DOCUMENT RESUME

ED 269 780 CS 209 737

AUTHOR Haswell, Richard H.
TITLE Change in Undergraduate and Post-Graduate Writing Performance: Quantified Findings.
PUB DATE [84]
NOTE 76p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Age Differences; Cohesion (Written Composition); Comparative Analysis; Developmental Stages; Graduate Students; Higher Education; Longitudinal Studies; Measurement Techniques; *Skill Development; Syntax; Undergraduate Students; *Writing Improvement; *Writing Research; *Writing Skills

ABSTRACT

To determine whether the development of writing skills changes across college years, a study examined the writing of average-age college freshmen, sophomores, and beginning juniors. Writing samples were obtained during the first week of the first semester, last week of the first semester, first week of the third semester, and first week of the fifth semester. The method was cross-sectional, resulting in four independent sets of 32 essays each for comparison. A fifth group of essays was added for further comparison: those of older "competent" writers, post-graduates working in a variety of occupations. A total of 128 essays was analyzed using 107 measures in the following categories: overall quality, ideas, support, organization, diction, syntax, and mechanics. The most significant finding was that writing did seem to change during the first two years of college, in varied and substantial ways, toward what may be characterized as competent, more mature writing. (Tables of statistical data and a six-page bibliography are included.) (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Richard H. Haswell
Washington State University
Pullman, WA 99164

ED269780

Change in Undergraduate and Post-Graduate Writing Performance: Quantified Findings

INTRODUCTION

Four years ago the editor of College English lamented that in three years not one article on the development of writing skills during college had crossed his desk: "I have yet to read about how most people continue to learn to write, if they do, after they attain a competence that gets them out of first-year writing courses" (Gray, 1982). Since then, this particular lacuna in scholarship has not been filled. Nor do earlier studies of post-freshman college writing compensate for the current lapse. Despite substantial studies of writing development in the schools and numerous studies of change within freshman composition courses, investigation of change across the college years is thinly scattered, piece-meal, non-replicated, and difficult to interpret (see Eblen, 1981; Freedman and Pringle, 1980; Hammond, 1984; Jewel, Cowley, and Rhum, 1969; Kerek, Daiker, and Morenberg, 1980; Kimmey, 1975; Kitzhaber, 1963; Maimon and Nodine, 1979; Miller, 1980; Stewart, 1978; Whitla, 1977). As a result, one can only barely begin to answer even the most basic of questions, for instance whether student writing changes for the better, if it changes at all. Decisions on undergraduate writing curriculum remain largely intuitive

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Richard H. Haswell
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS209737

because we have no agreed-upon ontogenetic data.

The present study intends to add to our stock of data, at least to that pertaining to the average-age freshman, sophomore, and beginning junior college student. It sampled writing of students making normal progress at a public, land-grant university, at four points: first week of the first semester, last week of the first semester, first week of the third semester, and first week of the fifth semester. The method was cross-sectional, so there were four independent research groups for comparison. A fifth group was added for further comparison, of older "competent" writers, post-graduates working in a variety of occupations. The writing of these five groups was analyzed by means of a deliberately broad spectrum of quantifiable measures.

The ultimate purpose of the investigation was to acquire data from which to draw implications for teachers and administrators in college composition programs. Those implications have now been written up in a full study (Haswell, 1986a). But that study is speculative and complex, and the original data is difficult to extract. Here I will provide only a straight-forward description of the research project and a report of the empirical findings, appending the briefest of conclusions. It is primarily for researchers, but other interested students of writing development may find information of interest among these data. The most significant conclusion--that under-class students here did in fact change, and in varied and substantial ways, toward older, competent writing--will be seen readily enough.

This report of findings is organized as follows:

I Elicitation of the writing . . . . . p. 3

II The research groups . . . . . p. 5

III Procedure of analysis . . . . . p. 12

IV Rationale for choice of measures . . . . . p. 13

V Description of measures . . . . . p. 17

VI Statistical procedures . . . . . p. 37

VII Findings . . . . . p. 38

Main table . . . . . p. 39

Additional tables . . . . . p. 48

VIII Conclusions . . . . . p. 59

IX Acknowledgments . . . . . p. 61

X Bibliography of works cited . . . . . p. 62

I. Elicitation of the writing

ELICITATION OF THE WRITING. On the second class meeting of the fall semester of 1980, all composition students at Washington State University wrote a 50-minute essay on one of two matched topics, without fore-knowledge or choice of topic. They wrote assuming that their essay, along with their answers to a brief questionnaire about age and course-work, would be used by their teacher for diagnosis of writing level (as it was). At the end of the semester, students, but only in selected freshman composition sections, repeated the process, each given the topic not written on earlier. Teachers turned all of the essays and questionnaires, over 2,500 of them, in to the researcher.

