry position, were all of the graduates of the classes of 1986-88 who attended Evergreen for all four years. These are the "native" or "indigenous" students at a college, those who did not immigrate or emigrate but were admitted as freshmen and completed their degrees without transferring credits from other institutions. Strictness with regard to this criterion - no transfer credits at all, in order to assess the effects of Evergreen alone - resulted in a small population of 165 cases, because most Evergreen students either transfer in or out, or accumulate some credits elsewhere. These students did not necessarily finish in four years, because some remained for all or part of a fifth year, and a few moved even more slowly through part-time studies. And the years were not necessarily consecutive. But all students in the study started as freshmen, stayed at the College, transferred no credits from elsewhere, and eventually graduated.

Studying this group of subjects offered several advantages. The most important is that such students who stay at one college or university give evidence of the greatest effect, the maximum value added, which the institution can generate. Secondly, a longitudinal study was ready-made: the data were already accumulated, without having to wait four years. Third, because it used only the College's own records, the study was not dependent upon volunteers. And finally, the subjects comprised an entire popula-

tion, not a sample. Because they are all possible cases, their cognitive development ratings and the variables which correlate with them hold in the real world. They are not subject to errors of statistical inference such as occur when generalizing from a sample to a population.

The limitation of studying indigenous students is that they are not typical of other students. They are not a sample from which one can make inferences about students who use the College in other ways. The freshmen, for example, do not represent all freshmen, because they are those who subsequently passed through the "filter" of graduation, which makes them atypical. In the long run, one might want to study the impact of the College on students who follow the different paths of transferring in, in different years; transferring out, in different years; and stopping out or studying part-time at different points, with or without accumulating credits elsewhere. This study cannot cast light down these other pathways.

The data for the study were SSEs collected from the College's files and rated by William S. Moore of CSID. The documents were submitted in a random order to the rater, who could not read any student's evaluations as a series and did not know, unless there was an internal reference, whether the document was written by a freshman, sophomore, junior, or senior. When a student changed programs during a year and wrote more than one evaluation, these were batched together and assigned one rating. Scores for gain in rating were computed by subtracting the freshman rating from the rating in the year of graduation. Students who stayed beyond four years received the rating for the final year, even if the fourth year rating was higher. Not all students, however, submitted all evaluations. In the freshman and senior years, on which gain-in-rating scores were based, approximately 85% of the students submitted SSEs. Seventy-two percent submitted both freshman and senior evaluations ( $n = 118$ ). The most important determinant of whether or not a student submits an evaluation is whether or not the faculty member requires it. In cases where the faculty member did not require it, some bias no doubt entered, with more evaluations probably received from students who could say they had done well.

The registrar's records provided data on sex, ethnicity, age at graduation, SAT score (or converted score for other measures), high school grade-point average, and proportion of units earned in different modes of study. The data are presented descriptively, because they pertain to a whole population and do not in most instances involve statistical inference. Statistical analysis was simple, involving correlational analysis,  $t$

test, and chi-square. Multivariate analysis was not very useful, because normal distributions could not always be assumed and many cases contained missing values. Discriminant analysis and factor analysis therefore served primarily as check-ups, to make sure all relevant combinations of variables had received consideration.

Researchers have observed different conventions about treating MID and other Perry-model ratings as ordinal or interval data, with Perry himself (1981) preferring the former.<sup>1</sup> Perry positions would seem to exemplify ordinal categories, and thus should be analyzed primarily with chi-square tests. The measure of central tendency should be the mode, that is, the most frequently represented position. But some researchers have forced the data a bit by treating cognitive development ratings as interval, so that they can compare mean or average ratings (e.g., Moore, 1982/86). There is some justification for this statistical bluff, and so this study acknowledges both positions.

On the one hand, the argument for treating cognitive development ratings as ordinal is clear: the positions do not represent increments of some underlying measure but are qualitatively different, representing different kinds or qualities, not just different quantities, of cognition. And the qualitatively different units should not be considered equidistant on a scale; it might, for example, be harder to move from Position 2 to 3 than from Position 3 to 4.

On the other hand, the counterargument for treating cognitive development ratings as interval data is twofold. The first argument is that the ratings are based on an extended, ten-step incremental scale from Position 2 through Position 5. The scale runs 222, 223, 233, 333, etc., with "223" indicating that the subject is predominantly

---

<sup>1</sup>For lay readers, "interval data" are those which can be measured in a linear manner, using a uniform interval of measure - for example, degrees of temperature or inches of distance. All sequences involve an order, but not all sequences are measurable in uniform intervals. Thus "ordinal data" are those which are ordered in stages, categories, etc., without any underlying unit. Some typical ordinal sequences are military ranks (in the sequence private, corporal, sergeant, lieutenant, etc.) and value preferences (in the sequence agree strongly, agree moderately, don't care, disagree moderately, disagree strongly). In these orderings, the categories form a sequence, but no measure guarantees they are equidistant. Military ranks, for example, are not all "one private" apart; and similarly, no measuring stick guarantees that the difference between "disagree strongly" and "disagree moderately" is equal to the distance between "agree moderately" and "agree strongly."

Translating these considerations to the present context, the issue is whether Perry positions, which are definitely ordinal, are *only* ordinal; or whether they are also a series in some more complex, connected, and measurable sense.



in Position 2 and partly in Position 3, and "233" indicating the subject is still residually in Position 2, but predominantly in Position 3. These ratings are then quantified as 2.00, 2.33, 2.67, 3.00, etc. The case for such quantification is that categories generated on such a scale are more interdependent than categories which are simply qualitatively different. They are quantitative in the sense that the rater has to inquire not just "which kind?" but "how much?" and "to what degree?". This way of looking at the ratings emphasizes the process aspect of development, the continuity, and the reality of transitions, which Perry (1981) judged to be as important as the stepwise aspect. The second argument for interval treatment is that the data consistently behave "as if" they were interval, in the sense that such treatment tends almost always to amplify and refine, rather than to contradict, the analysis based on ordinal treatment. Thus as long as the two treatments are moving in the same direction, it seems helpful to force the data and add the refinement of interval treatment. But if the two methods pointed in different directions, one would have to go with the ordinal treatment.

In the preliminary data analysis for this study (Thompson, 1989a), all Perry data were considered to be ordinal. But when interval analyses were computed, they proved almost always to be complementary, usually offering parallel but more refined measures through time, from the freshman to senior years. Therefore in the present version of this study, when ordinal and interval measures have pointed in the same direction and the latter refine the former, then means have also been computed. In line with this use of both vantage points, the graphs illustrating this study sometimes depict both mode and mean (Figures 5 and 6).

In any event, gain in cognitive development rating, computed by subtracting freshman from senior rating, is the figure which most merits treatment as an interval variable. It is a "mix" of movements from any initial baseline position or transition to any subsequent one (lower or higher), and so it is no longer qualitative or categorical. It has been computed specifically to allow *t* tests of the mean gain by different demographic and academic groupings (male in comparison to female, younger to older, coordinated studies to individual contracts, etc.) By thus computing gain, one can say comparatively whether women in comparison to men, or students who took more coordinated studies in comparison to individual contracts, etc., have gained more in cognitive development rating during four years at Evergreen.

These, then, are the methods used here, first, to study the freshman year; and then to survey the effect of the College on its graduates.

## THE FRESHMAN YEAR

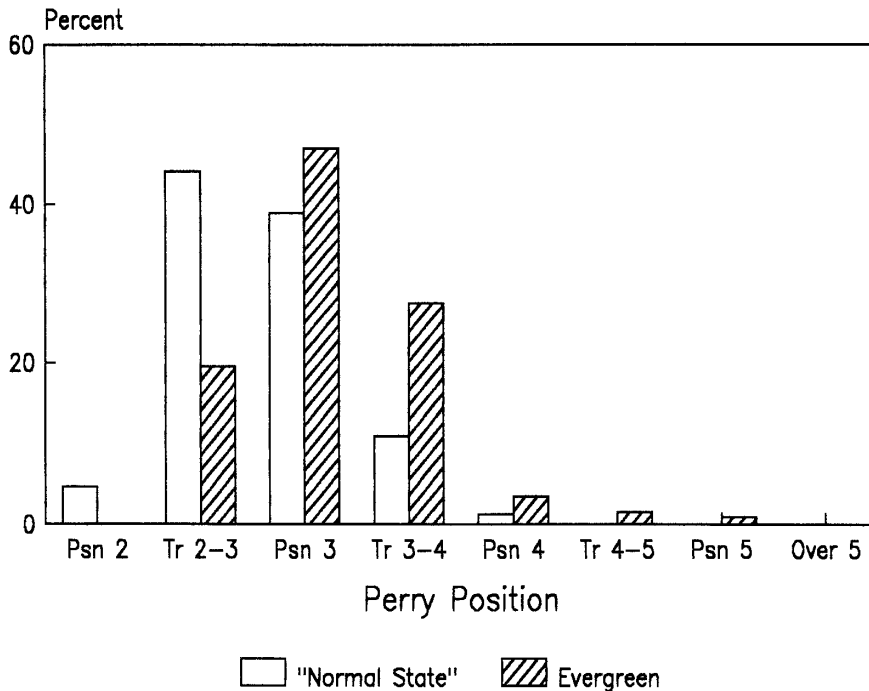
### The Washington Center study

The initial study by the Washington Center for Improvement of Undergraduate Education indicated, first of all, that Evergreen freshmen are an unusual group. The Washington Center pretest and posttest ratings can be combined (equally weighted and averaged;  $N = 478$ ) to approximate a mid-year rating. The Center for the Study of Intellectual Development (CSID) has provided aggregate data for all students who had been rated through 1984 (Appendix B). These data can be taken to represent cross-sectional views, by class, of a purely hypothetical "Normal State" college or university. The Evergreen and Normal State classes can then be compared, revealing that Evergreen freshmen occupy a higher modal position and that they rank, overall, significantly higher in their ratings ( $p < .005$ ), as can be seen in Figure 1 (next page).

In this graph and similar ones which follow, the tallest bar of each shade represents the mode. The mode is the position (or transition) in which the largest percentage of students is to be found. The higher-numbered and more complex Perry positions extend toward the right. Higher values are a matter of distribution (rather than of the height of the tallest bar). Higher ratings for a class or group are indicated, first, if the mode moves to the right; and second, if a higher proportion of students occupy the several positions toward the right. In this instance, the Evergreen freshman mode is higher, at Position 3, and the whole Evergreen profile is further to the right. Whereas most Normal State freshmen occupy the 2-3 Transition (44%) and Position 3 (39%), most Evergreen freshmen already occupy Position 3 (47%), and many are moving through the 3-4 Transition (28%). A few even appear in Position 4 and above (7%). Evergreen students thus tend to be both more advanced and somewhat more diverse in their cognitive positions.

Qualitatively, the difference between Normal State and Evergreen freshmen can be illustrated by reading what students say when they express the different modal positions. The modal Normal State student occupies the 2-3 Transition, where some awareness of diversity of opinion has surfaced; but she still believes that knowledge is

Figure 1:  
"NORMAL STATE" AND EVERGREEN FRESHMEN



Psn = Position; Tr = Transition.

received primarily from Authority and still operates primarily within a dualistic, right-wrong pattern of thinking, as in this illustration:

I tend to trust more what a professor says than what a student says. I have more faith in the teacher, that what he says is correct and concise. Whereas the student might be giving her opinion; it might not be the right one. The teachers are always more or less right (Belenky et al., 1986, p. 39).

The student is in transition because she has noticed that other students *have* opinions, but these remain, as Perry put it (1970), "pre-legitimate": you just wouldn't trust them or try to learn from them, because they are, in this epistemology, either right or wrong, and Authority's opinions tend more toward right (defined by this student as "correct and concise").

In contrast, a typical Evergreen freshman in Position 3 is less likely to locate knowledge-operations outside herself, and more likely to view herself as an agent who

generates knowledge through her own problem-solving activities. She is still reliant on Authority to teach methods and disciplines, but she is concerned less with *what* to learn and more with *how* to learn. Typically she seeks knowledge by grasping the methods and procedures of discrete academic disciplines, without yet trying to think about how one discipline might connect with another one:

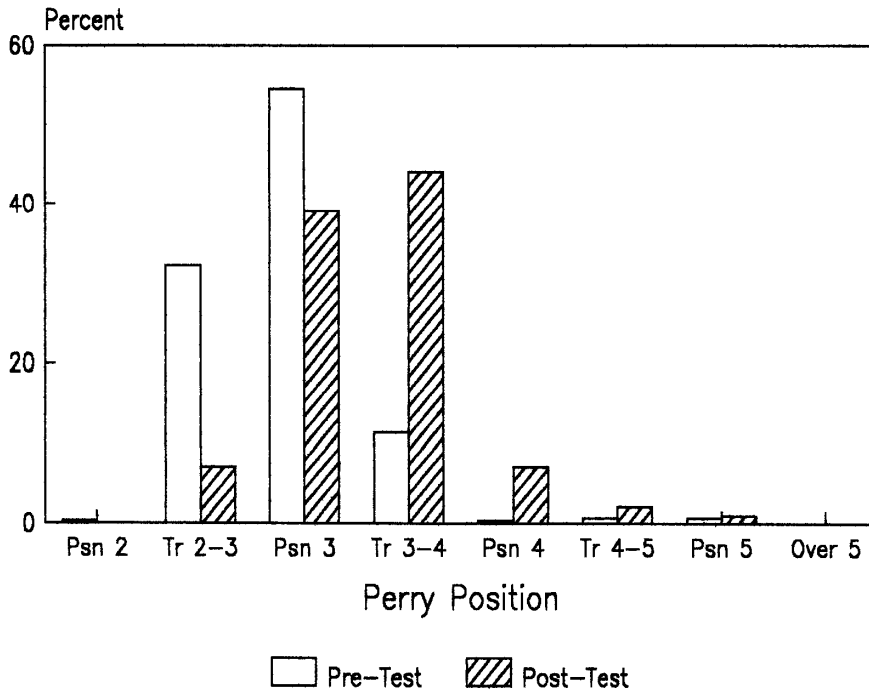
I came into the program expecting to expand my computer programming capabilities. I also expected to develop my writing skills and to supplement my math background as much as possible. I can happily say that this quarter has been successful in that I have accomplished my goals. . . . (Case 77, doc. 199.)

A sizeable number of Evergreen freshmen occupy even more advanced positions, particularly the 3-4 Transition. These students have acknowledged the diversity of opinions which surround them and are beginning to think independently, trying to work out the relationship between opinions and knowledge. They have moved on from concern with "how to learn" to broader, more intellectually self-conscious questions about "how to think," in a world of differing points of view.

Evergreen students thus differ from Normal State students because they arrived from high school with a cognitive head start. In one sense, the College cannot take credit for having done anything yet; but in another sense, it has served through its reputation as a magnet for students who tend to think in complex ways. These students might subsequently educate one another, simply by aggregating and interacting. This study cannot separate some portion of cognitive development which is caused by the curriculum and faculty from some other part, caused by the maturation process in general or by social interaction among students. But it can show that the above-average students who come to Evergreen continue to develop and do not merely rest on their laurels. They move ahead significantly (chi-square  $p < .005$ ) in cognitive development during the freshman year. The Methods section contained reservations about using means or averages; but within these reservations, one can say that the mean pretest rating was 2.92, and the mean posttest rating, 3.31. Thus Evergreen students, on the average, move farther ahead during the freshman year (mean gain = 0.39 position) than students elsewhere move from the freshman to the sophomore year (mean = 0.08), or even from the freshman year to graduation (mean = 0.18; data from Appendix B). In fact, the impressive gain revealed by the Washington Center data for the freshman year spurred the rest of this research project.

The magnitude of change is visible in Figure 2, which shows entering freshman

Figure 2:  
GAIN WITHIN EVERGREEN FRESHMAN YEAR



ratings at the beginning of the year (pretest) and later ratings near the end of the year (posttest). The big shift during the Evergreen freshman year is apparent especially in the great reduction of students in the 2-3 Transition (from 32% to 7%), and the great increase in those in the 3-4 Transition (from 11% to 44%), which has become the modal position. In qualitative terms, these students have begun to wonder about "how to think," have taken note of competing opinions, and are beginning to try to formulate viewpoints of their own. Another important difference is intellectual self-awareness, or awareness of intellectual agency as an aspect of the self. In their SSEs at the end of the year, students developing such awareness might say something as dramatic as the following, about discovery of what Belenky et al. (1986) call "self, voice, and mind":

After this [freshman] year my attitude towards my education has changed. . . . I have a desire to learn and I want to understand things just for my own selfish interest. I no longer feel the need to do well for my parents or

teachers first and foremost; rather, it is myself that I have to please first.

I desperately want to obtain self-satisfaction. (Case 47, doc. 145.)

This formulation has a characteristically subjective tinge. But developmentally, the important point is that the "I" has replaced the external figures, the parents and teachers, as the agent who appreciates knowing and learning. The locus of knowledge has come home to the self, and the student is poignantly, even passionately, aware that the "I" is enhanced and gratified in intellectual activity.

This rapid movement of Evergreen freshmen into the 3-4 Transition is an academic success story, particularly because the 3-4 Transition is as far as most students progress *by the time of graduation* (Moore, 1982/86; and Appendix B). What propels so much movement into the 3-4 Transition? This study cannot address change flowing from the general process of maturation, or the interaction of students outside the curriculum; but it can mention some aspects of the curriculum which at least in theory could foster movement from Position 3 to Position 4. Four such aspects are seminars, faculty roles, interdisciplinary subject matter, and narrative self-evaluation.

The experience which freshman students discuss most often in their SSEs is seminars, or "seminaring." All freshmen in Core programs are involved in seminars from the moment of their arrival for Orientation Week. Evergreen seminars are discussion groups devoted to dialogue (unlike university research seminars, often given over to serial monologue). In Evergreen dialect, "seminaring" is the participatory activity of student discussion, and students "seminar on" or "seminar about" a topic or more often a book (usually at the rate of a book a week, in a full-time program). As an organized group, a seminar consists of about 20 students and a faculty member (never a teaching assistant). Each faculty member in an interdisciplinary program leads one seminar, which is usually referred to by his or her first name ("I'm in Evelyn's seminar"). These groups usually reshuffle at the end of every quarter in year-long programs, so that students get to know most of their fellow students and work with most members of the faculty. The Evergreen verb "to seminar" emphasizes exchange of student opinions, with the faculty normally playing a guiding rather than a directive role. A seminar is thus primarily a dialogue among the students themselves, rather than a turn-taking conversation with the seminar leader. Seminars command a kind of primary group loyalty at the College and are generally perceived by students and faculty alike as the heart of interdisciplinary programs and the place where most of the action occurs.

Many freshman SSEs describe the discovery of an individual voice in seminars,

sometimes after a long period of fear and frustration:

When I registered, . . . I knew that the most difficult part for me would be participating in seminar. . . . At the beginning of the quarter, I found it very difficult to voice my ideas. I hesitated to speak for fear of being glared at by 17 total strangers. However, as I became familiar with these strangers, and discovered that they were actually listening instead of glaring, I began to increase my participation in the seminar. (Case 77, doc. 199, Fr.)

The most sophisticated description of a seminar which appeared among freshman SSEs describes the group as a model United Nations, with the author learning to manage the situation consciously from a perspective of contextual relativism:

I enjoyed participating in our [seminar], choosing to view it as a microcosm of the world. . . . I found I could equate the seminar with a group of nations, some whose voices were louder than others and who were unwilling to listen, some [who] let the discussion be others' responsibility, while still others had good things to say but couldn't be heard. It was interesting to try to find ways of encouraging the quiet ones to speak and the dominant ones to listen more. Using force was of course inappropriate, and finding alternative ways challenging. I discovered I had strength to speak out and discuss, as well as encourage others to do so. Like everyone else I sometimes cut people off and needed to listen more. (Case 57, doc. 158.)

From the content of SSEs, one can infer that exposure to differences of opinion in seminars is what propels many students out of "received knowing" and "procedural knowing" into self-conscious development of their own opinions - in other words, into the 3-4 Transition.