The older, college-graduate writing sample was gathered by paid research assistants. They contacted employers or managers in a supervisory

position at businesses, industries, and public organizations located at four population centers around the state of Washington: the Seattle, the Tacoma, the Spokane, and the Tri-Cities areas. These supervisors selected one employee, who was given space and fifty minutes to compose an essay on one of the same two topics given to the students. The employee also filled out a brief questionnaire. The writing was monitored and collected by the research assistant.

The two writing prompts used with all subjects were designed to elicit equivalent rhetorical responses, yet to offer a broad enough range of topic to allow writers of varying age and experience equal access to ideas for expression. Topic A read:

> We are all aware of the part that ideals of physical appearance play in our lives. But there remain many questions about the nature of these ideals, questions worth considering. Are conceptions of human "beauty" or "handsomeness" different for different age levels? All told, do these conceptions cause more benefit or more harm? Why are such standards created and maintained? How are they spread? Are they more difficult for males or for females to handle? What part do they play in politics, in the cosmetics and clothing industries, in dating customs, in attitudes toward physical handicaps?
>
> Clearly, these are just a few of many questions related to this central problem of the way people judge physical appearance. Imagine that a researcher wants to know your understanding of one of these questions. Focus on one and write a unified, organized, and well-developed essay setting forth your ideas.

Topic B, on "the part that codes of conduct play in our lives," paralleled Topic A in style and content.

II. The research groups

THE RESEARCH GROUPS. From the pool of student writing, 128 essays were selected, always using techniques of random sampling, to form four carefully delimited groups of 32 essays each. Here are the groups, with their specifications and with the acronyms by which they will be alluded to hereafter:

FR: written by 18-year-old entering freshmen enrolled in freshman composition, evenly divided by topic and by sex of writer.

SO: written by 19-year-old sophomores making regular academic progress (30-32 semester hours earned), enrolled in advanced composition, not transferred from another school, freshman composition but no other writing course taken, evenly divided by topic and sex of writer.

JU: written by 20-year-old juniors, 60-64 semester hours earned, otherwise identical to the SO group.

EF: written by 18-year-old freshmen, exiting the freshman course, otherwise identical to the FR group.

Not present in the FR and EF population was about 4% of the freshman class, namely those who had entered the Honors program or who had exempted out of regular freshman composition by means of an essay writing test. Both kinds of students were excluded from the SO and JU sample. The FR and EF groups sampled the rest of the entering-freshman population, who had to enroll in freshman composition as a university graduation requirement. SO and JU students were enrolled in either an advanced composition course taken mainly by non-humanities majors or a parallel course taken mainly by humanities majors. Sixteen writers (8 male and 8 female) were chosen from the former and six (3 male and 3 female) from the latter, a proportion matching the enrollment ratio of the two courses.

At the time more than a third of the post-freshman undergraduate population took one of these two courses. As with the FR students, very nearly all were taking the course as a requirement, this time for their major, and whether they took it as a sophomore or a junior depended almost entirely on their major's schedule of courses. In short, self-selection can hardly be a factor in the formation of any of the groups.

An effort was made also to exclude the influence of teacher. The 32 FR students represent 32 of 64 different course sections; the 64 SO and JU students represent 25 of 27 different course sections, no more than three students from any one section.