Because seminars are participatory, engaging, and even stressful, they loom large in the Evergreen environment and are especially likely to be mentioned in SSEs. But other dimensions of Core programs may be equally decisive for cognitive development, without being so frequently commented upon. One is the role of faculty, particularly in team teaching, which departs decisively from the learning structure of high school. While individual faculty members give expert or authoritative lectures, nevertheless students see faculty less as individual experts than as seminar leaders, co-learners, and members of a teaching team. In interdisciplinary programs faculty must function a good part of the time outside their own areas of expertise. The whole context of student contact with faculty thus tends to unseat what Perry (1970) called "capital-A" Authority of authoritative persons, and to foster the "lower case" authority of ideas that make sense through communication. While there are many ways to play the role of Evergreen faculty member, almost all ways tend to downplay "received knowing" and procedural expertise, and to present faculty members as people who are themselves trying to bring their backgrounds to bear upon new texts and issues, in order to reach

their own conclusions.

For most freshman, this new relationship with faculty presents a dramatic contrast to the familiar, traditional classroom. High school was organized around an implicit theory of knowledge. The separate classroom with the single teacher facing her students alone from behind a desk or lectern taught through its formal arrangements that knowledge is not dialogic, but is received from Authority. The traditional faculty-student relationship taught that knowledge is unitary in source and hierarchic in transmission. Indeed, the whole organization of the traditional classroom can be construed as a defense against dialogue. Such a context reinforces the cognitive structure of "received knowing" and tends to place a ceiling on ways of knowing. In contrast, lectures from different viewpoints and exchanges between team-teaching faculty in Evergreen Core programs may be, for some students, the first examples they have witnessed of adult differences of opinion expressed reasonably. Faculty participation in seminars carries the message that while teachers are still present to encourage learning, they need not do so as Authorities promulgating Truth. However the faculty member's role is played, it tends to be defined by a Socratic assumption that discovery of ideas and refinement of opinions emerge through dialogue. In comparison with typical high school experience, this is a radically new pedagogy and epistemology. In promoting the examination of different opinions, the role (or roles) assumed by faculty directly promote the epistemology of Position 4.

Further in the background, but equally conducive to the 3-4 Transition, lies the College's long-standing decision to emphasize interdisciplinary studies. Such studies appear in Core programs at the front of the curriculum, with disciplinary foundations filling in later, in advanced programs and contracts. The choice to teach interdisciplinary studies before disciplinary studies is in effect a decision *to teach relativism*. Core programs are thematic, and disciplines and texts are chosen because they *relate to* an interdisciplinary program theme, such as "Human Development," "Matter and Motion," "Classical and Modern," or "Society and the Computer." The College's accreditation report outlines the basic assumptions and strategy:

. . . Instead of the conventional institution's tendency to present a great work of literature or philosophy as . . . [an] illustration of the discipline of English or philosophy, our faculty tends to believe that the starting point in undergraduate education ought to be the hypothesis that assorted discipline-specific perspectives, methods of analysis, modes of inquiry and canons of evidence ought to be brought to bear upon important common subject matter. . . .

Conceptually, . . . Core programs contribute to fundamental educational



goals because they specifically address the most important obstacle to serious learning today: the multitude of voices arising from academic specialization. . . . The Core program attempts to bring into the academic foreground major conflicts and major agreements concerning the meaning of the past, the shape of the present and the prospects for the future. Indeed they "teach the conflicts" not only within the academy, but within the society (TESC, 1989, p. 21).

The design of such a program by the faculty is itself an exercise in constructed knowing, with the program theme setting the context for relating texts and disciplines to one another - a prime example of what Perry called "contextual relativism." In other words, just as the high school curriculum favored received knowing and the traditional college curriculum favors disciplinary multiplicity, so does the curricular design of interdisciplinary programs teach contextual relativism. The context teaches just as effectively as did the high-school classroom, but it supports a different level of cognitive complexity.

Finally, the practice of writing narrative evaluations begins in the freshman year and may itself be a stimulus toward cognitive development (Thompson, 1989b). At Evergreen, narrative faculty and student evaluations replace, rather than supplement, grades. Test-taking for grades forces the student to prove she is a received knower or problem solver (Positions 2 and 3), and only rarely an independent thinker (Positions 4 and above). In contrast, narrative self-evaluation transforms the author into an evaluator. Beginning with informal "in-house" evaluations written in most programs at the end of each quarter, students are introduced to the regular practice of asking what their education means to them. It is hard to deal with this question from the standpoint of received knowing or of methods and procedures alone, for the ritual itself calls for reflection on the relation of the knower to the known, and the impact of education on the self. Once the ritual has been completed, it enters the student's educational outlook, and what is studied thereafter is in some degree undertaken in anticipation of evaluating its impact upon the self. Narrative self-evaluation is thus itself a spur to forming one's own perspective and entertaining Position 4.

These same curricular factors - seminars, faculty roles, interdisciplinary studies, and narrative evaluation - also help to explain another aspect of gain during the freshman year: the arrival of an appreciable number of freshmen (10%) in Position 4 and above. The SSEs suggest that these students fully appreciated not just the particular texts and components of their programs, but what the faculty were thinking about - the overall interdisciplinary theme and structure. Within the seminar system, these students are like unofficial teaching assistants, able with faculty and peer support to in-

struct their fellow students in more complex cognitive operations. Helping these students to find a voice and an appropriate role in seminars is a particularly important seminar teaching skill.

On the other hand, at the end of the freshman year some students (7%) still remain in the 2-3 Transition. The good news is that they are in transition, not deeply entrenched in dualism (e.g., in religious absolutism); but the problem is that they are apt to feel defeated when seminar dialogue flies past above their heads. This is a problem about the sheer diversity of students in Evergreen classrooms, which increases when transfer students (not included in this study), who are probably more like Normal State students, join Evergreen students during the remaining three years. Belenky et al. (1986, ch. 9) urge positive learning experiences for every student, particularly those having the most difficulty overcoming previous learning disadvantages. From the broad distribution in Figure 2, one can see why striking a synergistic balance between, on the one hand, positive experiences for disadvantaged students, and on the other hand, supportive experience for the equally large group who are ready for much more advanced cognitive operations, is an important trick of the seminar teaching trade.

### **Hypotheses**

The results for the freshman year spurred further study. In an optimistic mood, one might have expected freshman gain to establish a constant or accelerating trend, with development breeding more development exponentially. But such an expectation would far overreach the pattern of evidence from other institutions (Appendix B). One thus has to consider a number of less optimistic hypotheses. For several reasons, a good showing in the freshman year might not predict a good showing for later years. The following hypotheses might rival the optimistic one:

1. *The artifact hypothesis.* Freshman gains might not be sustained in subsequent years because they might be, to some extent, artifacts of the rating system. Specifically, widespread freshman experience at Evergreen with seminars might have raised the ratings in a way that would not occur again in later years. The Perry model and the rating manuals based on it (Mentkowski et al., 1983; Moore, 1989) emphasize a contrast between Position 2, in which all knowledge and learning flow from Authority and are received passively, and the later positions, in which knowledge and learning are generated interactively, working with texts, teachers, and peers. Experience at Normal State with the lecture-and-test mode of instruction would tend to reinforce the

assumptions of Position 2, that knowledge is something received; whereas the seminar method at Evergreen undermines the assumptions of Position 2 and validates those of Position 3 and above, that knowledge is something generated among colleagues. A student attending a program based on seminars needs no stroke of genius to discover that she can learn from her peers; and her accelerated movement from Position 2 to 3 would justify no prediction of equally accelerated movement to 4 and 5. But in contrast, if a student is familiar only with lecture courses of 100 or more and with objective tests, then an interactive approach to learning is no artifact; it is a hard-won cognitive achievement that suggests unusual individual involvement in the learning enterprise. So the artifact hypothesis predicts that Evergreen students move easily, almost automatically, to Position 3; but that subsequent gains will be modest in relation to this initial surge.

2. *The sophomore slump hypothesis.* Folklore and research suggest that in the freshman year, student attention tends to focus more directly on going to college, but that the sophomore year is often different in quality (Feldman and Newcomb, cited in Martin, 1982, p. 87). When the initial challenge of "making it" in college has been met, academic anxiety tends to decline and interest can turn elsewhere. For traditional-age students it is time for social exploration (falling in love). Life may be more exciting than ever before, but hours and concentration devoted to academics may decline. The sophomore slump hypothesis thus predicts less rapid cognitive development, or little increase in cognitive development ratings, among sophomores.

3. *The community college hypothesis.* Two more hypotheses predict that Evergreen education will suffer from ceiling effects and that gains in the freshman year will not be sustained in the junior and senior years. One hypothesis was offered by a visiting professor from Hampshire College (which emphasizes highly individualized instruction), who claimed that Evergreen had invented the perfect two-year college, but that the emphasis on interdisciplinary studies and the absence of majors precluded upper-division teaching and learning. If this were true, then Evergreen gains should level off after lower-division work.

4. *The College of Relativism hypothesis.* Evergreen has a reputation for allowing students, in the language of the sixties, to "do their own thing." The Evergreen curriculum actually provides a great deal of structure; but viewed critically, it might seem to epitomize the cafeteria style. Because students can design their own sequence of studies without being confined to majors, the College might seem to encour-

age the kinds of individualistic or subjective judgment that surface in Perry Position 4. A ceiling on cognitive development might thus be built at this level into the basic structure of the curriculum. From this standpoint, Evergreen would be Allan Bloom's nightmare (Bloom, 1987), a college which institutionalizes relativism and subjectivism with a vengeance. This view predicts that a relatively high number of students might reach Position 4, but a comparatively low number would move beyond.

The optimistic hypothesis of accelerating cognitive development and these four rival hypotheses were all explored through a "retrospective longitudinal study" of native or indigenous graduates in the classes of 1986-88.

## **"INDIGENOUS" GRADUATES, 1986-1988**

The retrospective longitudinal study of 1986-88 graduates did not support the most optimistic hypothesis of gains rising right off the scale of other institutions, but it did reveal Evergreen in a strongly favorable light. Freshman gains seemed to be followed by a sophomore and junior consolidation. But the senior year contained a second, smaller surge in cognitive development, placing Evergreen in a quantitatively and qualitatively favorable position compared to results elsewhere, and effectively refuting the community college and College of Relativism hypotheses.

### **Description of the graduates**

The hypotheses listed above will be reviewed in the course of reporting on the retrospective longitudinal study; but to begin with, who were the 165 subjects, the population of indigenous students who never emigrated, never studied elsewhere, and finally graduated?

First, they were predominantly white (93%). Only 7% were persons of color. This low proportion could have been anticipated from the composition of the student body at the time: only 10% were then persons of color, and a substantial proportion were enrolled at a satellite campus where many students arrive with transfer credit and thus do not number among the four-year graduates. (This program is being studied by another research team.)

In terms of gender, the indigenous graduates were almost exactly equally men and women. In terms of age, only about 12% were born before the 1960s. In the vocabulary of student personnel administration, the modal four-year student was a "high school direct" who entered at 18 and graduated at 22. Half were age 23 or younger at graduation; 75% were 25 or younger. But 25% were of non-traditional age, having "stopped out" for three or more years; and 14% were over 30. Only 29% were first-generation college students. One or both parents of the other 71% had attended some college. The students' own economic position and that of their families was unknown, but the high percentage of college-attending parents suggests a fairly high percentage of

middle-class families which might be able to help a student financially to stay in school.

In terms of their high school preparation, the four-year graduates were diverse indeed. Their SATs and GPAs proved to correlate moderately with one another ( $r = .38$ ), but not with gain in Perry position. On SAT or equivalent tests (not required by Evergreen;  $n = 89$ ), the mean was 942. When these students took the tests, the national average was just under 900. Those at risk of low scores could avoid the test; and if reporting was voluntary, the tendency among those who took the test must have been to under-report low scores. The mean was thus not unusual, but the range certainly was: the lower quartile ran from 410 to 810; the middle 50% from 820 to 1070; and the upper quartile from 1080 to 1330. On GPA, the mean was a solid B (3.01), but the spread is again noteworthy. The lower quartile ranged from 1.33 to 2.56, the middle 50% from 2.57 to 3.41, and the upper quartile from 3.42 to 3.98. These dispersions suggest that some entering freshmen might have been struggling with basic math and literacy, whereas others could have been admitted to almost any college or university of their choice.

The four-year graduates clearly included some students whom a traditional admissions officer would have rejected, predicting they would have difficulty with college. But Evergreen has been more adventurous in admissions. These high-risk students whom one might regard as least prepared did not fail or drop out; rather, they had to graduate in order to appear in the study. The diversity of Evergreen students' high school backgrounds might have raised problems about how to reach common ground in seminar discussions, but poorer preparation by SAT and high school standards proved to be only a moderate, not a major, barrier to graduation from Evergreen (see comparison below, between all entering freshmen and those who later graduate).

The four-year students tended to follow individualistic rather than typical paths through the different modes of study in the Evergreen curriculum. Yet the simple average underlying this diversity is close to the pattern sometimes recommended within the College as normative or ideal. The average is approximately two years of coordinated study (interdisciplinary programs, usually basic or intermediate); a year of group contracts (usually more advanced and more focused on one or two disciplines); two quarters of individual contracts (either independent study or internship), and the equivalent of a quarter's work in four-unit courses.

The results for the indigenous graduates in the retrospective longitudinal study

can be reviewed under three main headings: first, *class profiles* - those results which yield cognitive development profiles by class, using primarily between-subjects data and moving from the freshman through the senior year; second, *individual gain* - those results which indicate how many individuals gained, stayed the same, or declined in cognitive development rating; and third, *demographic and academic groupings* - those results which locate tendencies toward more or less gain within the demographic and academic categories of sex, age, ethnicity, family's college experience, high school preparation, and mode of study.

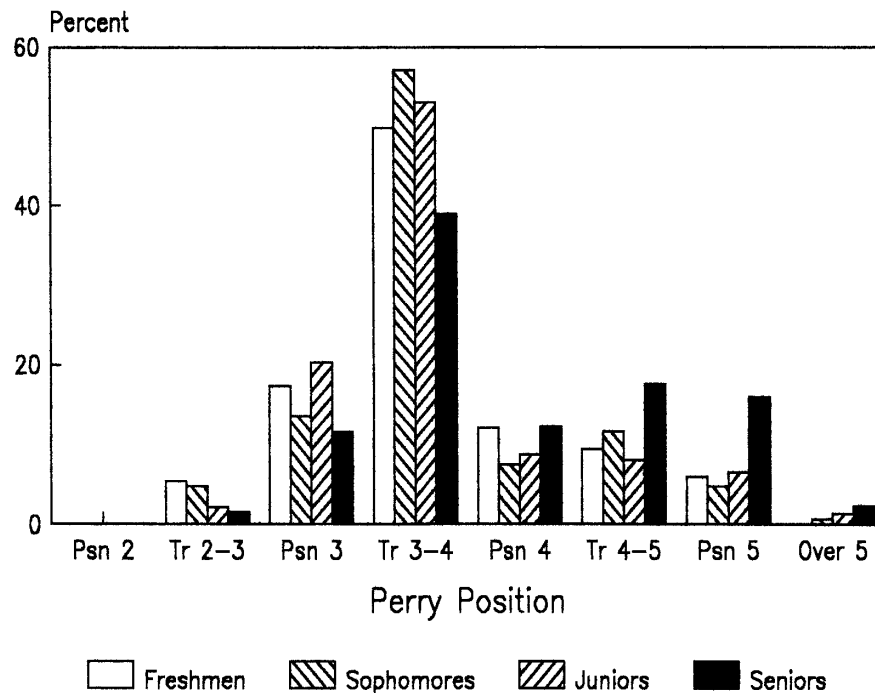
### **Class profiles**

#### *The Big Picture: All four years*

The overall picture can be presented as a profile of the graduates distributed as freshmen, sophomores, juniors, and seniors, through the different Perry positions (Figure 3). Initially this picture may be hard to read, but it can serve as a reference for the more discrete explanations which follow.

The main points which emerge in this picture are first, that the overall increase in cognitive development ratings is moderate between the freshman and the junior year; second, that although there is no sophomore slump, there is little sophomore gain; third, that students seem to become most "normal" or even cognitively conservative in the junior year; but that fourth and finally, the senior year heralds a substantial movement in cognitive development, visible as the black bars become taller toward the right side of the graph. In the senior year, the mode remains the 3-4 Transition (39%); but nearly half of the senior class (48%) has moved beyond that point, into the cognitive structures of Positions 4 to 5. Twelve per cent display a capacity for "late multiplicity" or simple relativism, Position 4; and 36% have moved on to the 4-5 Transition and Position 5, demonstrating capacity for what Perry (1970) calls "contextual relativism" and what Belenky et al. (1986) call "constructive knowing," including the capacity to construct models and theories. These results contrast markedly to the usual outcome of attending college, given the fact that seniors elsewhere "are primarily in transition between positions 3 and 4 and in stable position 4" (Moore, 1982/6, p. 2). These results definitely put to rout the hypotheses that gain at Evergreen was only an artifact of the rating system for the lower positions, or that Evergreen is only a big community college, or that it is a College of Relativism that does not help students to advance beyond Position 4. What will finally need explaining is the second surge of cognitive devel-

Figure 3:  
ALL FOUR EVERGREEN CLASSES



opment in the senior year, which places so many Evergreen students in Perry positions which comparative data indicate are not often achieved.

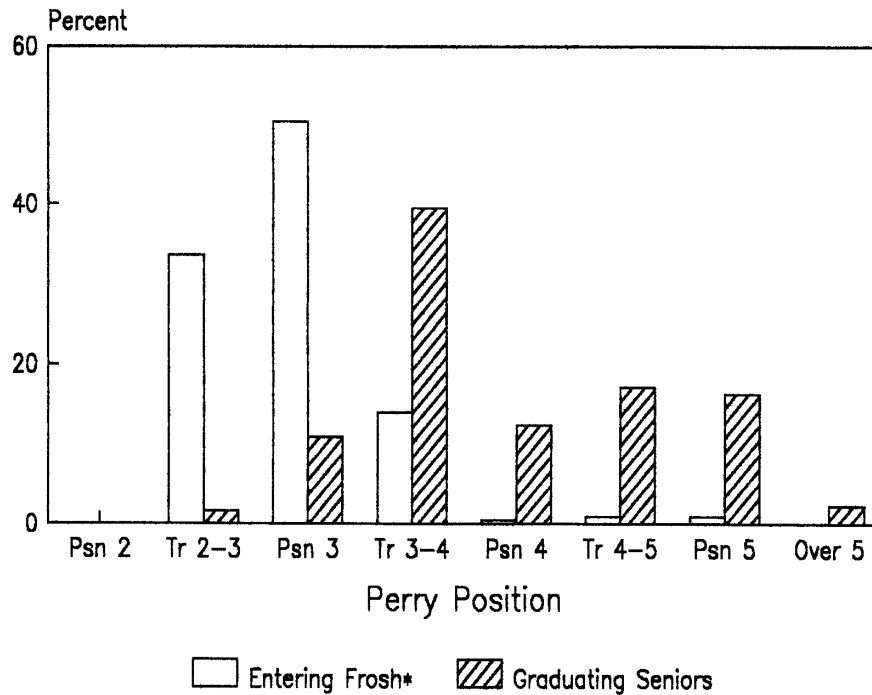
*Freshmen compared to seniors*

To illustrate the maximum effect of Evergreen education upon its students, the appropriate comparison is of the entering freshmen of the Washington Center study to the graduating indigenous seniors (Figure 4). To begin with, the senior profile is far to the right of the entering freshman profile, indicating change which might be characterized as dramatic. Furthermore, the senior distribution is no longer so tightly grouped. Four years at Evergreen thus produces not only dramatic movement in cognitive development, but considerable diversity in position achieved.

This diversity can be appreciated through a series of qualitative illustrations.



Figure 4:  
ENTERING FROSH AND GRADUATING SENIORS



\*Includes SSE-MID correction factor.

The 12% of students who reach Position 4 are directly aware of what Perry called "multiplicity," or what others, particularly those critical of this style of thinking (e.g., Bloom, 1987), would call a simplistic relativism.<sup>1</sup> In Evergreen SSEs, this position is often in evidence when students discuss the multicultural emphasis of the curriculum:

I have expanded my cultural awareness immensely. . . . By being made aware of the significance of cultural differences, . . . I have become much more open to, and accepting of, beliefs that differ from my own. I have learned that we all have so much that we could learn from each other if we would just open our minds and hearts. . . . (Case 61, doc. 168.)