Most important, these four student groups were selected to form groups equivalent in initial (pre-test) verbal skill. In order to do so, the Verbal Composite score was used from the Washington State Pre-College Examination. The WSPC is a sophisticated program of assessment (Noeth, 1979) taken by 95% of entering freshmen at Washington State University. The Verbal Composite is expressed in standard score units (range = 20.0 to 80.0, $\bar{X}$ = 50.0, SD = 10.0) and is derived from four subtests: reading comprehension, English usage, spelling, and vocabulary. It was assumed and later found to be the single best predictor of general compositional quality of the sample essays, correlating at .63 with the holistic evaluation (M1). With the first random drawing of essays, the group means of the Verbal Composite indicated a trend upward with age: FR, 47.7; EF, 50.6; SO, 50.5; JU, 51.8. Comparative ranking divulged that the freshman sample was deficient in Verbal Composite scores above 60 compared to the SO and EF samples, and the junior sample deficient in scores below 40. Since it was felt that this imbalance might compromise group comparisons, random selection was continued (leaving the EF and SO selection intact) until the spread of Verbal Composite scores between groups was 4 or

less at each of the 32 ranks. The change from the original sample was small: 5 FR essays and 2 JU essays were replaced. The final selection shows groups with highly equivalent means and distributions of all pre-test verbal indicators.

	FR GROUP		EF GROUP		SO GROUP		JU GROUP	
	$\bar{X}$	(SD)	$\bar{X}$	(SD)	$\bar{X}$	(SD)	$\bar{X}$	(SD)
High-school GPA	3.3	(0.50)	3.2	(0.42)	3.3	(0.37)	3.3	(0.41)
High-school English GPA	3.2	(0.69)	3.2	(0.50)	3.3	(0.53)	3.3	(0.50)
Vocabulary	51.3	(9.68)	51.1	(9.45)	49.5	(9.54)	49.7	(9.89)
Usage	48.7	(9.34)	49.9	(7.24)	49.1	(10.93)	48.4	(8.77)
Spelling	51.8	(8.05)	52.0	(10.22)	51.0	(9.11)	50.3	(10.09)
Reading	51.8	(11.13)	50.6	(10.04)	52.1	(9.18)	52.6	(9.04)
Verbal Composite	50.7	(9.64)	50.6	(9.25)	50.6	(9.97)	50.5	(9.10)

The departure from a purely random selection in order to match the groups may affect the way one interprets the results of analysis of individual measures. It means that essays of the total 18-year-old freshman population sampled here are likely to be slightly worse than our sample, and essays of the total 20-year-old junior population slightly better. Worse or better, of course, only in traits that correlate with the Verbal Composite score, which means with just a small minority of the 107 traits measured here. Technically, the non-randomness of the selections of the FR and JU sample possibly compromises the procedures used for testing statistical differences between group means, since such procedures require independence of observations. Most statisticians, however, will discount this possibility because the changes entailed by the matching of groups are so slight.

The matching of the groups was needed to counter a much more serious compromise, threatened by the tendency of verbally weak students to drop out of college. This means a purely random selection of students (such as the first drawing) may show a spurious improvement in writing skill with age and educational experience--spurious because it looks like maturation has caused the

improvement whereas only a shift in population may have. Longitudinal studies are especially vulnerable to changes in population. Maimon and Nodine (1979) ended with 57% of their original sample; Kerek, Daiker, and Morenberg (1980) with 48%; Jewell, Cowley, and Rhum (1969), working with matched pairs, with 20%; Eblen (1981) with only 6%. But cross-sectional studies such as the present one, which appear to counter this attrition by collecting at one point in time random samples from different age groups, may still suffer from it. Stewart (1978), for instance, gave O'Donnell's Aluminum passage to freshman, sophomore, junior, and graduate students in education. But the trends he found in syntax may be attributable only to an increase in the intrinsic verbal-ability level of the groups. Of course the study here does not perfectly control the effects of this attrition, since of all those pre-college students performing at any one level on the WSPC, say at a Verbal Composite of 54, the less verbally skilled may tend not to make it to the junior level in college. The matching of groups by initial verbal skill is imperfect to the degree that this particular diagnostic examination is imperfect. For a thorough discussion of the problems in inferring ontogenetic change from both longitudinal and cross-sectional studies, see Rest, pp. 106-145, and Wohlwill, pp. 122-157.

Another compromising factor was the differences among the groups in chosen academic field. Percentage breakdown by declared major:

	FR	EF	SO	JU
Agriculture	8.7%	0.6%	3.6%	6.3%
Business and economics	30.4%	28.6%	50.0%	31.3%
Engineering and technical	21.7%	28.6%	14.3%	12.5%
Education	13.0%	4.8%	7.1%	9.4%
Humanities and social sciences	17.4%	23.0%	14.3%	31.3%
Biological and physical sciences	8.7%	14.3%	10.7%	9.4%
	100.0%	100.0%	100.0%	100.0%

The only anomalies are the high proportion of business and economics majors in SO and of humanities and social science majors in JU. The actual percent of the latter in FR and SO may be higher than indicated here, since undeclared majors represented 28.1% of FR, 34.4% of EF, 12.5% of SO, and 0.0% of JU, and (according to the university curriculum office) humanities and social science students tend to declare major later than business students do. Nevertheless, influence of chosen academic field is a factor possibly uncircumvented by this study.

The EF group cannot be compared with the other three student groups as securely as they can with each other. So as not to forget this, I have isolated the EF results in the main table of findings below. The EF group may represent a somewhat higher calibre of performance than the other groups do, in part--predictably--because there were more no-shows at the end-of-the-semester testing, and in part--unfortunately--because the students were selected from only nine sections, increasing the chances of undue influence of one or two teachers. Actually, hard evidence for a different calibre of student emerges in only one place, the holistic rating of their essays (see Measure 1, below). Their mean score was 34.5, a 24% increase on the first-week freshman essays and beyond the 6%-20% range reported by research into the effect of freshman composition on students (see Bailey, Brengle, and Smith, 1980; Davis, 1979; Faigley, 1979; Gorrell, 1983; Grabor, Hines, and Miranda, 1974; Meyer, 1982; Morenberg, Daiker, and Kerek, 1978; Witte and Faigley, 1981). I would like to argue that we are seeing simply the effects of an unusually good freshman-writing program, but in truth it operated quite traditionally, with readings from a

moderate, comprehensive rhetoric by Hans Guth and an anthology of essays by very sophisticated writers like Joan Didion and James Baldwin, a required writing assignment of around 400 words a week including three in-class essays and an optional research paper, and an instructor corps composed mainly of teaching assistants (92%).

The EF group certainly cannot have differed much in initial writing ability, academic interest, or course performance, since not only their distribution of pre-college verbal scores and of declared major matched that of other groups, but their distribution of final grades turned out almost exactly that of the FR students. The safest explanation for the jump in holstic--other than the instructional effect of the course itself--may be simply that the EF essays differed so much in writing occasion. Striving for a grade at the end of a course cannot equate with providing a diagnostic at the beginning, with reasons easily imagined for the one arising out of a superordinate effort and for the other out of something less, even something deliberately less. At the end of the semester students wrote essays averaging 433 words, a full third more than at the beginning, more even than the post-graduation writers did (see Measure 2 below).

Besides the enormous jump in production rate, evidence for a significantly different attitude appears in the distribution of holistic ratings by initial writing competence. If the FR and EF writers are grouped by high-school verbal-composite score, it becomes obvious that the lower-ability students advanced the most during the course.

GROUPING	FR		EF		DIFFERENCE
Verbal Composite Score	N	Holistic $\bar{X}$	N	Holistic $\bar{X}$	FR $\bar{X}$ from EF $\bar{X}$
31-40	6	17.17	6	24.33	+ 7.16
41-50	9	22.44	9	36.78	+14.34
51-60	12	30.83	13	35.92	+ 5.09
61-70	5	42.40	4	40.00	- 2.40

(The holistic measure here is Measure 1 below.) In their study of freshmen at the University of Michigan, Bailey, Brengle, and Smith found the same phenomenon of students who performed worst at the beginning of the course progressing the most during it.

The 32 subjects in the post-college group (hereafter labelled as MA) were selected by a different standard of writing ability. The supervisors who chose them were asked not to select "the best writer we have," but rather "a competent writer," one who "handles extended writing assignments on the job with no problems."