The student is "open to" different opinions, but this is the kind of openness that Bloom fears. The student is not yet inquiring whether some opinions rest on more sensible

---

<sup>1</sup>Perry does not like the term "relativism" for Position 4, because it suggests a grasp of relationships, particularly of the significance of one concept relative to another, which does not begin until Position 5 (personal communication).

premises, or are more logical in form, or are more relevant to the context, than others. The statement is thus a good example of Position 4, and also of the value of multicultural education as a stimulus to cognitive development. From the standpoint of the Perry model, this multiplicity or relativism is a necessary and desirable step. Though conservatives might decry it, simple relativism is a threshold that must be reached: there is probably no way to arrive at more complex cognitive positions without first encountering the ineradicable component of relativity and subjectivity in all points of view.

The 4-5 Transition (18%) can be understood as the waypoint between the multiplicity or relativism just illustrated, and Position 5 (16%), contextual relativism or constructivism. The epistemology of Position 5 changes the student's relation to authority and often helps to stabilize her sense of identity. It can bring a new feeling of confidence that one is no longer the subject of external intellectual authorities and can now function autonomously in the world of ideas (considered broadly, to include the full range from scientific to aesthetic ideas). The student often realizes that she herself is now doing what the faculty, or others who claimed some intellectual authority, had been doing all along, and that she has in an important sense joined an intellectual community (Perry, 1970). The relativism of Position 4 functions as a solvent, dissolving certainties and absolute truths; but this uncertainty opens up the new question of Position 5 (and beyond): how can one think about the relative merits of different inductive generalizations and provisional truths? A "constructivist" or "contextually relativistic" epistemology answers this question by suggesting that the individual can construct limited-range humanistic, scientific, or artistic models (ideal types, middle-range theories, or ideas of the state of the art, to draw examples from different disciplines) which help one to address a given range of phenomena without laying claim to certitude. In Evergreen SSEs, a typical expression of this kind of complex thought is formulation or appreciation of systems theories:

I learned that politics and economics are a system, an integrated part of society, each one affecting the other and affecting the rest of society as society affects it. These individual systems should not be viewed independently but rather viewed in their relationship to one another. (Case 69, doc. 183.)

This sort of thought does not claim that anything is true absolutely, only that some statements about the relationship among things are more reasonable than others. Such thinking has moved toward much greater complexity than the modal entering freshman position (Position 3), illustrated above by the student who conceived learning as simply

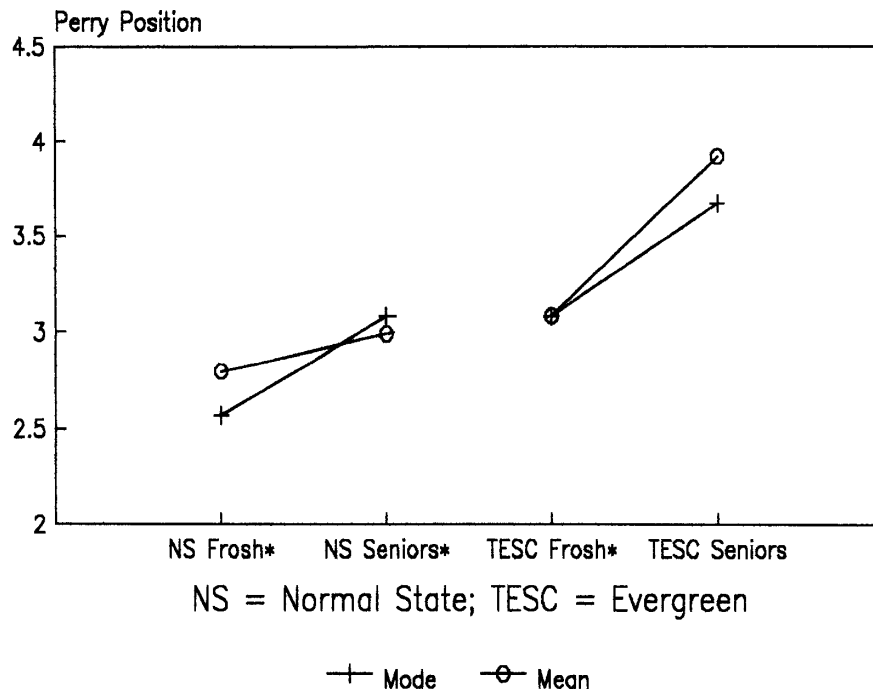
skill-acquisition (improving her programming capabilities, math background, and writing skills).

Besides comparing Evergreen freshmen and seniors internally, one can make a rough comparison between freshman and seniors at Evergreen and those at Normal State. A year-by-year comparison is impossible because the CSID normative data are cross-sectional, not longitudinal: they offer a set of class portraits of everyone enrolled in each class, but they do not offer a time-lapse picture, like the Evergreen longitudinal study, of the same students passing through each year. The Evergreen students are, in contrast, the same ones, as they pass from matriculation through the "filter" of graduation. But some comparison is possible for just the freshman and senior years, comparing the "Normal State" freshman cross-section with the Washington Center freshman cross section (averaged for mid-year), and the Normal State seniors with Evergreen's fourth-year seniors (eliminating those who stayed for a fifth year, who would give Evergreen an artificial boost). The rationale for comparing these senior groups is that the Normal State seniors are about to graduate, or are about to become the "indigenous graduates" of Normal State. (They may have transferred from one institution to another, but "Normal State" is a construct, not a place; and so it is fair to say that the Normal State seniors have all been attending *some version* of Normal State.) Though all of the Normal State seniors may not graduate, it is unlikely that the difference between Normal State seniors and Evergreen seniors could be accounted for by the failure of some of the former to graduate.

If one is willing to entertain a comparison based on these assumptions, then the results appear in Figure 5. Measured by change in the mode, the overall freshman-to-senior increases were about the same: from 2.57 to 3.08 at Normal State, and from 3.08 to 3.67 at Evergreen. (The fractional figures, which would not usually represent a mode, result from the correction factor which makes MID and SSE ratings comparable.) But measured by change in the mean, the Evergreen freshman-to-senior increase was definitely the greater: from 2.80 to 2.99 at Normal State, and from 3.08 to 3.92 at Evergreen. The mean increase at Normal State was a small fraction of a position (0.19), whereas at Evergreen it was nearer to a whole position (0.84).

These differences in terms of cognitive development between Normal State and Evergreen education become clearer when one considers the modal or most common outcome. The usual function of Normal State is to help the preponderance of students in the 2-3 Transition to move beyond dualism and received knowing into separate, pro-

Figure 5:  
NORMAL STATE & TESC FRESHMEN & SENIORS

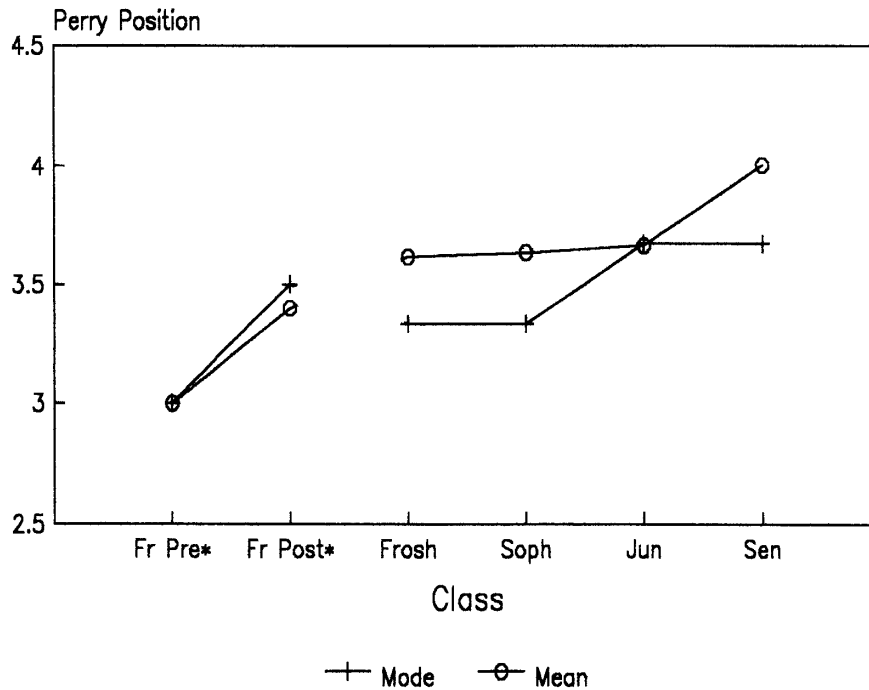


\*Includes SSE-MID correction factor.

cedural knowing, or mastery of disciplinary methodologies, which is typical of Position 3. "Normal" education enables the "Normal" student independently to apply methods of problem-solving in domain-specific and domain-appropriate ways. This would be the kind of gain or value added most frequently experienced at Normal State between matriculation and graduation.

Because its indigenous graduates end up in more diverse positions, one cannot say that Evergreen has any one typical function. But the senior ratings indicate that Evergreen is engaged in a distinctly "nonnormal" enterprise. It tends to matriculate students who are able, or are soon able, to accomplish the intellectual operations most in evidence among Normal State's seniors. And it successfully encourages complex, relativistic thought. In qualitative terms, *this is Evergreen's value added: its graduates tend to think differently, using different epistemologies or theories of knowledge than*

Figure 6:  
FRESHMAN AND SUBSEQUENT GAIN



\*Includes SSE-MID correction factor.

*graduates of other colleges; and these epistemologies are suited to problems of greater complexity.* Although no data exist comparing elite and nonelite colleges which have used the Perry model, one can at least speculate that Evergreen, though a relatively inexpensive public institution with relatively open admissions policies, would rank from the standpoint of promoting cognitive development among a fairly small stratum of institutions, most of them probably more costly to attend and more restrictive in admissions.

So far, the comparison between freshman and senior ratings has emphasized Evergreen's high overall gain. But freshman gain can also be compared to subsequent gain (Figure 6). This graph shows that although subsequent gain was impressive, nevertheless all gain after the freshman year added up to no more than gain during the freshman year, which was definitely the year of the College's maximum impact on its

students. The lines to the left represent the Washington Center's measures, and the lines to the right represent the longitudinal study. The mode shifted from Position 3 to the 3-4 Transition, or one-half position, in the Washington Center study; and from the beginning to the end of the 3-4 Transition (one-third of a position) in the longitudinal study. (The difference in magnitude is probably an artifact, because the former study reported ratings by half-positions and the latter by thirds.) The mean gain was the same in both studies (0.40), indicating that increase in cognitive development rating during the freshman year was equal to all subsequent gain, from the sophomore through the senior years.<sup>2</sup>

One more comparison is possible for the freshman year, between all freshmen and those who graduate (Figure 7). Are students who finally graduate any different than freshmen in general? Comparison of all freshmen (the Washington Center posttest) with the freshmen of the longitudinal study indicates that the two means and distributions are significantly different ( $p < .005$ ; the means are 3.33 for all freshmen, corrected, and 3.61 for freshmen who later graduate). The difference stems from the higher number of freshmen from the Washington Center study who finished the year in Position 3, in contrast to the higher number of freshmen who will later graduate who have already reached Position 4 and above. If there were no difference, it would suggest that students leave the College according to criteria independent of cognitive development (e.g., finances, social life, etc.). Although the measure used here is a rough one, the difference in means and profiles suggests that position in the Perry scheme is related to retention: those students initially employing more complex cognitive operations are a little more likely to stay at Evergreen for four years, and those who initially employ less complex cognitive operations are a little more likely to leave.

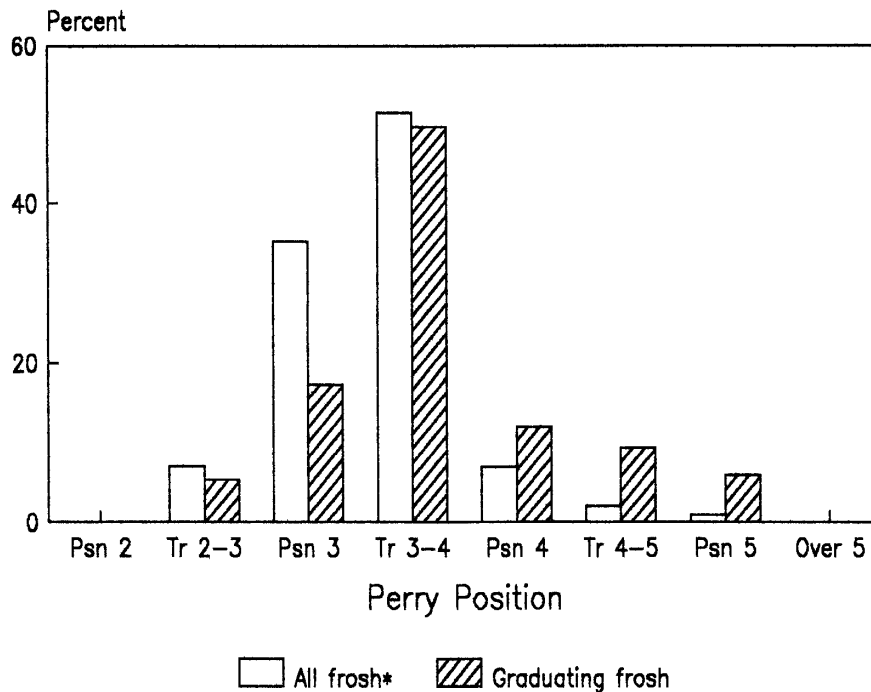
### *Sophomores and juniors*

But the impressive gain of freshmen is not sustained in the sophomore and junior years, which seems to uphold the artifact hypothesis, the sophomore slump hypothesis, and possibly the community college hypothesis. Figure 8 indicates that no actual slump occurs in the sophomore year, but evidence of development is notable mainly in its absence. The modal Perry position (3.33) does not increase, and the

---

<sup>2</sup>This very substantial proportion of gain in the freshman year should be encouraging to the Washington community colleges which have adapted Evergreen's coordinated studies programs to their campuses. These programs appear to have great potential for encouraging cognitive development.

Figure 7:  
ALL FRESHMEN AND THOSE WHO GRADUATE

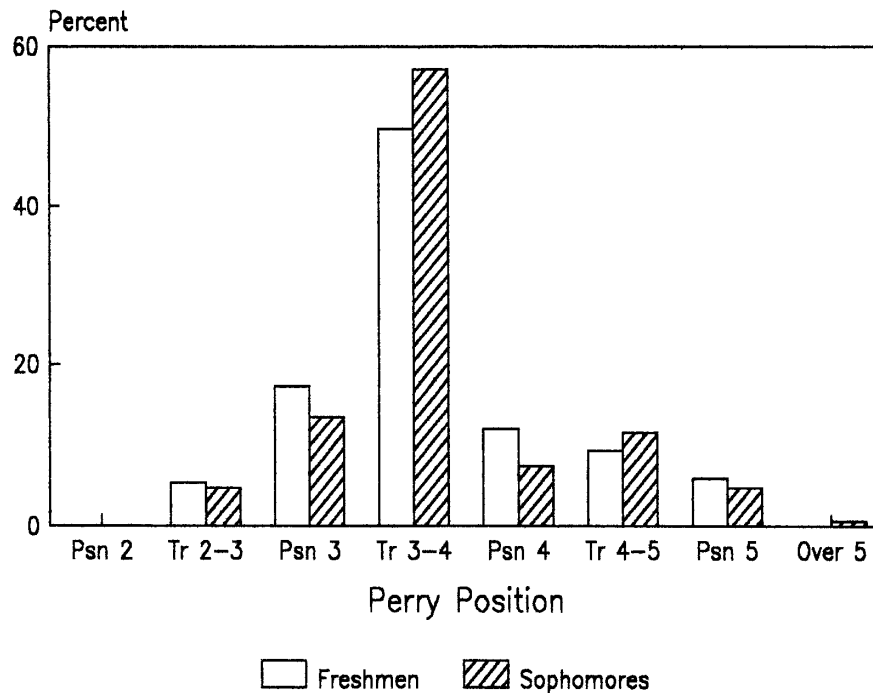


\*Includes SSE-MID correction.

mean increases only marginally (from 3.61 to 3.63). The proportion of students in Position 3 declines and the proportion in the 3-4 Transition increases, accounting for some overall gain; but the percentage in Position 4 and above actually declines, from 28% to 25%.

This represents a tightening up, a normalization, around the 3-4 Transition, a trend which continues even further in the junior year (Figure 9). The most striking feature of this graph is the increase, rather than decrease, of juniors in Position 3. Juniors apparently tend to "regress" to preoccupation with methods, problem-solving, and what Belenky et al. called "separate" and "procedural" knowing. The junior mode moves to 3.67 (still within the 3-4 Transition), but the mean increases hardly at all (from 3.63 to 3.66). The proportion above Position 4 stays exactly the same (24.6%), whereas the increase in Position 3 is accompanied by decline in the 3-4 Transition.

Figure 8:  
FRESHMEN AND SOPHOMORES

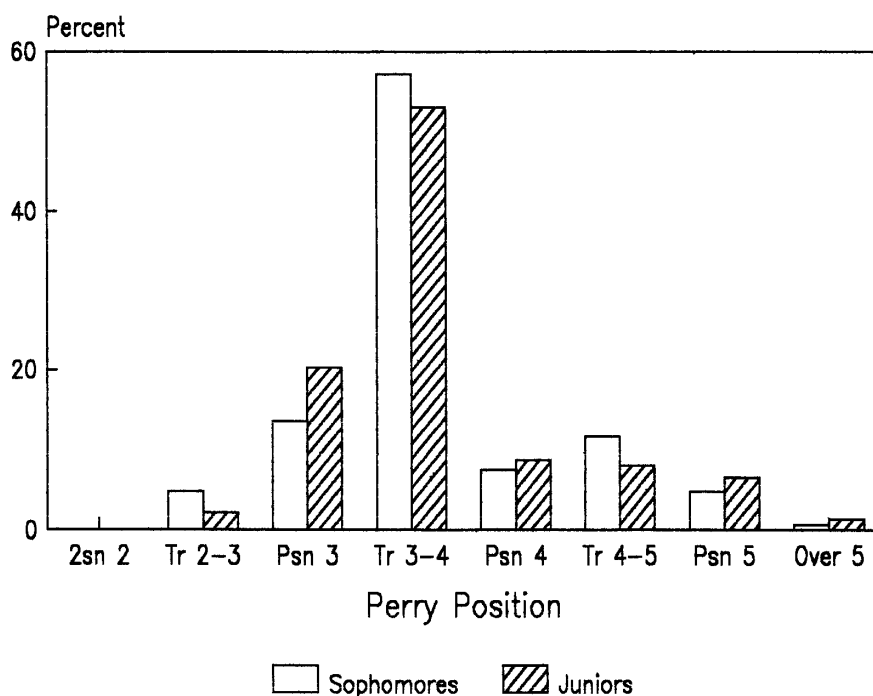


The faculty, one might infer, teach the sophomores and juniors to use more consistent, but not more complex, cognitive structures; or the sophomores and juniors teach this to one another. The result is that overall, the juniors seem to be the most methodical and conventional students of all. The result in these two years is normalization, with the mode still considerably above the cross-sectional data for other colleges and universities, but with little positive change in cognitive development ratings and fewer students far ahead of the norm.

The artifact and sophomore slump hypotheses are generally congruent with the data and together help explain why there is so little sophomore-junior change. The artifact hypothesis is useful in suggesting that experience in seminars probably generates the high gain within the freshman year, and the hypothesis correctly predicts that the same rate of gain cannot be sustained, because no comparable stimulus arises in the



Figure 9:  
SOPHOMORES AND JUNIORS



sophomore year to create similar "hothouse" development toward Position 4. A weakness of the artifact hypothesis, however, lies in the implication that the initial rapid gain is an artifact in the sense of an *artificial* inflation. Seminars presumably bring about real, not just artificial, changes in the way students think. If seminars in interdisciplinary programs promote rapid movement from the epistemology of received knowing to that of generative and collaborative learning, this is a substantive gain, not a kind of grade inflation. It cannot be said that in producing the results which register so rapidly and impressively in the Perry model, seminar instruction is just a kind of teaching-to-the-test. Instead, one can read the same results as saying that interdisciplinary studies, seminar participation, and other aspects of the Evergreen environment initiate changes in the freshman year which are so large as to require time for assimilation and elaboration.