The expertise of the 32 subjects thus chosen embrace a broad variety of occupational areas. Here is the complete inventory:

accounts manager (card company)
administrative secretary (advertising firm)
administrator (research firm)
agent (insurance)
assistant director (nursing)
architect
commercial artist
coordinator (youth services)
contract manager
development manager (social services)
executive secretary (arbitration firm)
field sales manager (signs)
finance manager
flight attendant
geologist
human resources specialist
legal assistant
loan representative (bank)

marketing representative
office assistant
(conference office)
orientation officer
personnel director
(hotel)
personnel representative
(communications firm)
project engineer
research scientist
resource manager
(lumber company)
secretary
(hospital)
superintendant
(recreation department)
technical editor
technologist
(radiology)
transcriptionist, senior
(insurance firm)
travel agent

On the job, this group performed, according to its own report, professional writing activities ranging from the composition of memos, reports, correspondence, contracts, proposals, advertising, and publicity, to cleaning up the writing of superiors. The ages of the MA group ranged from 31 to 58 with a mean of 38.8 and a median of 37 (S.D. = 7.26). Of course, no conclusions can be made about writing maturation from a comparison of this group with the student groups. The MA writing serves as a terminus ad quem, a mark by which to judge the development inferred from the underclass writing. As such, the MA group was intended to function as a more reasonable criterion than the kind of professional writers sometimes selected for that purpose, for instance essayists published in national slicks like Atlantic Monthly, Harper's, or The Saturday Review (e.g., Hagen, 1971; Gebhard, 1978). The workaday "competency" of the MA writers here may serve as a more informative comparison for undergraduate writers.

III. Procedure of analysis

PROCEDURE OF ANALYSIS. Analysis was performed on copies of the essays, typed exactly as written except for erasure of name and age of writer. For

15 of the measures, chiefly subjective ones, essays were analyzed by a team of 14 English-department teaching assistants, with an average of 2.5 years experience teaching college and 2.2 years high-school composition. For each measure they trained as a group, then worked in pairs analyzing a small number of essays (usually 15), first independently and then in collaboration for adjustment. A record of both independent and adjusted scoring was kept, the first to calculate inter-rater reliability. The report of this reliability figure in the main table of findings below will identify these group-analyzed measures. Every effort was made toward precision of measurement, and many times results were discarded on evidence of inaccuracy or unreliability. I executed all other measures.

IV. Rationale for choice of measure

RATIONAL FOR CHOICE OF MEASURE. Finding the best measurement for a particular writing trait is a complex problem, with at best arbitrary procedures. I will list here the rules followed in this study. (The numbering system used hereafter to refer to particular measures--as in "M30" or "Measure 30"--follows that below in the next section and in the main table of findings in section VII.)

(1) Measures were designed to apply to the entire essay. For computation, the easiest method would have been to analyze samples of the same size, say 30 T-units or 300 words. But the method is also easily misleading. With one writer 30 T-units or 300 words might include the end of the essay, where T-units in fact tend to be shorter than at the beginning, thereby benefitting another writer with a long essay were length of T-unit under measurement.

Of all the measures, only M8 did not count the entire essay. Except for measures of size (M2, 5, 7, 8, 21, 23, 25, 53-55, 68) or variance (M91), most traits then needed to be expressed as a rate of some sort.

(2) As the ideal denominator for a rate I took, when I could, the once-only opportunity. So correctness of spelling (M92) is expressed as a percent of total words that are spelled correctly, because each word used presents one and only one opportunity to be spelled right. M42 does not count number of passives per T-unit, say, because some writers fill their T-units with more verbs than other writers do. Rather M42 measures the number of passive formations per opportunity to make a passive, that is, the percent of those predicate forms which allow passivization (main verbs, present participles, infinitives, etc.) that have been made passive. The advantage of a once-only opportunity rate is that three different ways to express rate can be calculated from one measurement: percent, frequency, and proportion. If in writing 100 verbs a writer makes 20 of them passive, then we can say 20% of the verbs are passive (percent), or the passive rate is .2 or 1 every 5 verbs (frequency), or of verbs able to be made passive 20% have been and 80% have not.

Unfortunately, once-only opportunity rates are often difficult to construct. M42 is not truly one, since even a predicate in the right grammatical form for passivization will not offer the opportunity if it is intransitive. M42 is a valid rate only if the transitive/intransitive verb ratio remains constant among groups being measured. And much of the creativity of language simply cannot be confined to a once-only opportunity. One noun offers the opportunity to be modified a limitless number of times. A sentence, a T-unit, even a phrase offers more than one opportunity for expansion by syntactic