The sophomore slump hypothesis is incorrect insofar as it predicts regression, but is compatible with the data in suggesting that social life has its attractions at Evergreen as much as elsewhere, and that when the anxiety of coping with new academic demands is overcome, energy is freed for investment in other areas of life besides cognitive development. Instead of regressing, most sophomores and juniors seem to stay put or to "temporize," which in Perry's terminology (1970) means staying in the same position for more than a year.

Although the artifact and sophomore slump hypotheses both have some merit in explaining the lack of sophomore and junior development, another hypothesis should be formulated to complete the picture. A *consolidation* hypothesis is also compatible with the data, the Perry model, and developmental theory in general. Piaget wrote of rhythms of accommodation and assimilation. ("The filtering or modification of . . . input is called assimilation; the modification of internal schemes to fit reality is called accommodation" [Piaget & Inhelder, 1969, p. 6].) From this perspective, the impressive gain of the freshman year would entail accommodation, or construction of new cognitive structures. These skeletal structures would need time to flesh out, to organize the data of memory and to assimilate new intellectual experience. Thus when Figures 8 and 9 show movement from both sides toward the 3-4 Transition, this can be interpreted as assimilative or consolidative activity, rather than as developmental arrest. And there is no reason to think, on the basis of data from other institutions, that the time Evergreen students spend in Position 3 and especially the 3-4 Transition is inappropriately long. In other words, a real consolidation of gains and preparation for further development is as compatible with the graphs as are the other hypotheses.

Two kinds of development have to go on for a student to work fruitfully in interdisciplinary studies and to move on to more complex cognitive positions. First, disciplinary foundations must be laid within the cognitive structure of Position 3, with its emphasis on methods of problem-solving; and second, the possibility of relativism must enter the cognitive picture, which is what happens during the 3-4 Transition. If these are the cognitive operations on which sophomores and juniors are focusing, then this is developmental work which can legitimately take one or two of the college years.

In freshman programs, students probably learn the ideology more than the effective practice of interdisciplinary studies. They are in favor of such studies and can probably handle cross-disciplinary assignments of the compare-and-contrast type. But only later do they fill in the subordinate competencies, the disciplinary skills and in-

formation, that are requisite for real interdisciplinary thinking. This assimilative process involves what cognitive scientists call "chunking" and "automatization." "Chunking" means learning, through the experience of repeated practice, to organize fairly large units of procedure and content; and "automatization" means rendering cognitive processes habitual and unconscious, so that one doesn't fall off one's cognitive bicycle while trying to remember consciously how to ride it. Assimilation must also involve some grasp of the "tacit dimension" (Polanyi, 1967) of a discipline, the *modus operandi* of those who feel at home in a given knowledge-domain. Interdisciplinary students must navigate such passages from novice to expert (or at least to journey-person) in more than one domain.

In the sophomore and junior years, both coordinated studies programs and group contracts contribute to building and connecting disciplinary foundations; and this work on foundations is the interdisciplinary equivalent of pursuing a major. Coordinated studies programs are interdisciplinary in structure, but they do not fail to teach disciplinary foundations; rather, disciplinary competences and the bridges or connections between them are constructed simultaneously. Evergreen faculty in lower-division coordinated studies programs probably focus on interdisciplinary issues without noticing how deeply their students are preoccupied, according to SSEs, with skill development and disciplinary competence. For example, students often comment in SSEs that "I worked on my writing," or other skills, without presenting much evidence of grasping the interdisciplinary themes.

In group contracts, disciplinary work tends to be intensive. If the contract is interdisciplinary, nevertheless the disciplines at this level of study are usually closely connected (e.g., chemistry and biology in "Molecule to Organism"). Thus within fairly specific knowledge-domains, students are accumulating factual knowledge and basic skills such as would register on tests of level of disciplinary information and of reading, writing, critical reasoning, and computational skills. One would predict that the students in Position 3 and in the 3-4 Transition would show considerable gain in standardized tests of skills (probably not across-the-board, but in areas of academic concentration and role-identification). These gains should appear in Evergreen's assessments of math and writing skills.

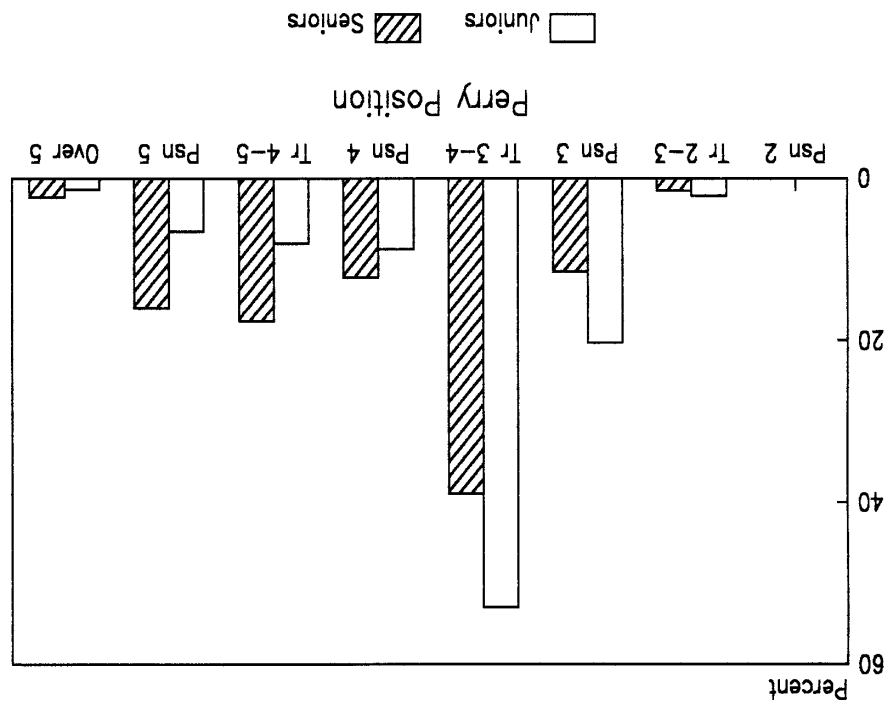
The other kind of work to be done in the sophomore and junior years is forward-looking - the work of the 3-4 Transition, with which the largest number of students are occupied. The task is to apprehend the relativistic implications of multiple

ways of knowing. Nelson (1988a, 1988b) has explained how students at this point tend to form a "mosaic" of ways of knowing, typically with a fault-line between scientific truths and the procedures leading to them, regarded as objectively certain; and personal truths, regarded as subjectively certain. The mosaic is itself an achievement, requiring much consolidative effort in Position 3. But problems are bound to arise in the 3-4 Transition about integrating the separate parts. The different "objective" modes of knowing may prove to be incommensurate or contradictory (psychology and economics professors, for example, may explain the same phenomena in irreconcilable ways). Or one may notice that these modes of knowing are not so objective after all, but are tinged with the subjectivities of culture and ethnicity, of class and gender. The neat tiles of the Position 3 mosaic thus lose their square boundaries and start to bleed into one another.

Perry uses a different metaphor, that of the "Trojan horse," to explain this kind of change, when a student lets an idea through the gates that subsequently conquers a well-defended pattern of "separate," "procedural" knowing. In the context of interdisciplinary studies and team-taught programs, students in Position 3 are necessarily exposed to "Trojan horse" influences which move them toward Position 4 - that is, toward concerns with diversity of opinion, with status of authority, and with independent thought. They encounter differences of perspective and opinion among their peers in seminars, and also among the faculty, not only in lectures but when, almost every quarter, they move to another faculty member's seminar. At first different faculty viewpoints might be perceived as personality differences, so that the student can resolve diversity by picking which faculty members she likes best. But finally there is no avoiding the situation that faculty, students, and texts also disagree substantively, offering opinions which are potentially legitimate but at some level incommensurate or contradictory. The student constantly faces this uncertainty about truth during the 3-4 Transition, no doubt with both attraction and repulsion; and so there is much work, both exploratory and defensive, to be done.

In sum, the artifact and sophomore slump hypotheses, although not without merit, do not fully explain the sophomore and junior years. There is work to be done in consolidating Position 3 and approaching Position 4. The complexity of these tasks is such that a student can quite understandably spend more than a year consolidating skills, encountering relativism, and perhaps cultivating an openness to diversity.

Figure 10:  
JUNIORS AND SENIORS



*Juniors and seniors: A "senior surge"*

The community college and College of Relativism hypotheses are both undone by a glance at the junior and senior class profiles (Figure 10), which together depict the indigenous graduates as upper-division students. From the junior to the senior year, the mode remains in the 3-4 Transition (3.67); but the mean rises by a third of a position (from 3.66 to 4.00), and the distribution rises markedly in the higher positions, with the proportion in Position 4 and above rising to almost half (48%).

The community college hypothesis was in some difficulty from the outset, because the comparative data suggested that in cross-sectional analysis, only about 15% of Normal State students reach the 3-4 Transition, whereas the Evergreen graduates were already pushing 60% in the sophomore and junior years. These are simply not the ratings one would expect from community college students. The graph of Ever-

green graduates as upper division students further undermines the community college hypothesis. Only if Evergreen's critic meant that upper-division work is by definition disciplinary study, without an interdisciplinary matrix, would such results be compatible with his perception. But if he were to admit that upper-division work involves thinking in more complex ways, and if these ways are captured in Perry positions 4 and above, then Evergreen must be offering upper-division work of impressive quality.

The College of Relativism hypothesis is also doomed to failure, because the proportion of students in Position 4 dips after the freshman year and ends up only regaining the freshman level (12%). In the senior year, fewer students occupy Position 4 than the positions beyond it - the 4-5 Transition (18%) and Position 5 (16%). What needs explaining is why the *opposite* of this hypothesis is true: why so few students occupy Position 4, while so many reach higher positions. The problem is to find consistent explanations for a cluster of factors which appear together: (1) the high number and the persistence of cases in the 3-4 Transition, for the first three years at Evergreen; (2) the relatively few cases in Position 4; and (3) the relatively high - even extraordinary - number of cases in Position 5 and above. Is there any coordinated explanation of why there are so many students in the 3-4 Transition, so few in Position 4 itself, and so many beyond?

One explanation might be that the positions are, if sequential, then not equidistant: it might hard to get to Position 4, and then easy to move further toward Position 5. But this explanation does not make good theoretical or empirical sense. First, Perry's theory suggests the opposite, that it would be harder, not easier, to move from Position 4 to Position 5, because the transition is developmentally complex and entails "a revolution in identity" (1970, pp. 114-122). Second, other studies simply do not suggest it is easy to move ahead from Position 4. The data (Appendix B) are scant, because so few students advance beyond Position 4; but of the 337 seniors who took the MID prior to 1984, 4.7% reached Position 4, 1.5% reached the 4-5 Transition, and none reached Position 5 (Moore, 1982/86). This pattern derives from cross-sectional data, but there is little reason to think that because the students in the present study are destined to graduate, this would reverse the pattern.

Another possibility is that the low number of ratings in Position 4 is attributable to use of SSEs and not the MID instrument, which is usually used to generate cognitive development ratings. Preliminary research showed that results from SSEs tend to run a bit higher than results from MIDs, and that the SSEs tend toward a less normal distri-

bution, with a disproportionately high number rated in the 3-4 Transition. This swelling of the 3-4 Transition could occur for two reasons. One is that the SSE stimulus, to present "the student's own evaluation of personal achievement," can itself generate more transitional material, especially in the form of "before" and "after" images. Another is that some students might be shy about presenting a Position 4 evaluation to their faculty member. Perhaps few students write Position 4 SSEs *because this is an evaluation*. From a strict multiplistic stance, a student would not want to evaluate anything at all; all endeavors are of equal value to whomever has undertaken them, and the evaluative task seems at odds with one's whole way of thinking. In particular, a SSE written from Position 4a, the oppositional stance which says nobody can call me wrong and everyone has a right to his or her own opinion, might be hard for a student to discuss with a faculty member or include in a permanent transcript. A reasonable strategy, if the student found the situation awkward, would be to emphasize the skills and procedures she had mastered and to de-emphasize the component of relativism. If this is occurring, part of the swelling of the 3-4 Transition and the dip in Position 4 could be a side effect of using evaluations rather than MID essays as data for cognitive development ratings.

What next needs to be explained is how so many seniors could appear in the 4-5 Transition and Position 5. What happens to seniors that does not happen during other years? Three possibilities (which might operate in conjunction) are first, that some seniors were simply ready to move to a higher position; second, that many seniors have an especially important culminating experience to evaluate during the senior year; and third, that the occasion of writing the last of a long series of evaluations is itself a stimulus to contextually relativistic thought.

Some movement toward Position 5 is to be assumed, without any special explanation; and when a student arrives at Position 5, contextual relativism, relativistic evaluation comes naturally. The student can see her achievements in the context of possibilities for that person in that situation. In fact, it is only in the higher positions that SSEs become evaluations strictly speaking: they stop being statements about likes and dislikes, and lists of achievements; and they become deliberations about the possibilities which the situation contained, and value judgments about how well the opportunity was utilized. A student can, in other words, make better sense than ever before of the task of evaluation itself from the perspective of Position 5.

Second, students may have an especially important culminating experience to

evaluate in the senior year, and they may have been under pressure to think about this year more extensively and in more complex ways because it is so consequential. Evergreen does not require a senior thesis or culminating project, but many students regard their final program, contract, or internship as such a project. If, in the sophomore slump, studies yielded somewhat to social or part-time work agendas, now academics and "life after Evergreen" tend to return to the top of the list. Choosing a senior program is an act of consequence. When one completes one's major at Normal State, senior course choices are not absolutely crucial (an economics major, for example, is probably just rounding out the major, and she will probably not face large-scale choices about who she will be in the world when electing a cluster of senior courses). But if at Evergreen the whole question of how to use the senior year is an open one, if one can change fields or can undertake full-time activities as different as an internship or an interdisciplinary program, then the choice of a senior program and mode of study needs to be reflected upon and rationalized in terms of the kind of person one wants to be. Academic advisers confirm that much concern goes into choice of senior programs, and choice-making and deliberation about consequential choices becomes, in effect, a pedagogy: in making consequential choices as seniors, students are truly learning to think in complex and consequential ways. The high Evergreen senior ratings thus may indicate that students have been taking the senior year especially seriously.

Finally, it is possible that the evaluation process itself is a stimulus for seniors to achieve Position 5. The task of evaluation is itself a spur to cognitive development. The stimulus has always been there: the process of self-evaluation has called regularly for taking an overview and seeing one's own personal achievement in the context of a whole program of studies. Evergreen thus creates a "culture of evaluation" (Thompson, 1989b) in which students learn to reflect upon their individual work in the complex, integrated contexts of seminars and programs. In their content, the evaluations often indicate a pressure to organize and dramatize, to see one's education as a kind of "Pilgrim's Progress" (Perry's metaphor) with continuity from the beginning to the end of each program. The pressure to wrap up the plot is especially strong in the senior year, when the curtain is about to fall. Many graduating seniors write evaluatively from the synoptic perspective of saying what their whole education has added up to, and the search for a synoptic perspective is a stimulus toward Position 5. In this sense, the task of evaluation is itself educative, and it perhaps spurs seniors, looking backward for the meaning of four years' effort, to arrive at Position 5.



These explanations are speculative, and are not necessarily adequate to understanding the senior surge in cognitive development ratings. But the phenomenon is there to see in Figure 10, as a major finding of the present study. To the initial finding of great impact of the College on cognitive development in the freshman year, one can add that many students contribute to another wave of gain in the senior year.

The natural issue for future inquiry would be what happens to transfer students. When the College does not have four years to work its full effects, how much of the impressive gain revealed in this study is available to transfers? And to which transfers, immigrating and emigrating in different years? Do transfer students tend to experience the same gain as freshmen in their first year at Evergreen? Does it depend on being in a seminar? Do they stay as long in transition or consolidation as indigenous sophomores and juniors, or are they under more pressure towards change? And do they too tend to experience more rapid change in the senior year?

### **Gain at the individual level**

The freshman, sophomore, junior, and senior class profiles have served to depict the broad-brush pattern of Evergreen education. But this overview of class profiles may have given the false impression of all students marching incrementally towards higher positions. The view presented so far has said nothing about individual behavior, which is variable. When, in these profiles, the mean or mode is rising, everyone could be rising a little; or alternately, individuals could be passing one another in both directions, with some students regressing, others moving ahead, and the latter only slightly exceeding the former on the average. The next steps are, then, to locate individual gain, and to find out to what variables, demographic and academic, it seems to be related.

### *Defining gain*

In this study, individual gain is determined conservatively by subtracting the freshman rating from the senior rating, even if a previous rating in the sophomore or junior year was higher. The data impose the limitation that the baseline is performance at the end of the freshman year, rather than at matriculation. With gain so defined, it is normally distributed, with a mean of 0.345 and a standard deviation of .81. Fifty-nine percent of the individuals in this study experience some gain, by this definition. Referring to the percentage of all subjects, one can say that half of the graduates gain at

least one-third of a position; 42%, more than one-third position; and a quarter, a full position or more.

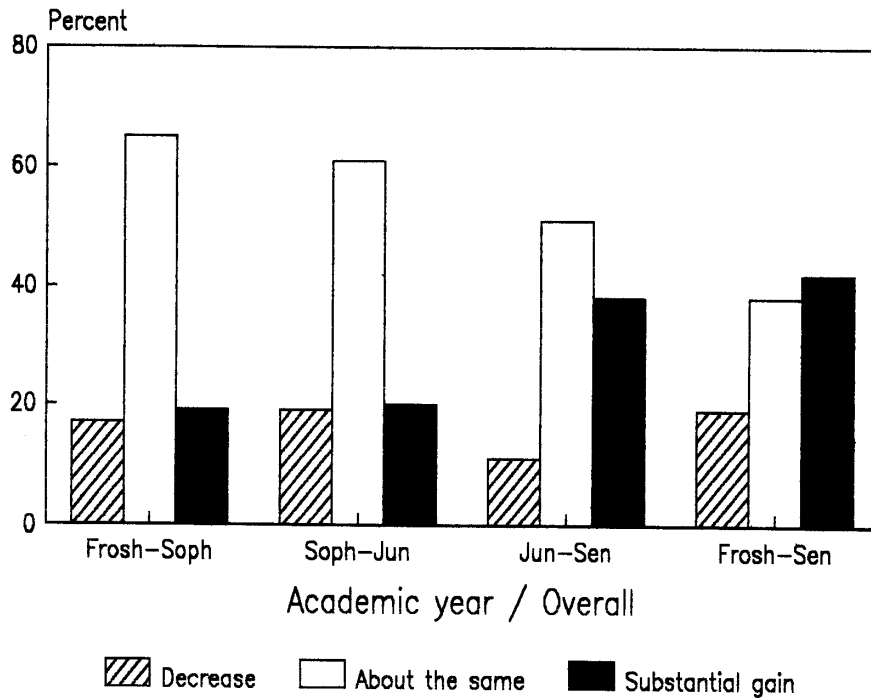
Individual gain could have been defined more liberally, resulting in the impression of more gain. For example, preliminary research comparing MID's to SSEs (see section on "Methods") suggested that upper-division students' ratings on different writing samples tend to fan out, with the author picking from a repertoire of cognitive structures to address the issue at hand. In the absence of a probing interviewer to seek out the subject's highest potential, many SSEs thus might not illustrate the subject's most complex thought. If this were the case, then the highest rating attained, not the chronologically latest rating, would be the best indication of gain. By this more liberal criterion, 75% of the individual subjects in this study would exhibit gain in cognitive development rating, the mean gain would be 0.60, and 25% would have gained more than a full position.

#### *Variability in individual gain*

To determine how much individual movement goes on beneath the surface of the class profiles, individual gain scores can be assigned to three categories: "decline," or negative changes of more than one-third position; "about the same," covering the range of plus or minus one-third of a position; and "substantial gain," or gain of more than one-third position. The criterion for gain is set high in considering this issue because the rating system cannot claim great reliability in measuring small increments (Moore, 1982/86). Whereas such small unreliabilities tend to cancel out in considering other issues, they would confound the present one. With a broad category of "about the same," the picture presented in Figure 11 is affected less by the rating system.

By this criterion, most students stay in about the same position from any one year to the next, but considerable mobility and exchange of positions occurs during each year. Initially, almost as many individuals decline as achieve a substantial gain, so that only in the senior year can one say individuals who gain (38%) greatly exceed those who decline (11%). The bars to the far right represent individuals' cumulative scores from the end of the freshman year through graduation. At graduation, 19% register decline, 38% are within one-third of a position of their starting-point, and the largest group of individuals (42%) has registered gain of more than one-third position.

Figure 11:  
INDIVIDUAL GAIN (BY YEAR & OVERALL)



### *Explaining decline*

Instances of decline - which turn out to be almost 20% of all cases - may come as a surprise if one expected steady incremental gain by everyone. But the Perry model does not presuppose such a pattern. A strict developmentalist might claim that development is not subject to backsliding, and either one has developed or one has not. But the Perry model does not insist that development is always upward and linear. The model accounts, first of all, for a student being ahead in one area and still behind in another (as in Piagetian *decalage*). Secondly, it suggests that when a student encounters a new field of endeavor, she tends to recapitulate or recycle through the developmental stages (Perry, 1981) - presumably more rapidly, if one is working from a higher position, but not instantaneously. Third, the model is not measured perfectly by the MID instrument (Moore, 1982/86) or by the present adaptation of it; in all measure-

ments some variance is due to unreliability. And finally, Perry (1970) mapped three paths leading to stasis or actual decline, three routes which students might take to circumvent greater cognitive complexity: (1) retreat, or active denial of the legitimacy of multiple versions of truth; (2) temporizing, or delay as a defense against a new accommodation coming too soon; and (3) escape, by which Perry meant temporizing permanently and refusing the stimuli toward growth.

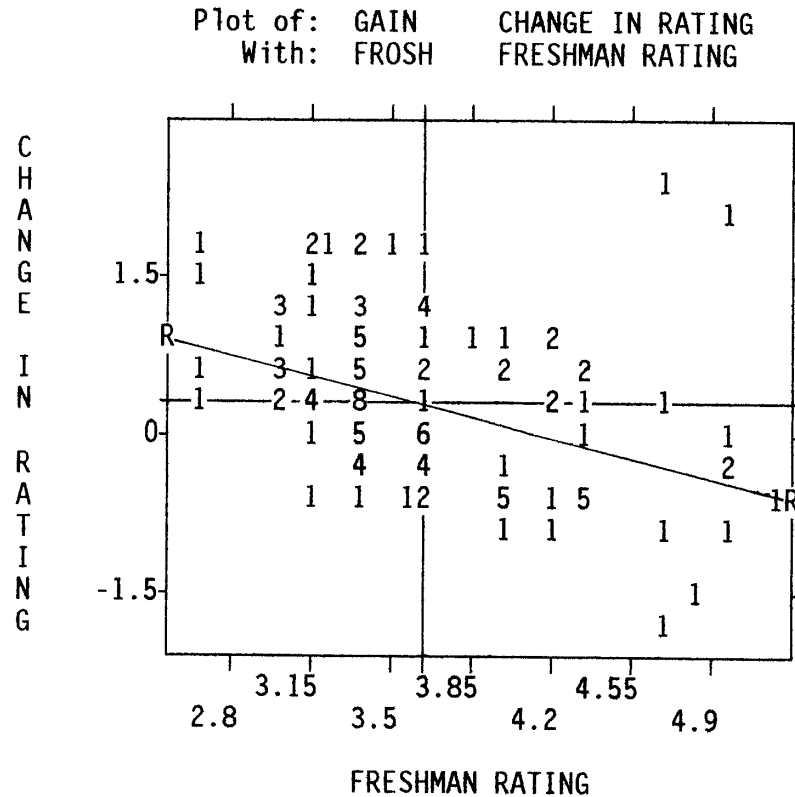
In other college contexts, retreat might involve return to an anti-intellectual religious orthodoxy, but this is unlikely at Evergreen. Individuals who decline may be exercising two other options (or a combination): if an Evergreen student wants to avoid cognitive complexity but does not wish to drop out of college (and hence out of this study), then the main alternatives would be to minimize academic engagement, or to adhere to a simplifying ideology or "-ism" (what Perry researchers call a "New Truth"), particularly one which identifies the bad guys and explains virtually everything without invoking complex cognitive operations.

#### *Tortoises and hares*

One of the most suggestive but problematic outcomes of this study is that individual gain seems to follow a "tortoise-and-hare" pattern, with slower starters tending to end up as higher gainers. On the average, students who started with higher cognitive development ratings as freshmen end up with higher ratings as seniors, but they tend to have gained only moderately; whereas students who started with lower cognitive development ratings tended to accrue more gain by the time of graduation, by a significant margin. Such a pattern might suggest that the College is doing a fine job of stimulating its initially less advanced students, but that it might not be as supportive of the cognitive development of those initially more advanced, who might even tend to coast after the freshman year without being sufficiently stimulated in the direction of more complex cognitive operations. But such a pattern might also be a statistical artifact, a regression to the mean, in which especially high initial scores are typically followed by lower ones, and vice versa. This statistical phenomenon no doubt contributes to the tortoise-and-hare pattern, but it does not completely eliminate its significance. And overall, the data definitely show that the *opposite* pattern does not exist - that Evergreen education is not patterned like Reaganomics, in which the rich get richer and leave the initially disadvantaged even further behind.

For readers willing to consider cognitive development ratings as more than or-

Figure 12:  
GAIN AND FRESHMAN RATING



118 cases plotted. Regression statistics of GAIN on FROSH:  
 $r = -.40252$ ;  $r^2 = .16202$ ; S.E. of Est., .74712;  $p = < .0005$ .

dinal data, the tortoise-and-hare pattern stands out in highest relief as a correlation between freshman rating and gain (Figure 12). Treated in this manner, freshman rating turned out to be the second best predictor of individual gain, second only to senior rating (and more accurate than sophomore or junior rating.) Senior rating explained 56% of gain, freshman rating 16%. But the correlation between gain and freshman rating was negative ( $r = -.40$ ;  $p < .0005$ ). Thus as in Aesop, so at Evergreen: whereas the slower-starting tortoises who end the freshman year below the mean tended to register gain when they graduate, the hares who start out above the mean tended not to gain in rating. Figure 12 shows this pattern as a plot of individual cases. The reference lines dividing the quadrants represent mean freshman rating and mean gain. The upper left quadrant is full of students who began with below-average ratings, but reg-

Table 1:  
GAIN AND FRESHMAN RATING

Crosstabulation: GAIN CHANGE IN RATING  
By FROSH FRESHMAN RATING

FROSH->  GAIN	Count			
	Exp Val	Below	Above	Row
	Col Pct	median	median	Total
No gain	16 27.0 27.6%	39 28.0 65.0%	55 46.6%	
Gain	42 31.0 72.4%	21 32.0 35.0%	63 53.4%	
Column Total	58 49.2%	60 50.8%	118 100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5	Eta
15.11965	1	.0001	27.034	None	.37495

istered higher than average gain; and the lower right quadrant is as crowded with students who began with higher ratings but ended with below-average gain. The regression line (from R to R) depicts the overall tortoise-and-hare trend as a moderate downward slope (negative correlation).

For statistically more rigorous readers who regard cognitive development ratings as ordinal data and thus not subject to linear correlations, the tortoise-and-hare pattern appears in Table 1 as a crosstabulation and chi-square test. Of the tortoises who rated as freshmen below the median (3.33), 28% posted little or no gain, and 72% gained one-third of a position or more; whereas of the hares who initially rated above the median, 65% posted little or no gain, and only 35% gained a third of a position or more ( $p = .0001$ ). The eta coefficient, squared, indicates that freshman rating accounted for 14% of variation in gain. The statistically conservative reader might balk at further dividing ordinal data at any point but the median; but if one uses three categories in order to relate freshman rating to (1) decline in rating at graduation, (2) same rating, and (3) higher rating, then the crosstabulation shows that of the tortoises, only

7% register decline in rating at graduation, whereas 45% of the hares actually decline ( $p < .0005$ ). If category (3), higher rating, is further divided into "moderately higher" (one-third through two-thirds of a position) and "much higher" (a full position or more), then the hares' count falls below the expected value about equally in both categories.

This tortoise-and-hare pattern seems to allow a prediction: whereas most findings offered in this study are simply descriptive statistics about a particular population, this pattern is statistically significant for purposes of inference, so that considering these subjects as a sample of past and future indigenous graduates, one can predict that the tortoises will gain more than the hares until something causes change in the student body or the curriculum.

But is this pattern real, or only a statistical artifact called a regression phenomenon? If real, then the tortoise-and-hare pattern expresses actual differences in the behavior patterns of initially high-scoring and low-scoring students, with the former somehow becoming less motivated or less effective in cognitive development, and the latter more motivated or effective. But if the pattern is a statistical artifact, then it is an example of a regression effect in which substantial irregularities appear when data are gathered at just two points in time, but then tend to level out, or to follow the rule of "regression to the mean," when more measurements are made.<sup>3</sup> There is no conclusive test to determine whether or not the tortoise-and-hare pattern is a regression effect, but one can look further for potentially corroborating evidence.

One corroborative strategy is to see if the low-scoring (tortoise) freshmen actually end up above the senior mean of 4.00, and the high-scoring (hare) freshmen below it. Such results would definitely exceed those explicable as a regression phenomenon. But the data do not take so decisive a turn. The freshman tortoises start with a mean rating of 3.18 and end up with a senior mean 3.83, and the freshman hares start at a mean of 4.07 and reach a senior mean of 4.16. If the "pull" between the initial rating and of the senior mean were simply equal, then the freshman tortoises would have reached 3.59, and the seniors 4.04. So the data fall in the gray zone, where both explanations, real results and regression effect, are plausible contributing factors.

---

<sup>3</sup>Such convergences occur, for example, when unusually high and low batting averages, appearing early in the baseball season, tend to level out as more cases accrue through time. When initially high batting averages settle down and lower ones rise, one cannot conclude that the initially hot batters lost their spirit and the others found it - only that both groups are "regressing to the mean."

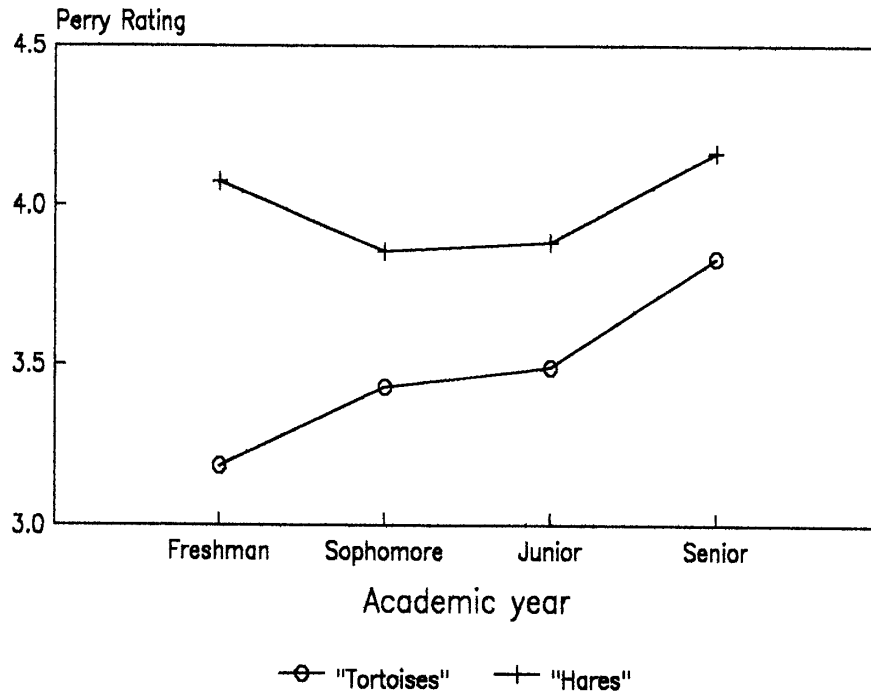
A second mode of corroboration is to use discriminant function analysis to see where freshman rating fits in among variables that help to predict gain. Such analysis indicated how variables might be combined to predict the three categories of no gain, moderate gain, and high gain. This exploration was more suggestive than conclusive, because the number of complete cases was small ( $n = 59$ ) and the proportion correctly classified was moderate (71%). The variables included in the discriminant function were freshman rating, ethnicity, high school GPA, age at graduation, units in the study mode of courses, and SAT. (In this kind of analysis, ratings in the other academic years and in sex, parents' education, and the other modes of study did not help to predict location of individuals in the three categories of gain.) Freshman rating emerged as most important among these variables: any student with a low initial rating was predicted by discriminant function analysis to post gain. Individuals who might predictably decrease tended to include those usually thought to be well-prepared for college, by the traditional criteria of high school GPA and SAT. In other words, they disproportionately included the achievers who had done best in high school by conventional measures of success. This again suggests a real tortoise-and-hare phenomenon and not just an artifact; but the number of cases and moderate proportion correctly classified are still inconclusive.

A third mode of corroboration is to ask whether one sees a real pattern or merely a regression effect when introducing more points in time. Figure 13 shows the mean Perry position from the freshman through the sophomore, junior, and senior years for both groups - the tortoises in the upper line, and the hares in the lower. In each year, each group was significantly different from the other (including the senior year, when they tend to converge [ $p < .02$ ]). The initially larger difference grew smaller through time. One indeed sees what can be regarded as a regression effect, but this is not all. Regression to the mean by itself would take the form of two relatively flat lines converging on one point. What one sees in addition is a definite "sophomore slump" for the hares, in contrast to the steady gain of the tortoises. In considering the sophomore and junior years, a question had arisen earlier as to whether the class profiles indicated a slump, or a consolidation. The answer is now clearer: a sophomore slump for the hares was accompanied by sophomore gain for the tortoises, followed by sophomore-to-junior consolidation for both groups. Both groups then went on to participate in the "senior surge."

In all years, hares displayed more variance than tortoises (the hares' standard



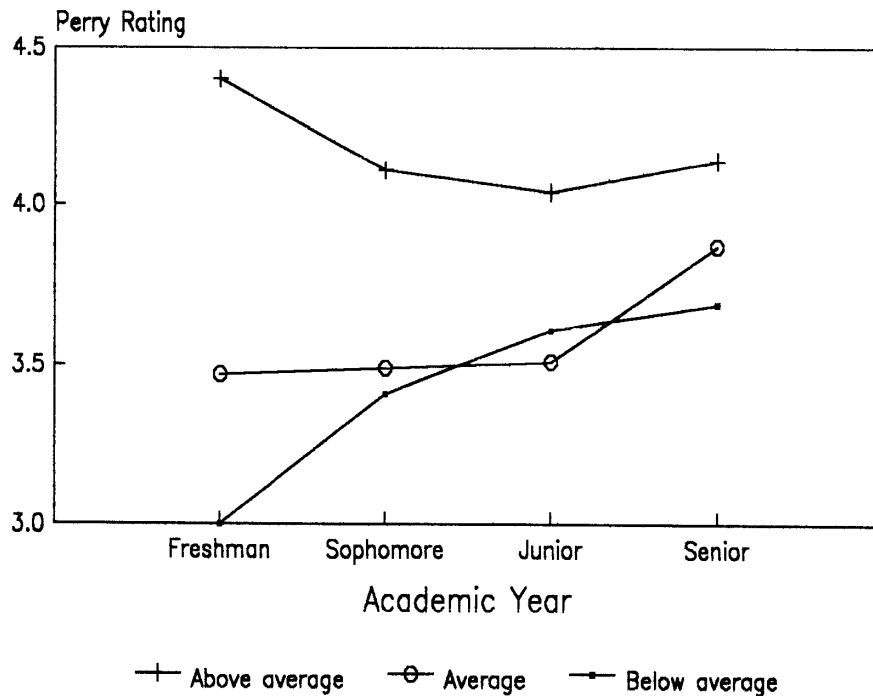
Figure 13:  
"TORTOISE AND HARE" MEAN RATINGS BY YEAR



deviations were freshman, 0.47; sophomore, 0.57; junior, 0.66; senior, 0.88; whereas the tortoises' sequence was 0.21, 0.49, 0.50, and 0.62). This could be an instance of the general tendency of higher ratings to vary more, as mentioned in the "Methods" section. But it also suggests that more than one subpattern might be contained in the hares' ratings. The frequencies suggest that while the tortoises moved in subsequent years toward higher positions in a pattern of normal distribution, the hares tended either to stay in the 3-4 Transition, or to move on to Position 5. (The low number of students ending up in Position 4, noted earlier, is thus a low number primarily of hares in Position 4).

Figure 14 gives a clearer idea of what is happening by presenting three rather than two subpatterns. Here, the freshman ratings are categorized three ways, as "below average" (including ratings through Position 3), "average" (spanning the 3-4

Figure 14:  
THREE GROUPINGS: MEAN RATINGS BY YEAR



Transition), and "above average" (beginning with Position 4). This depiction again shows convergence which can be regarded as regression to the mean, but it also reveals other trends. The initially below-average students (i.e., those of them who remained in the College) made steady gains, the highest overall. The students with initially average ratings tended to remain for two more years in the 3-4 Transition, perhaps consolidating freshman gains (which do not appear here, in the absence of a freshman pretest); and they then contributed heavily to the senior surge. It was predominantly *their* surge. Finally, the students with the initially highest ratings truly slumped, not just as sophomores but also as juniors, and then participated moderately in the senior surge. This is the point that matters: in the three-group split, the students who started out highest stand out as having tended not simply to gain less, but actually to have declined. In the two-group tortoise-and-hare illustration, split at the median (Figure 13), the hares

caught up, as seniors, with their own freshman ratings, and were able to post overall gain. But Figure 14 shows that their gain depended heavily on students just above average, now assigned to the middle group, who spurred the senior surge. The group with the highest freshman ratings actually declined from the 4-5 Transition and a mean rating of 4.40 as freshmen, to Position 4 and a mean rating of 4.14 in the senior year.

The outcome of these considerations is that some of the tortoise-and-hare pattern is no doubt a regression phenomenon, but definitely not all. And finally, one can conclude from these several kinds of evidence that the situation is definitely not the *opposite* of the tortoise-and-hare pattern. Without being able to sort out exact amounts of the "real" phenomenon and the regression effect, one can definitely say this is not the Reaganomics pattern in which the rich get richer while the poor get poorer. The initially cognitively more advanced students definitely do not account for most of the gain. While it is true that Evergreen's tortoises, as a group, never attain the mean position of the hares, nevertheless the gap definitely narrows, suggesting the initially less advantaged students perhaps profit even more than others from Evergreen's particular kind of liberal arts and sciences education.

When a finding like this emerges (even with some uncertainty, as in the present instance), it has to be interpreted in connection with teaching and learning experience. Who might be the tortoises in the student body, and who the hares? This question seemed puzzling when put informally to members of the faculty, and the categories did not seem to resonate clearly with local myths and experiences, except with the general observation that intellectually astute and well-prepared students probably have to gear down their level of seminar discussion to stay in touch with the discourse of their peers.

While faculty did not identify the tortoise-and-hare pattern as a familiar phenomenon, this is no surprise, since they tend to be in contact with students for only a year or two and are unlikely to form an overview of the individual student's progress from matriculation to graduation. The tortoise-and-hare pattern did, however, make experiential sense to another researcher, now the Vice President for Student Affairs, who read student transcripts from the college's first decade and developed a Student Typology of Academic Growth (Martin, 1982, pp. 90-95). Her experience both as an adviser and as a research-reader of transcripts indicated that the tortoise-and-hare pattern is true to life. Her research revealed that some students "throw themselves into the pool" of Evergreen experience without reservation when they arrive (personal communication), while other subtypes are "withholders" and "deferrers" who are initially

somewhat skeptical of the College's novelties and immerse themselves slowly, as if saying "prove it to me." The general pattern of student development which she found in Evergreen narrative transcripts was one of "big jumps," followed by rather long periods of slower-moving "exploratory academic behavior," followed perhaps by another "quantum leap" (Martin, personal communication). Those students who leap into the pool but then wait a long time before experiencing further development might number among the hares of the present analysis, whereas the more conservative skeptics with prove-it-to-me attitudes, the withholders and especially the deferrers, might exemplify the tortoises.

Assuming that tortoises and hares are real students and not mythical beasts, what sort of explanations or hypotheses would account for their evolution? If students who start with lower ratings gain in ability to handle more complex intellectual problems and graduate successfully, surely this is a commendable educational outcome; but why might the students who start from higher positions gain less? It is the hares who are problematic. No one explanation of their pattern seems obvious, and perhaps a combination of factors affect each case, or different factors explain different cases. One explanation might be that the baseline data for this study were the *end-of-freshman-year* SSEs. It is possible that the hares, the students who start with higher ratings, were first glimpsed right after an advance in cognitive position within the freshman year, so that they were in an assimilative stage and were less likely to advance again. This would be compatible with the Vice President for Student Affairs' view of students who might be the hares of this study.

Another explanation has to do with self-evaluations as the data for ratings. Perhaps performance anxiety initially stimulates complex thought in writing self-evaluations, but later, when the task becomes more familiar, it is approached with less serious reflection. Freshmen (and other new students) often express anxiety about writing their first formal evaluation and going to discuss it with their faculty member. Perhaps a kind of Hawthorne effect generates heightened concentration as students formally observe and evaluate themselves for the first time. And perhaps evaluation later becomes a less stimulating kind of business-as-usual, a routine mode of production. The hares could in some instances exemplify this pattern.

Further possible explanations of the tortoise-and-hare pattern raise philosophical issues about the empowering and liberatory, or the conservative and socializing, functions of higher education. A third explanation might be that some freshman thinking

has a high-flying quality that tends to be rated well up on the Perry scale, but that this sort of thinking is subsequently abandoned by the same minds as they become more seasoned. Some freshman SSEs, enthusiastic in tone, suggest that the novice evaluator's thought is initially not dualistic but uncritically holistic-utopian. These late adolescents may be pursuing relatively naive vision-quests and might be settling initially for spurious connections and totalities, so that subsequent decline in Perry position might represent wiser or at least procedurally more logical thought, more respectful of the problem-solving potentials of Position 3. From this standpoint, one function of higher education would be to connect utopia with reality, to make it less like a cloud and more like a tree, rooted in skills and methods of realization. As these skills are acquired, the self-evaluations that describe the process may, especially in the absence of a probing interviewer, emphasize achievements at Position 3.

The previous explanations all find good or respectable reasons for the hares' lesser gain. But alternately, it might be regarded forthrightly as a sign of something wrong. A long line of educational criticism (e.g., Shor, 1987) suggests that decline in variety and vitality of cognitive functioning might in some instances be a real outcome of higher education, which perhaps systematically replaces vision and critical capacity with the kinds of routinized cognitive behaviors that are useful for functioning productively but numbly in technological society. More students really did rate above the 3-4 Transition in the freshman than in the sophomore and junior years, and the instances of decline may offer a revealing glimpse of a process of bureaucratic and scientific rationalization which emphasizes the pragmatic positivism of Position 3 problem-solving at the expense of broader cognitive development. From this perspective, higher education is socializing modern white-collar workers effectively, but with an anti-intellectual and dehumanizing effect. If this is so, then all one can say is that Evergreen is not quite as guilty as most institutions, because it later returns a relatively high proportion of students - if not the same ones - to more complex cognitive functioning.

No one of these explanations seems satisfactory in itself, but at least an unexpected trend has been identified, so that it can receive attention in other assessment efforts. Evergreen education definitely does not follow the Reaganomics pattern; but how real and how important the opposite tortoise-and-hare pattern might be depends on further research and discussion. This study at least advances the hypothesis that the College and its curriculum are more helpful to those students who begin in the 2-3 Transition and Position 3, and that they are not quite as helpful to those who are per-

forming more complex cognitive operations by the end of the freshman year. If the goal is to stimulate all groups equally, then the limited evidence presented here suggests that this might not be occurring. The question is perhaps even whether seminars and the curriculum are offering the appropriate level of stimulation to those students who are best able to understand the content and underlying principles of interdisciplinary studies.

One naturally wonders, however, about comparable patterns at other institutions. One cannot put Evergreen in perspective without comparative research. Perhaps the tortoise-and-hare pattern is normal and occurs in other colleges and universities. Certainly every teacher of freshmen has encountered students who seemed impressively bright, but then just did not seem to have travelled as far, by the time of graduation, as others who initially seemed more confused or plodding. Perhaps Evergreen's pattern will be replicated elsewhere in comparative research; or perhaps such research will reveal the opposite, that Evergreen is unusual and that other institutions practice "cognitive Reaganomics."

### **Gain by demographic and academic groupings**

Variation in gain has another dimension besides the overall movement of classes and the particular behavior of individuals. It may also vary with membership or categorization in relevant demographic and academic groupings. Demographic groupings represent social characteristics such as age, sex, ethnicity, class, and status. The academic variables represent high school background and the different Evergreen modes of study. In this discussion, however, one must keep in mind that all references to such variables - to women, older students, persons of color, etc. - are to *the subjects in this study*, the indigenous graduates, and not to all students in the College who might belong to the category. As a more general cautionary note, this is not a study of a sample which looks for relationships and differences of such magnitude that one can make inferences or predictions about a larger population. The subjects themselves constitute a population, not a sample, and so the tendencies and differences reported here definitely existed, and are not subject to errors of statistical inference. But in saying that a tendency or difference existed in this population, one cannot immediately conclude either that it will persist among future indigenous graduates, or that it is of such magnitude as to have implications for policy. In general, the following results are informative about how Evergreen works, and they suggest that the College works more or less as it is

supposed to. The small differences in mean or modal cognitive development rating for different subgroups of students are not of such magnitude as to imply that the groups with slightly lower ratings are ill-served, or that problems exist to be remedied by changing policy.

The relationship between demographic and academic variables and Perry positions cannot be expressed as one summary statistic. One important figure is the proportion of each subgroup who gain in Perry position: in the area of gender, for example, do as high a percentage of women as men experience gain (defined as one-third position or more)? This is a revealing statistic but only a start, because all amounts of gain are not equal. Hypothetically, 65% of younger students might gain by a third of a position each, while 55% of older students might gain by a whole position; so one should also ask what proportion of the members of any group experience moderate gain (defined as above one-third position, but less than a whole position) and what proportion experience high gain (a full position or more). Further insights into the demographic distribution of gain come from linear correlations (interval variables), and from changes in mean and modal position (nominal and ordinal variables).

And finally, one can look to see if differences between subgroups increase or decrease when the freshman year is compared to the senior year. The favorable hypothesis is that differences decrease. In most social science research, the goal is to find independent variables which predict a dependent variable. But the present kind of educational outcomes research hopes to demonstrate the opposite, that demographic variables do *not* distinguish among graduates, who should have had equal access to the educational opportunities afforded by the institution. If education works fairly, then all groups should gain, but differences should be ameliorated through time, and demographic variables should become less relevant and determining. Gain should end up about the same for both groups on any dichotomous variable (e.g., male-female), or for reasonable categorizations of interval variables (e.g., for age: students direct from high school, young adults, and older students). On the present evidence, this pattern of equalization seems to occur at Evergreen. The evidence points in the right direction, but it should be recalled (from the Methods section) that the measures used here are perhaps less reliable for seniors than for freshmen.

*Demographic variables*

*Gender*

Gender is a particularly interesting demographic variable because of theoretical considerations introduced by Belenky et al. in *Women's ways of knowing* (1986). The authors suggest that an environment which helps students to find a "voice" and which promotes "connected knowing" will be particularly favorable for women's cognitive development (ch. 6, 10). In many ways Evergreen seems to promote voice and connected knowing: it emphasizes seminars, where one speaks out and learns from one's peers; it assigns faculty the role, in coordinated studies, of colearners; it provides close student-faculty contact in small seminar groupings within full-time, often year-long programs, where students and faculty become well acquainted (the academic program is often the student's primary group at Evergreen, in contrast to the fraternity or sorority elsewhere). The fit between these practices and Belenky et al.'s prescriptions seems fairly close (Thompson, 1989b). And the MID rating system seems to respect connected knowing, because it registers when students learn from their peers and identify with faculty as colearners (rather than as external Authorities). Indigenous graduates, since they are those with the most experience of Evergreen's teaching modes, should be the ones in whom the greatest effects of connected knowing would be apparent. One might thus predict that the ratings of SSEs will reveal relatively high gains in Perry position for women.

But this study does not find the predicted result. One limitation is that the baseline was the rating of the SSE written at the end of the freshman year. At that time, women were on the average slightly ahead of men (women's mean = 3.66, mode = 3.33; men's mean = 3.57, mode = 3.33). Women may have matriculated with a head start in cognitive development, but it is equally possible that they had discovered their voices and had profited from connected knowing already in freshman seminars, with the predicted change occurring before the first data were recorded. When they were later very slightly less prone to gain, this could be part of the general tortoise-and-hare pattern.

After the freshman year, males were favored with just a little more gain. On the average, both sexes gained about one-third position, but men tended to gain a little more (.43), women a little less (.275). In overall terms, only a few more men than women experienced gain (55% and 52%, respectively, in Table 2), but the men were



Table 2:  
GAIN AND GENDER

Crosstabulation: GAIN CHANGE IN RATING  
By GENDER GENDER

GENDER-> GAIN	Count Exp Val Col Pct	Female	Male	Row Total
No gain	31 30.3 47.7%	24 24.7 45.3%	55 46.6%	
Gain	34 34.7 52.3%	29 28.3 54.7%	63 53.4%	
Column Total		65 55.1%	53 44.9%	118 100.0%

marginally ahead in high gain (26% for men and 21.5% for women). Controlling for age indicates that more high-school direct males tended towards gain than high-school direct females, and that older women were less prone than older men to gain a position or more. At graduation, men's cognitive development ratings ended up ahead of women's by the same slight margin as in the CSID comparative data (Appendix B), indicating that Evergreen had accomplished nothing dramatic that favors women (men's mean = 4.02, mode = 3.67; women's mean = 3.98, mode = 3.67). But in terms of amelioration of differences, men and women were more, not less, alike in cognitive development rating at graduation than as freshmen. Gender seems to be a social determinant of decreasing relevance.

One must conclude, however, that indigenous female graduates did not gain more than men, as might have been predicted from the opportunities at Evergreen for finding a voice and experiencing connected knowing. Why did not fact, in this instance, conform to theory? The reason would not seem to be that the rating system is less sensitive across-the-board to women's voices, because it did register their higher average as freshmen. This bit of evidence moves against the possibility (which has been of some concern to Perry researchers [Thompson, 1988]) that the rating system might not be sufficiently sensitive to women's achievements. The remaining possibili-

Table 3:  
GAIN AND ETHNICITY

Crosstabulation: GAIN By ETHNIC CHANGE IN RATING ETHNICITY

ETHNIC->	Count	White	Of Color	Row Total
GAIN				
No gain		50	5	55 46.6%
Gain		60	3	63 53.4%
Column Total		110 93.2%	8 6.8%	118 100.0%

ties are that connected knowing did not occur, or that it did occur but did not promote a male-female difference. Offsetting factors might have favored males. But the most interesting hypothesis for further investigation would be that connected knowing did occur, and proved to be just as advantageous for men as for women.

### *Ethnicity*

The variable which permits the least interpretation is ethnicity, because the number of persons of color in the population was only 12 ( $n$  = only 10 for freshman and senior ratings, and 8 for gain); and these were members of three different cultural groups which might differ as much from one another as from the white category. In Table 3, it is not sensible to offer percentages of students experiencing gain, because with so few students of color, a difference of just one case would change the percentage appreciably.

The ratings for these few students of color displayed more variance than for other students, with more ratings appearing at both the low and the high ends of the continuum of gain (though one would expect, with a higher number, regression toward the mean). In these few instances, whereas a lower number of students of color may experience gain, nevertheless the instances of high gain were offsetting, so that students

of color rate higher, on the average, than other students, both as freshmen and as seniors. The freshman difference was apparent in both the mean and the modal positions (students of color, mean = 3.75, mode = 3.67; others, mean = 3.60, mode = 3.33). The groups moved a little closer together as seniors, possibly indicating amelioration of socioeconomic differences. At a recent conference, Perry researchers expressed concern about possible bias in the rating system (Thompson, 1988), if the more familiar modes of expression of the dominant culture were to receive higher ratings. But the rating system was definitely not so biased as to obscure the particularly high ratings of a few students of color. In general, these few clues suggest variability, but not overall inequality, in the cognitive development of students of different ethnicities. The clues are so weak as to support the need for more careful study of the experience of students of color at Evergreen (in progress, by another research team).

### *Age*

No linear correlation appeared between gain and age at graduation, but other relationships emerged. Overall, differences in cognitive development by age group did not tend to equalize at graduation, but tended to persist. Young high-school direct or traditional age students, those who graduate by age 23, predominated among the indigenous graduates (61%). Older indigenous graduates can be subdivided into young adults in their 20s (25%), and those 30 and older (14%). Here in particular, it is important to remember that these students were not typical of the whole student body, especially because older returning students who had started college elsewhere were excluded from the indigenous population by definition.

The indigenous high-school direct students were 56% female, but the older indigenous graduates were not, as one might guess, primarily women who had delayed their education. Rather, 57% of the young adults (ages 24-29) were male, and so were 70% of the indigenous graduates over 30.

By age grouping, the proportion of students demonstrating gain was not quite equal. As Table 4 indicates, fewer older students posted gain (46%, in contrast to 57% for younger, high-school direct students, those who graduated by age 23); and the gain which they did experience tended to be moderate rather than high (15% high gain for older, in contrast to 29% high gain for younger). Also, older students tended to start from a slightly lower position (3.54, in contrast to 3.66); so their tendency toward lower gain cannot be explained as an instance of the tortoise-and-hare pattern. As se-

Table 4:  
GAIN AND AGE

Crosstabulation: GAIN CHANGE IN RATING  
By AGE AGE

AGE->	Count Exp Val Col Pct	Hischool direct	Older students	Row Total
GAIN				
No gain	33 35.9 42.9%	22 19.1 53.7%	55 46.6%	
Gain	44 41.1 57.1%	19 21.9 46.3%	63 53.4%	
Column Total	77 65.3%	41 34.7%	118 100.0%	

niors, their modal position was 3.33 (mean = 3.93), whereas younger students rose to a mode of 3.67 (mean = 4.04). Further breakdown into three age groups, with the age 24-29 young adults and the over-30 groups considered separately, indicates that the young adults fell disproportionately into the category of "hares" who rated high at the end of the freshman year but in the long run did not register gain. The differences between groups did not decrease from the freshman to the senior years but remained the same, with the implication that age differences are not ameliorated through time. This pattern differs from findings at Alverno College about older students, who tended both to start in higher positions and to graduate in higher positions (Mentkowski & Strait, 1983, Fig. 15; Mentkowski, 1988, Fig. 1). The differences found at Evergreen were not so large as to suggest that older indigenous students are ill-served or are not treated equally, but they might suggest that some older students are less inclined to cognitive flexibility. Older students may tend to have more stable cognitive structures in place than younger students, as a result of greater life experience. And understandably, some older students are more economically realistic and more vocationally oriented, coming to college with the intent of grasping the how-to-do-it, problem-solving structures of Position 3.

*Parents' education*

Parents' education merits consideration as a demographic rather than an academic variable, first, because it is the nearest thing to a measure of economic class and social status available from the College's records; and second, because it is taken into account in admissions policies which aim at equal access to education and at transforming the College into a representatively multicultural environment. The assumptions are that students who belong to the first generation of their family to attend college tend to come from families lower in class and status, and tend more often to be members of minorities. They tend, in other words, to be less socioeconomically advantaged. First-generation college students comprised 29% of the population, and students whose parents (one or both) had attended some college comprised 71%. Controlling for other variables indicates that first-generation and college-parent students did not differ in gender, but that first-generation students tended to be a little older at graduation, pointing perhaps toward different financial burdens.

Graduates who represent the first generation of their family to attend college tended to have arrived at Evergreen a little better prepared, by the criterion of high-school GPA; and they tended also to receive higher ratings in the freshman year. One might instead expect the children of more educated parents to have grown up in high-verbal environments, to have done better in high school, and thus to rate higher in cognitive development, at least initially. But 57% of the first-generation students were above the median high school GPA (3.07), in contrast to 47.5% of the children of college-attending parents. And children of college parents did not get off to a more rapid start when one uses the end-of-freshman-year Perry position as a baseline. Instead, the first-generation students rated slightly higher as freshmen (first generation mean = 3.68, college parents mean = 3.59; both modes = 3.33). These differences could follow from differential selection to attend Evergreen: are the more successful among less advantaged high school students the ones who dare to go to an alternative college? Or perhaps the differences reflect differential retention: did first-generation students who were less prepared or did not do as well as freshmen simply leave the College in discouragement, so that they could not appear in a study of indigenous graduates? Or were first-generation students under more pressure to succeed as freshmen, and therefore more motivated in concentrating on their freshman evaluations? Such unanswered questions point toward the value of further study, charting the different patterns of use of the College - arriving from different backgrounds at different ages, and transferring

Table 5:  
GAIN AND PARENTAL EDUCATION

Crosstabulation: GAIN CHANGE IN RATING  
By PARED PARENTAL EDUCATION

PARED->	Count Exp Val Col Pct	First ge neration	College parents	Row Total
GAIN				
No gain	19 16.3 54.3%	36 38.7 43.4%	55 46.6%	
Gain	16 18.7 45.7%	47 44.3 56.6%	63 53.4%	
Column Total	35 29.7%	83 70.3%	118 100.0%	

in and out in different years.

After the freshman year, first-generation students tended to be hares who were passed by the tortoises. As Table 5 indicates, a smaller proportion of first-generation students experienced gain (46%, in contrast to 57% of those with college parents). And by the senior year, the students of college parents have overtaken by a small margin. Their mean gain was 0.40, compared to the first generation's 0.22. The result is that graduates with college parents ended up at a mean rating of 4.01, and first-generation graduates at 3.97. The difference was very small, but the ratings converged over time, suggesting equalization of differences.

Overall, gains in cognitive development at Evergreen did not seem to be parceled out among the indigenous graduates in proportion to predetermining factors of gender, ethnicity, or age, or by class and status as reflected in parents' educational history. Persons of color, members of both sexes, younger and older students, and graduates from relatively more and less educated families all ended up roughly equal in position at graduation and in gain. All between-groups differences declined by the senior year, except for age differences, which simply stayed the same. Small differences remained, some of which merit further study; but none were of the magnitude that would suggest revisions of curriculum or policy.

*Academic variables*

Of the academic variables in this study, two pertain to preparation for college, which is traditionally inferred from high school grade-point average and Scholastic Aptitude Test score (or comparable Washington State test score, converted by a standard table). The question is whether these traditional predictors of achievement point to gain in cognitive development. The other academic variables are the different study modes at Evergreen: coordinated studies, group contracts, individual contracts, and courses. The question here is whether the proportion of a student's work in any particular study mode bears a relationship to gain in Perry position.

*SAT and High School GPA*

High school GPA and SAT correlated moderately with one another ( $r = .38$ ,  $p = .001$ ), so that they comprise a general factor (which emerged in factor analysis) of preparedness for college. But by itself, SAT can be eliminated as a predictor of freshman Perry position, senior position, or gain. Cases were relatively few ( $n = 89$ ), because reporting of SAT scores to Evergreen was optional. These scores were not related to any other dimension of this study except GPA. The chi-square statistic for higher versus lower SAT was a perfect zero, and its relationships with freshman position, senior position, and gain were small and uninteresting. One cannot consider SAT a relevant variable, let alone an important admissions criterion, if one's interest is in enhancing cognitive development.

In contrast, high school grade-point average correlated significantly, though at a low level, with freshman Perry position ( $r = .23$ ,  $p = .02$ ). Treated nonparametrically, the Mann-Whitney 1-tailed  $p = .04$ . When GPAs were split at the median (3.07, or almost exactly a "B") into lower and higher, then a stronger correlation of .40 ( $p = .004$ ) appeared for the freshmen in the higher category. In other words, high-school high-achievers tended to be more predictable and to be slightly more advanced in cognitive development. This was one of the few relationships in this study which was firm enough that one might make an inference about future indigenous graduates: one can be sure, with less than 5% chance of error, that students matriculating with higher GPAs and then subsequently graduating from Evergreen will rate a bit higher in Perry position as freshmen, until something changes about high school, the students admitted to Evergreen, or the freshman curriculum. This correlation suggests that high school

*"Indigenous" Graduates, 1986-88: 67*

Table 6:  
GAIN AND HIGH SCHOOL GRADE POINT AVERAGE

Crosstabulation: GAIN      CHANGE IN RATING  
By GPA      HIGH SCHOOL GPA

GPA->	Count Exp Val Col Pct	Below median	Above median	Row Total
GAIN				
No gain	17 19.1 39.5%	22 19.9 48.9%	39 44.3%	
Gain	26 23.9 60.5%	23 25.1 51.1%	49 55.7%	
Column Total	43 48.9%	45 51.1%	88 100.0%	

GPA and cognitive development rating were both functioning as measures of academic performance. Although the mean cognitive development ratings for the two groups were significantly different (3.50 for freshmen with lower GPAs, 3.69 with higher;  $p = .04$ ), this difference was not large enough to have policy implications.

Table 6 shows that the initial advantage of high-achieving high school graduates disappeared after the freshman year, and high school GPA lost any predictive relationship with Perry position. More of the students who started out below the median GPA registered gain (60.5%) than of those who started out above the median (51%) - another tortoise-and-hare phenomenon. By the senior year, differences appear to have been ameliorated, so that the means were almost the same (senior ratings were 4.03 and 4.05, respectively, for the subgroups above and below the median GPA).

*Distribution of modes of study*

These findings suggest that Evergreen is not catering to an elite group of students in terms of high school preparation, and that a student need not initially be educationally advantaged to profit from the College, from the standpoint of cognitive development. But are all modes of study offered at the College of equal benefit, from this standpoint? Officially, all study modes are equal: a student can enroll in any year



in any kind of program she chooses, within the limitation that one program is sometimes the prerequisite for another one. And yet some modes and patterns of study are widely regarded as "more equal" than others. First, all students are advised to experience coordinated studies in a Core program when they first arrive on campus (or if they arrive as transfers, to take an interdisciplinary program or group contract with a seminar component). And second, students who make their way through Evergreen primarily in individual contracts and courses are sometimes thought to be missing a large part of the point of the College - to experience liberal education in an interdisciplinary environment of collaborative learning. These students may be part-timers who choose these modes of study because they are employed. Coordinated study programs and group contracts tend to be full-time, daytime programs, and so individual contracts and courses are the easiest modes of study to pursue while working. But perhaps something important is sacrificed. Do students who take more coordinated studies tend to come out ahead in cognitive development? Do those who take more seminars tend towards more gain? And do those who take more individual contracts and courses tend to miss out or fall behind, according to the present measure of cognitive development?

One might address these questions most directly by sorting out the mode of study of each student during each year, and then correlating the study mode with the amount of gain in cognitive development rating. But this sort of directness was not practical, because many students changed from one program or mode of study to another during a year and submitted more than one evaluation. Instead, the method used here was to sort students both into arbitrary quartiles and into rational categories (e.g., less than one quarter, more than one year), according to the number of units earned in each study mode. (The results were checked by computing "study mode ratios," or proportions of each student's whole program earned in each mode; but the results were nearly the same.)

The data revealed, first, the distribution of study modes elected by the indigenous graduates. The best indicators are the median and the middle 50% of students. The indigenous graduates seemed to prefer coordinated studies, which are interdisciplinary programs, team-taught, usually full-time and year-long (and usually beginning or intermediate in level, rather than advanced). The median indigenous student, if there were such a person, took just over half of her credits in coordinated studies, and the median number of units was 96, or exactly two years. The variation was wide ( $SD = 42$ ), with the middle 50% of students ranging from 68 to 122 units in coordinated

studies. But most indigenous students (87%) took more than a year of coordinated studies, and some (a surprisingly high 15%) were enrolled in coordinated studies for almost all of four years.

Group contracts are disciplinary or interdisciplinary offerings, usually full-time, taught by one or two faculty members at the intermediate or advanced level. The median duration of enrollment by indigenous graduates in group contracts was just over two quarters, with the middle 50% ranging from one quarter to four, and 30% of indigenous students enrolling in group contracts for more than a year.

Individual contracts, which include both individual study projects and internships, were less popular with indigenous students than one might expect. A third of the indigenous graduates took less than a quarter of individual contracts, and thus probably did not have an internship experience (though a few internships are incorporated into programs and group contracts). The median number of contract units was 27, or a little less than two quarters. A year of individual contracts - 48 units - marks the upper quartile in this mode, and only 4% were primarily individual contract students, i.e., were enrolled in contracts for more than two years.

Conventional courses, finally, have had a variable place in the curriculum, depending on the current status of part-time studies. When these students attended the College, part-time studies were not emphasized, and so course offerings were limited. The median number of units in courses was 12, with the middle 50% of the indigenous graduates taking from four to 22 units of courses. Courses were definitely the minor option: 64% of the students took 16 units or less of courses.

### *Coordinated studies*

An unexpected pattern emerged when gain in rating was crosstabulated with coordinated studies (Table 7). Among students who took a low number of credits in coordinated studies (up to one year), gain was less likely: only 37.5% of students in this category achieved a gain in Perry position by graduation. Those who enrolled in one to two years of coordinated studies programs much more frequently registered gain (62.5%), indicating the general value of seminar and interdisciplinary experience. But students above the median, those taking two to three years of coordinated studies, tended in fewer instances toward gain: only 40.5%. Finally, a small number of students who went beyond three years of coordinated studies were most likely of all to gain; the proportion was a quite high 69%. This descriptive statistic comes close to the

Table 7:  
GAIN AND COORDINATED STUDIES

Crosstabulation: GAIN CHANGE IN RATING  
By COORD YEARS IN COORDINATED STUDY

COORD-> GAIN	Count Exp Val Col Pct	One year	Two years	Three years	Four years	Row Total
No gain	10 7.5 62.5%	18 22.6 37.5%	22 17.4 59.5%	5 7.5 31.3%	55 47.0%	
Gain	6 8.5 37.5%	30 25.4 62.5%	15 19.6 40.5%	11 8.5 68.8%	62 53.0%	
Column Total	16 13.7%	48 41.0%	37 31.6%	16 13.7%	117 100.0%	

level of significance where one could draw inferences about other students in other years ( $p = .065$ ). It is not the result of a fluke in classification of a few cases, because the same pattern emerges from different modes of classification (either by dividing units of coordinated studies into quartiles, or by considering the proportion of the student's total program spent in coordinated studies).

This is a complex pattern, difficult to interpret. Overall, the value placed at Evergreen on coordinated studies is confirmed, if the point is that a student might tend to benefit in cognitive development by taking as much as two full years of coordinated studies. But beyond two years, the situation is more problematic. Benefit was a little less likely to accrue, and one might even claim that some students spent too long in coordinated studies, if cognitive development were the sole criterion. They might have remained too long in programs which were not advanced enough, or where too many other students were freshmen or new transfers. The level of instruction in many coordinated studies programs is not pitched to the third- or fourth-year student, and the task of "seminaring" with a coordinated studies peer-group might not have been challenging enough to induce development. Perhaps the activity of participating in seminars becomes ritualized, so that it is no longer a method of growth through inquiry; or perhaps the mode of study becomes too abstract, too much a matter of words and books and

opinions, without enough attention to the relationship between theory and practice. Following these clues, one might want to advise students not to spend more than two years in coordinated studies. Yet it was the small group of 16 students who enrolled in more than three years of coordinated studies who were most likely of all to gain in cognitive development rating. This reversal at the end of the continuum is hard to explain. A look at the names of the programs in which these students were enrolled revealed no obvious consistency. Some areas of the College, particularly management studies and Native American Studies, offer coordinated studies all the way to the senior year; but these particular programs were not disproportionately responsible for gain or no gain. Teaching experience and anecdotal evidence suggest the possible explanation that a few students take to interdisciplinary coordinated studies in a way that becomes almost a commitment to a way of life. Perhaps such students are the ones more likely to stay in coordinated studies through all four years and register gain.

#### *Group contracts*

A rather different pattern shows up in the relationship between gain and quarters of group contracts (Table 8). The pattern is statistically significant ( $p = .03$ ), allowing one to make inferences about past and future indigenous graduates. The inference is that according to the present measure, group contracts do not foster cognitive development as consistently as the other study modes. Students who did very little work in this mode - less than an academic quarter - were surprisingly prone to gain: 79%, which is the highest tendency toward gain to emerge in this study. Those who spent up to a year in group contracts were only a little more likely than not to register gain, and those with more than a year of group contracts were slightly prone to no gain (57.5%). But this does not necessarily speak against the value of group contracts; it only means that they are not especially conducive to the kind of cognitive development measured by the Perry scheme. In effect, these results confirm the Perry model as a measure of liberal arts and sciences values. Group contracts tend to emphasize problem-centered or disciplinary skills. The consolidation hypothesis, presented to explain what goes on in the sophomore and junior years, asserted that a main task in these years is to work up the disciplinary skills that make interdisciplinary education a reality and not just an ideal. Group contracts should contribute effectively to this sort of consolidation, and their effect should appear more positively in an outcome study that focused on disciplinary

Table 8:  
GAIN AND GROUP CONTRACTS

Crosstabulation: GAIN CHANGE IN RATING  
By GROUPCON UNITS IN GROUP CONTRACTS

GROUPCON->  GAIN	Count				
	Exp val	0 to 15	16 to 48	Over 48	Row
	Col pct	units	units	units	Total
No gain	4 8.9 21.1%	28 27.3 48.3%	23 18.8 57.5%	55 47.0%	
Gain	15 10.1 78.9%	30 30.7 51.7%	17 21.2 42.5%	62 53.0%	
Column Total	19 16.2%	58 49.6%	40 34.2%	117 100.0%	
Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5	
6.94344	2	.0311	8.932	None	

skills. Nevertheless, the data for group contracts suggest a hypothesis for further study regarding the substantial number of transfer students who take mostly group contracts with a vocational and disciplinary focus (e.g., in management or social services). The evidence from indigenous students suggests that opportunities for cognitive development can be lost if interdisciplinary liberal arts and sciences studies are sacrificed in order to take disciplinary group contracts selected on grounds of vocational relevance.

### *Individual contracts*

Individual contracts are so diverse that it is hard to generalize about them, particularly because the recording system available from the Registrar for this study did not distinguish contracts for advanced independent study from those for internships. Perhaps the very diversity of contracts causes them to gravitate toward the mean, in terms of gain in Perry position; but the data suggest no disadvantage, in terms of cognitive development, incurred by working heavily in this mode. Nor were the students in contracts identifiable by controlling for other variables. Age, in particular, is not related to enrollment in individual contracts. Among all students, older returning stu-

*"Indigenous" Graduates, 1986-88: 73*

Table 9:  
GAIN AND INDIVIDUAL CONTRACTS

Crosstabulation: GAIN CHANGE IN RATING  
By INDCON UNITS IN INDIVIDUAL CONTRACTS

INDCON->	Count Exp Val Col Pct	First quartile	Second quartile	Third quartile	Fourth quartile	Row Total
GAIN						
No gain	17 14.1 56.7%	10 11.8 40.0%	15 15.5 45.5%	13 13.6 44.8%	55 47.0%	
Gain	13 15.9 43.3%	15 13.2 60.0%	18 17.5 54.5%	16 15.4 55.2%	62 53.0%	
Column Total	30 25.6%	25 21.4%	33 28.2%	29 24.8%	117 100.0%	

dents may prefer the contract mode for completing college; but this was not true of indigenous students.

In Table 9, the first quartile consists of those students who studied very little - less than a quarter - in the individual contract mode. They were a little less likely to end up in a position of gain. Those in the second quartile spent about one quarter in an individual contract and were a little more likely to register gain. Perhaps most importantly, the third and fourth quarters are about equal, with 55% registering gain. The third quartile contains students who took two or three quarters of contracts, and the fourth, a full year or more (in most instances, between one and two years). No downturn in gain was associated with registration in this many units of individual contracts. These students were definitely not missing out, from the perspective of cognitive development. They probably exemplify a theme mentioned repeatedly in self-evaluations, that individual projects are an especially rewarding learning experience. One might argue on other grounds that there is such a thing as "too many individual contracts," but one cannot prove the point by referring to any tendency to miss out on cognitive development as measured here.

*Seminars and balanced programs*

But to put this finding in perspective, the indigenous graduates must be distinguished from other graduates who may have transferred from community college or from Normal State, without much experience of interdisciplinary studies or seminars. The indigenous graduates tended to have accrued a high number of units in coordinated studies, so that they had an interdisciplinary background to build upon. What can be said is that nothing is lost and something is apparently gained if students *with this background* take a year or more of contracts. But one cannot generalize from the experience of students with this background to others who might lack it.

From an overall perspective, this study tends to confirm the widely-held intuitive view that an interdisciplinary liberal arts and sciences background is advantageous from the standpoint of cognitive development, but that a mix of different modes of study is valuable, particularly because coordinated studies programs often do not offer advanced work. The earlier discussion of class profiles pointed to the high gains of freshmen involved for the first time in seminars, but this should not be taken as a case for offering only programs with seminars, at the expense of a balanced program of studies. One might even say that students ought to go beyond what Evergreeners call "seminaring," which is essentially a habit of talk, and develop some kind of senior or culminating project - an advanced independent study project or a connection through an internship of theory to practice. The case for a balanced program emerges if one classifies coordinated studies and group contracts together, as programs involving seminars, in contrast to individual contracts and courses, which are far less likely to offer seminars. In Table 10, the breakdown is by quartiles, because breakdown by academic quarters or years of study would provide groupings too unequal in size. The first quartile represents students who took up to 65% of their academic work in programs with seminars, amounting to about 120 units. The second quartile goes up to 77%, or 144 units - exactly three years. The third quartile extends to 88%, or 163 units (ten quarters); and the top quartile goes all the way to graduation. The pattern is reminiscent of the results for coordinated studies. The table shows that solid grounding in two years of programs with seminars tended to result in gain, and that the probability of gain was even higher (67%) with two to three years of seminars. But over two years of enrollment in seminars was less conducive to increase in rating. Something must be lacking. It is not advanced study, because group contracts are included here. Two possibilities are insufficient opportunity for self-directed individual work, or failure to realize the

Table 10:  
GAIN AND SEMINARS

Crosstabulation: GAIN CHANGE IN RATING  
By SEMINARS % CREDITS WITH SEMINARS

SEMINARS->  GAIN	Count Exp Val Col Pct	First quartile	Second quartile	Third quartile	Fourth quartile	Row Total
No gain	14 14.6 45.2%	10 14.1 33.3%	18 13.6 62.1%	13 12.7 48.1%	55 47.0%	
Gain	17 16.4 54.8%	20 15.9 66.7%	11 15.4 37.9%	14 14.3 51.9%	62 53.0%	
Column Total	31 26.5%	30 25.6%	29 24.8%	27 23.1%	117 100.0%	

College's avowed value of relating theory to practice. But then again (as with coordinated studies, considered separately), the seminar "addicts" return to a higher probability of gain.

This look at demographic and academic variables has pointed toward some particular conclusions, but the overall conclusion is that except for fairly minor points, this study does not predict which subgroups will gain most from Evergreen, in terms of cognitive development. This is not a failure, because it suggests that at Evergreen, cognitive development - impressive from the perspectives of overall gain and of class profiles - is not distributed unfairly to some groups and not to others, either by socioeconomic variables or in different modes of study. The College can report that by this outcome measure, the results look good, and that nothing appears to be broken or in need of fixing. No findings about variations among subgroups suggest changes in academic or admissions policy. Change may, of course, be advocated on other grounds, to maximize other values than cognitive development as conceived in the Perry model - for example, greater diversity in the student body. But these findings about cognitive development do not argue for or against any particular change.



## CONCLUSIONS AND APPLICATIONS

### Research agenda

A study in the present genre - a post hoc analysis in the *bricoleur* tradition - cannot settle many points definitely, and so the first conclusion must be that a long agenda remains for further research. Ideally, further research would begin with a similar retrospective longitudinal study using a stratified sample of all sorts of students. The most pressing unresolved issue to pursue in such a study is what happens to transfer students when they make use of the College at different points in their lives, transferring in and out in different years and staying for different durations. Are transfers a homogeneous group, or do they differ markedly between, for example, younger and older? Does the first year at Evergreen have the same powerful effect on transfers as on indigenous freshmen? Is the senior surge replicated among community college transfers, or are they, as seniors, in an assimilative phase, following their initial experience in interdisciplinary studies?

Additional items on the research agenda should be, first of all, the effect of the College on students of color. They were too few, among the indigenous graduates, for their presence to register significantly; but they would be more numerous in a transfer student study or in research directed toward them specifically.

A second uncertain issue was the effect of the College on male and female students. On the basis of Belenky et al.'s *Women's ways of knowing* (1986), it made sense to hypothesize that women would move ahead in cognitive development at Evergreen, because of curricular emphasis on finding one's "voice" in seminars, and on engaging in collaborative learning and "connected knowing." But the expected result did not emerge from the data. One possibility is that the rating system does not give equal emphasis to women's voices or styles of expression, and another is that seminars and collaborative learning might be equally advantageous to men. These are intriguing rival hypotheses which deserve further research.

Thirdly, this study raised an important issue about "tortoises and hares," who were, respectively, those students who rate high at the end of the freshman year, but

subsequently register little or no gain; and those who initially receive low ratings, but tend toward more gain by the time of graduation. The issue raised by the tortoises and hares is whether the College is equally stimulating to students who employ more complex cognitive structures. This suggests more focused research on the effect of advanced studies.

This project also points toward ways of improving the art or craft of the cognitive development rating system based on Perry's model. One improvement would be to refine with more precision the relationship among ratings based on different kinds of writing samples and portfolio materials - MID essays, student self-evaluations, and other kinds of academic writing. Another is to look into student self-evaluations not to rate them, but to reconsider the rating criteria (to rate the rating system, so to speak). Perry's own position is that Perry positions are relatively fixed cognitive structures or epistemologies, and that time, gender, and ethnicity probably bring about variation only in style of expression, and not in fundamental ways of knowing (personal communication; and Thompson, 1988). But certainly it is time, after more than 20 years, to review documents such as MIDs and SSEs to see again, as Perry did in the 1950s and 1960s, what language students actually use to describe their learning experiences, and to find out if any changes have been wrought by time, especially in relation to the gender, race, and class of contemporary college students.

And finally, this study hopes to influence research agendas elsewhere, with a payoff in terms of comparable data. Comparisons in this study have been awkward, with no reference point except the aggregate comparative data for what has been referred to as "Normal State." Many issues raised here - for example, the ratings for "tortoises" and "hares" - are hard to interpret without comparable data for other colleges and universities in general, and liberal arts and sciences institutions particularly.

### **Conclusions**

Aside from the standard researcher's conclusion that more research is necessary, this study has reached some tentative substantive conclusions.

None of these conclusions, however, are more important than the conclusions and hypotheses which occur independently to readers. The study was written to provide information, to stimulate thought, and to provide a context for discussing cognitive development at Evergreen and elsewhere. The most important conclusions will occur to members of the Evergreen faculty, administration, or student body, or to their peers

anywhere else who are interested in assessing or understanding cognitive development. Ideally, members of the Evergreen faculty will find information and a vantage point here that will in the long run have some effect on their teaching. It was the hope of Evergreen's Assessment Study Group (the College's assessment steering committee) to support studies which would result not merely in conclusions good for public or legislative relations, but which would promote an internal dialogue about teaching and learning. Dialogue moves forward through comparing different interpretations, and thoughtful conclusions can of course take the form of dissent. In this context, the most important conclusions of this study are those of the reader, to which the author's conclusions, which follow, are secondary.

One justifiable conclusion is that The Evergreen State College is a successful institution of higher education from the standpoint of fostering cognitive development. By the outcome measure employed here, it is fulfilling its mission, which is to provide an alternative form of liberal arts and sciences education:

The role of the institution is to provide high quality undergraduate education to appropriately prepared students by offering a unique curriculum of liberal arts and sciences characterized by interdisciplinary studies. This is done through close faculty-student contact at all levels of the curriculum, and collaborative teaching and learning activities. . . .

Evergreen's fundamental mission is to assist students in learning how to learn and how to continue developing their skills in a world of increasing diversity, interdependence, and moral complexity. The highest priority is placed on the quality of undergraduate instruction.

Evergreen approaches that task with the tools of a traditional college: the disciplines of the humanities, arts, the natural sciences, and the social sciences. However, those disciplines are transformed at Evergreen into teaching and learning experiences characterized by: interdisciplinary learning communities . . . ; internships and applied projects which bridge theory and practice; small classes and narrative [evaluation] which require, even at the beginning level, active involvement of students; independent study options and self-evaluations where students take responsibility for their own learning; and a campus environment which celebrates diversity as a resource for learning (The Evergreen State College, 1989, p. 6).

Beginning with this mission statement, one might say more forcefully that Evergreen's *raison d'être* is to provide an alternative that addresses educational problems which have become ever more widespread since the College was founded, and which are attracting more and more attention. The main problem is pervasive bureaucratization in mass undergraduate education, including large and impersonal classes, departmentalized specialization, remote or entrepreneurial faculty, and loss of a sense of aca-

demic community (see, e.g., Smith, 1990). Evergreen's alternative could be judged favorably if it were offered to an appreciative clientele and simply did as well as other institutions of higher education, in Washington or nationally, by reasonable outcome measures. "As well as" would be good enough, and "better than" would be an extra increment of value which would recommend Evergreen's special features, as outlined above in the mission statement, for consideration at other institutions.

By the outcome measure used in this study - a measure which seems especially appropriate to undergraduate liberal arts and sciences education - Evergreen seems to have done substantially better than the mainstream, as depicted in the comparative data. The data for Normal State are not fully up-to-date; they are directly comparable for only the freshman and to some extent the senior years at Evergreen; and some reliability is lost after the freshman year in comparing MID essays from other institutions with portfolio materials from Evergreen. But after enumerating these qualifications, the evidence still suggests that Evergreen is doing better than most mainstream institutions in stimulating cognitive development. Using fully comparable MID ratings, Evergreen freshmen advance as far in modal position and much, much farther in mean position than students represented in the comparative data. And if seniors elsewhere can be compared to graduating seniors at Evergreen, then the latter have advanced just as far from their freshman baseline in modal cognitive development rating, and much farther in mean rating. A substantial percentage end up in Position 4 and above - ratings which are not often seen in the comparative data. The cognitive structures used by Evergreen graduates in writing self-evaluations definitely seem more complex than one would have predicted from the data about students elsewhere.

But as soon as this has been said, one also needs to consider a second conclusion, that most Evergreen students, most of the time, are still working as problem-solvers within Position 3, early multiplicity or separate, procedural knowing; or they are in transition to what Perry called "late multiplicity," or what most other educational writers call relativistic thinking. Only in the senior year do a substantial number of the indigenous graduates move beyond Position 3 and the 3-4 Transition; and even then, they never become the majority. This is a finding with definite relevance to teaching. Though they move ahead of their peers at other institutions, Evergreen students still tend to be at one with most of American higher education, conceiving knowledge and learning as problem-solving and procedural knowing. The students' self-evaluations reveal that the great bulk of learning activity remains in this heartland, and though fac-

ulty may have much more complex objectives in mind, most students are perceiving what is taught within the perspective of a plurality of skills and methodologies, without yet fully appreciating the faculty's concerns with contextual relativism and more complex cognitive operations.

This leads to a third conclusion, that students and faculty tend to operate from quite different positions in the Perry model. Students, even in the senior year, are still unlikely to occupy Positions 4 and above, whereas faculty probably function from Positions 4 through 9. This is one justification for having teachers and paying them salaries: they tend to be able to help others to grasp complex ideation in their disciplines; and at Evergreen, they tend also to be skilled in designing interdisciplinary coordinated studies programs, which one might think of as local monuments to contextual relativism and constructivism.

The cognitive structures employed by faculty members vary with time and task. The Perry model does not suggest that either faculty or students attain some highest level of development and then stay there forever. On the contrary, Perry (1981) argues that the positions are less like an invariant sequence than like a spiral or helix, through which one recycles every time one enters a new area of inquiry. In other areas of life than the academic, and in academic circumstances involving more novelty, emotion, or stress, one tends to revert to less complex ways of construing the situation (see also Kegan, 1982). Perry has a saying that "every guild regresses its new members to the lowest rank," meaning that whenever one starts a new enterprise, one has to recapitulate the developmental process. This would apply especially to teaching in interdisciplinary studies, in which faculty are constantly challenged to work in areas where they lack expertise. In this circumstance, they are true colearners. They may operate confidently in their field of expertise from a dominant cognitive position which is quite advanced (5 through 9), and yet they may have to spend much of their time and effort in coordinated studies making sense of new information, methods, and perspectives - the work that goes on within Positions 2, 3, and 4.

With this qualification about its spiral or helix structure, the Perry model still suggests that faculty members tend to operate from more complex perspectives. The faculty member always models her own structures and styles of cognition; that is part of teaching. In the words of an Evergreen faculty member, "I want to be a [Position] 9 with my students and want them to see me that way." A faculty member's dominant cognitive position is one dimension of her integrity. But in the words of another col-

league, "there's something to be said for being understood." The model, by providing a language for talking about successively more complex ways of thinking, suggests that whatever the faculty member is trying to communicate, it must be relevant to the student's current cognitive schema or epistemology.

This implies that faculty must be vigilant and sensitive in tracking the great diversity of cognitive activities going on in every classroom discussion. This diversity might rank as a fourth major conclusion: because some Evergreen students make more progress in cognitive development than students elsewhere, and because students are generally not forced to pass through a gauntlet of prerequisites numbered 101, 201, 301, etc., and because indigenous Evergreen students are joined particularly in the junior year by an infusion of transfer students, the diversity of epistemologies or cognitive perspectives in play at one time in an Evergreen classroom is probably extraordinary. And furthermore, this is not the quiet diversity of passive listeners in lecture halls; it is the active diversity of participants not all of whom can easily understand one another. One might even argue that the prevailing norm of gentleness, mutual acceptance, and nonconfrontiveness in Evergreen seminars is more than good modern manners; it might also be to some extent a defense against the substantial incomprehension which would be revealed if anyone were so confrontive as to lift the veil. Seminar conversations have a tendency to scurry off into safe byways when real incomprehension and misunderstanding begin to emerge. Ideally, diversity is the very basis of collaborative learning; but one of the functions of the faculty member would seem to be to keep the students who are least and most advanced in cognitive perspectives from ending up isolated from, or frustrated by, each other.

### **Applications**

At this point, the Perry model suggests some specific applications of theory to practice. These may be new ideas, or they may be only a description of what effective college teachers are already doing. But qualifications are in order, lest it appear that the thrust of the argument is simply to institutionalize the Perry scheme as a prescriptive model. The suggestions are offered as ideas to try on for size, rather than as a template to impose on the classroom. In applying any theory or research to educational practice, one must make the application "contextually relative," in Perry's terms. The point cannot be to impose any one set of ideas or results on the teaching situation. A classroom is not simply a stack of Perry positions. The environment is complex, and

different teachers function in relation to different classes and students with different methods and objectives. The Perry model illuminates just one more dimension in a complex educational context. Cognitive structure is only one component of classroom diversity. Other components include academic background, and social and psychological variables - e.g., class, gender, ethnicity, and psychological type or learning style. From the perspective of contextual relativism, applications of the Perry model to teaching have to fit appropriately into the faculty member's larger awareness of what goes on in the classroom. Ideally, Perry research contributes to a heightened sense of the diversity of student perspectives. And it can also contribute to a sense of community, if the faculty member uses the scheme to remember that she also, at some past time, operated from precisely that position, perspective, or epistemology in which a student or class seems stuck.

With these notes of caution, the Perry model suggests applications of theory to practice which derive from the ideas of "developmental instruction" and of a "Plus One rule" mentioned in the introduction. The model suggests that among the many goals of the classroom teacher, one important and conscious goal can be *to induce transitions*, from one stage to another. In helping students reach transitions, the teacher has to offer access to the *next* position - not one that is too far over the horizon to be understood. The "Plus One rule" holds that students can hear and appreciate a position which is about one stage ahead of what they are currently producing in their own oral and written work. For example, a student currently producing her own work (her seminar participation, papers, labs, artwork, etc.) from Position 3 or the 3-4 Transition can "hear" Position 4, though probably not yet Position 5. A teacher has to offer access to the next position, not one that is too far over the horizon to be understood. Discourse from more than a position ahead probably cannot be assimilated or accommodated. It can be written down dutifully in fragments, and memorized; it can be translated into the less complex structure that the student was able to hear (which accounts for some of the odd translations of their own remarks which teachers hear from students); or it can be experienced as nothing more than baffling static ("The lecturer is talking all over the place"), or as a kind of verbal assault administered by Authority. But it is unlikely to contribute to change, which occurs at the threshold of the next accessible position. The problem, then, is to keep lectures and especially seminars relevant to the developmental potentials of the students present - or in Vygotsky's well-chosen words, relevant to "the level of potential development as determined by inde-

pendent problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, in Bruner, 1986).

In lecturing, one of the main problems is to make sense to students who presently reside in different cognitive positions. Most good lecturers use a recursive style, offering some material which develops new ways of thinking, but finally bringing the discussion back to foundations which all students can fully appreciate. In seminars, the dominant mode of instruction at Evergreen, the situation is more complicated, because part of the teaching is done by faculty, and part by peers. The seminar leader can often sit back and listen to students teaching one another through what can be heard as a dialogue among cognitive positions. But not all positions are represented, nor are they equally articulate; and some are too far apart to communicate without translation. The "Plus One rule" and a concern for change thus entail some suggestions for the seminar leader's role: (1) let the students proceed alone, whenever they can teach one another; (2) model the higher "Plus One" position which is not supported from within the group; (3) attend to those students who are alone in their position; (4) mediate and translate mutual incomprehension.

A final point, more a cautionary note than a conclusion: the indigenous graduates, it must be remembered, are those on whom a college has had its greatest impact; but they are not typical of the student body as a whole. The difference arises because throughout the sophomore, junior and senior years, indigenous students are joined by transfers from the community college system and from other colleges and universities. According to the comparative data, many of these students will be using the cognitive structures of Position 3. The present study shows that some continuing students, especially seniors, have been advancing toward the 4-5 Transition and Position 5, and thus will be using epistemologies which are over the horizon for some of their new classmates. Furthermore, in upper-division classes these differences in cognitive perspective are sometimes layered with other differences, for example, in age or in vocational versus academic orientation. The diversity of epistemologies in upper-division seminars can be extreme. In general, diversity is a value, and students can teach each other in seminars; but it is not hard to imagine groups in which cognitive positions are far enough apart to cause as much anxiety and frustration as learning. In such circumstances, the "Plus One rule" suggests that the function of the faculty remains to model more complex positions; but much of their work would seem to be to build bridges across Position 4. They must help students in the 3-4 Transition to hear and appreciate



the discourse of those in the 4-5 Transition and above, which will not always be easy.

And the students in the higher positions will especially need interaction and a feeling of community and comradeship with the faculty, to support their cognitive development. Perry is helpful on this point (1970, pp. 121-126), emphasizing both the capacity and the need of students in the higher positions for a sense of community with the faculty. Such a sense of community is precisely what the Evergreen alternative, with its small classes, close student-faculty contact, and emphasis on collaborative learning, should be able to offer. A sense of community between teachers and students - a real "community of learners" - is no doubt one of the most generous benefits which can flow from small classes and direct contact between students and faculty.

## REFERENCES

- Belenky, M. F., Clinchy, B. McV., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.
- Bloom, Allan (1987). *The closing of the American mind: How higher education has failed democracy and impoverished the souls of today's students*. New York: Simon & Schuster.
- Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Council of Presidents and State Board for Community College Education (1989). *The validity and usefulness of three national standardized tests for measuring the communication, computation and critical thinking skills of Washington State college sophomores: General report*. Bellingham, WA: Western Washington University Office of Publications.
- The Evergreen State College (1989). *Constancy and change at Evergreen: A self-study report to the Northwest Association of Schools and Colleges*. Olympia, Washington: The Evergreen State College.
- Hutchings, P., Amiran, M., Halonen, J., & Scholnick, S. (1989). Getting portfolios into practice. Presentation, AAHE Fourth National Conference on Assessment in Higher Education, Atlanta, GA.
- Kegan, R. (1982). *The evolving self: Problem and process in human development*. Cambridge, MA: Harvard University Press.
- Knefelkamp, L. L. (1974). *Developmental instruction: Fostering intellectual and personal growth in college students*. Unpublished doctoral dissertation, University of Minnesota.
- MacGregor, J. (1987). *Intellectual development of students in learning community programs 1986-87*. Olympia: Washington Center for Undergraduate Education, Occasional Paper No. 1.
- Martin, G. E. (1982). *Encounter with education: Impact of an alternative college on student development*. Unpublished Ed.D. dissertation. Teachers College, Columbia University.
- Mentkowski, M. (1988). Paths to integrity: Educating for personal growth and professional performance. In S. Srivastva & associates, *Executive integrity: The search for high human values in organizational life* (pp. 89-121). San Fran-

cisco: Jossey-Bass.

- Mentkowski, M., Moeser, M., & Strait, M. J. (1983). *Using the Perry scheme of intellectual and ethical development as a college outcomes measure: A process and criteria for judging student performance*. 2 vols. Milwaukee: Alverno College Productions.
- Mentkowski, M., & Strait, M. J. (1983). *A longitudinal study of student change in cognitive development, learning styles, and generic abilities in an outcome-centered liberal arts curriculum*. Milwaukee: Alverno College Productions.
- Moore, W. S. (1982/86). *The Measure of intellectual development: A brief review*. Farmville, VA: Center for Applications of Developmental Instruction. (Originally 1982; revised 1983, 1984, 1986.)
- Moore, W. S. (1986). *Understanding students-as-learners through the Perry scheme: A rater training workshop*. Farmville, VA: Center for Applications of Developmental Instruction.
- Moore, W. S. (1989). *The measure of intellectual development: An instrument manual*. Farmville, VA: Center for the Study of Intellectual Development.
- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.
- Nelson, C. (1988a). Varieties of critical thinking in science: From data to dialectics. Workshop materials presented at The Evergreen State College.
- Nelson, C. (1988b). Addressing process and content: Teaching moves based on Perry and *Women's ways of knowing*. Workshop materials presented at conference on "Conflict over content: Knowledge, values, and higher order inquiry," University of Chicago.
- Perry, W. G., Jr. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart & Winston.
- Perry, W. G., Jr. (1981). Cognitive and ethical growth: The making of meaning. In A. Chickering & associates, *The modern American college*, pp. 76-116. San Francisco: Jossey-Bass.
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Polanyi, M. (1967). *The tacit dimension*. Garden City, NY: Anchor Books.
- Shor, I. (1987). *Critical teaching and everyday life*. Chicago & London: University of Chicago Press.
- Smith, P. (1990). *Killing the spirit: Higher education in America*. New York: Viking.

References: 87

- Thompson, K. (1988). *A report on "The Perry clan gathering."* Olympia: Washington Center for Improving the Quality of Undergraduate Education.
- Thompson, K. (1989a). *Assessment in integrated studies: First results from the Evergreen assessment project.* Paper presented at the A.I.S. Conference on "Integrated Studies: Achievements and Consequences," Sonoma State University, October 14, 1989.
- Thompson, K. (1989b). *Narrative assessment in collaborative learning environments: Encouraging connected and constructed knowing.* Paper presented at the AAHE Conference, Chicago, IL. Olympia: Washington Center for Improving the Quality of Undergraduate Education, Occasional Paper No. 2.

## Appendix A:

### Scheme of Cognitive and Ethical Development \*

Dualism modified ↓	Position 1	Authorities know, and if we work hard, read every word, and learn Right Answers, all will be well.
	Transition	But what about those Others I hear about? And different opinions? And Uncertainties? Some of our own Authorities disagree with each other or don't seem to know, and some give us problems instead of Answers.
	Position 2	True Authorities must be Right, the others are frauds. We remain Right. Others must be different and Wrong. Good Authorities give us problems so we can learn to find the Right Answer by our own independent thought.
	Transition	But even Good Authorities admit they don't know all the answers <i>yet!</i>
	Position 3	Then some uncertainties and different opinions are real and legitimate <i>temporarily</i> , even for Authorities. They're working on them to get to the Truth.
	Transition	But there are <i>so many</i> things they don't know the Answers to! And they won't for a long time.
Relativism discovered ↓	Position 4a	Where Authorities don't know the Right Answers, everyone has a right to his own opinion; no one is wrong!
	Transition (and/or)	But some of my friends ask me to support my opinions with facts and reasons.
	Transition	Then what right have They to grade us? About what?
	Position 4b	In certain courses Authorities are not asking for the Right Answer; They want us to <i>think</i> about things in a certain way, <i>supporting</i> opinion with data. That's what they grade us on.
	Transition	But this "way" seems to <i>work</i> in most courses, and even outside them.
	Position 5	Then <i>all</i> thinking must be like this, even for Them. Everything is relative but not equally valid. You have to understand how each context works. Theories are not Truth but metaphors to interpret data with. You have to think about your thinking.
Commitments in Relativism developed ↓	Transition	But if everything is relative, am I relative too? How can I know I'm making the Right Choice?
	Position 6	I see I'm going to have to make my own decisions in an uncertain world with no one to tell me I'm Right.
	Transition	I'm lost if I don't. When I decide on my career (or marriage or values) everything will straighten out.
	Position 7	Well, I've made my first Commitment!
	Transition	Why didn't that settle everything?
	Position 8	I've made several commitments. I've got to balance them—how many, how deep? How certain, how tentative?
	Transition	Things are getting contradictory. I can't make logical sense out of life's dilemmas.
	Position 9	This is how life will be. I must be wholehearted while tentative, fight for my values yet respect others, believe my deepest values right yet be ready to learn. I see that I shall be retracing this whole journey over and over—but, I hope, more wisely.

\*From Perry, W. G., Jr., "Cognitive and Ethical Growth: The Making of Meaning." In A. Chickering and Associates, The Modern American College. San Francisco: Jossey-Bass, 1981, Chapter 3, pp. 76–116.

Appendix B:

MEASURE OF INTELLECTUAL DEVELOPMENT:

COMPARATIVE DATA

-----Per Cent-----

<u>Class</u>	<u>N</u>	<u>Mean</u>	<u>Pos 2</u>	<u>Tr 2-3</u>	<u>Pos 3</u>	<u>Tr 3-4</u>	<u>Pos 4</u>	<u>Tr 4-5</u>	<u>Pos 5</u>
Fresh	1695	2.80	4.7	44.1	38.9	11.0	1.3		
Soph	367	2.88	1.9	42.0	37.6	15.3	2.7	0.5	
Jun	358	2.91	2.5	33.0	47.2	15.4	1.4	0.3	0.3
Sen	337	2.98	1.8	29.7	46.9	15.4	4.7	1.5	
<u>Age</u>									
18	378	2.87	1.1	40.5	45.0	11.4	2.1		
19	229	2.81	1.3	48.9	38.9	7.9	3.1		
20	200	2.87	0.5	41.0	44.5	11.5	2.5		
21	116	2.91	0.9	35.3	46.6	15.5	1.7		
>22	99	2.90		43.4	41.4	10.1	2.0	2.0	1.0
<u>Gender</u>									
M	526	2.92	1.7	40.1	37.2	15.7	4.3	0.5	0.1
F	1287	2.89	1.0	37.2	47.0	11.8	2.4	0.3	0.1

Pos = Position; Tr = Transition

Source: Moore, 1982-83, revised April 1984