parallelism. And conversely not every word can be modified or made parallel (e.g., articles) in the way that every word can be spelled right. In these multiple-opportunity traits, expressions of rate by percent, frequency, and proportion are not necessarily synonymous. If we are measuring free modification per T-unit and a writer constructs 20 free modifiers in a 40 T-unit essay, then the frequency rate is .5 or 1 free modifier every 2 T-units. But this does not mean 50% of the T-units have free modification, since some T-units very well may contain more than one free modifier. Similarly, the proportion of T-units with free modifier to T-units without will not necessarily be 50% to 50%. Two essays could both have the same number of free modifiers in the same number of words yet one a higher-free-modifier rate per T-unit if its T-units are longer. The main problem with any multiple-opportunity trait is that the denominator, no matter which one is picked, will remain undefined. As a result, choice of a measure for multiple-opportunity traits is finally a choice among evils.

(3) I chose incidents per word for multiple-opportunity traits when I judged the length or complexity of an expression of the trait to be relatively unimportant. In M85, whether an instance of parallelism takes up 3 or 15 words seems less useful to know than that the writer made that choice of syntactic form at all. Traits measured this way (M10, 11, 26-31, *48, 85) are expressed as a frequency rate because a percent figure would be meaningless, the numerator (incidents) not representing a subset of the denominator (words). The word was chosen as denominator to eliminate the variability among writers of larger units, such as clause, T-unit, or sentence. This choice occasionally made a considerable difference. Compared to students, MA writers have a lower rate of non-additive logical connectors per word (M27) and a higher rate per

T-unit. (This is because MA writers write longer T-units.)

(4) When I felt it useful to know the relative length of an expression of a multiple-opportunity trait, I measured it by percent of total words in the essay embraced by structures of the trait. Percent of words in free modification will not tell us the number of incidents of free modification in an essay. But an incidence rate also may mislead as to what portion of the writer's ideas are given over to the rhetorical purposes served by free modification. The free modifiers, though numerous, could be brief and perfunctory. With a words-in-structure rate, proportions may still be calculated. M71-73 show us, for instance, that 29% of FR words in free modification were placed at the terminal position, contrasting with 41% of MA words. But I avoided couching traits only in the form of such intra-class proportions. Fifty percent of a writer's free modification may occur in the final position when in fact only two words of the total essay function as free modifiers.

(5) For this reason, I avoid expressing traits as ratios--e.g., the ratio of verbs to adjectives. Exceptions are M39-41, 56, 70, and *88, where other measures help ground these more abstract, relational figures.

(6) In *post-facto analysis occasionally a computational shortcut was taken. The means for group rates were determined by summing individual scores and dividing the sum, instead of by averaging individual rates, as is done in all the non-parametric pre-set measures. This shortcut will always be identified as "whole-group" rate. Usually there will be little difference in the resulting group mean. Substantial difference may occur when there is large variance in individual scores or a large portion of zeros.

(7) As already noted, when a quarter or more of a group records zeros for

a trait, I relegate the measurement to a nominal level. That is, I record only whether the trait is present or not in an essay, and consequently measure group performance by percent of writers using the trait at all (e.g., M17, 18, 32, 45, 46, 84).

In the present report, 107 measures are presented. Of these, 84 were pre-set, determined on before analysis of the essays was initiated. The remaining 23 measures were developed after analysis began to help interpret the pre-set measures or to fill in gaps that the analysis rendered important. These 23 post-facto measures, of course, contravene the assumptions of hypothesis testing, and throughout this report they are flagged with an asterisk. No inferential statistical tests were run on them.

V. Description of measures

DESCRIPTION OF MEASURES. The following arrangement of the 107 measures is intended solely for easy reference. The category of "Organization" may be the only one to counter tradition or intuition. By it I mean roughly any assistance to the reader in conveying relationships of ideas beyond that assistance conveyed by the semantic base. Both "organization" and "coherence" then serve the same purpose. (For confirmation of this point, see Alan C. Purves' factoring of reader response, p. 431.)

A.	Overall Quality	Measure 1
B.	Ideas	Measures 2-11
C.	Support	Measures 11-19
D.	Organization	Measures 20-36
E.	Diction	Measures 37-52
F.	Syntax	Measures 53-91
G.	Mechanics	Measures 92-107

A. Overall Quality

1: Holistic evaluation. Sum of 7 independent impressionistic ratings on a scale of 1 (low) to 8 (high). College English teachers (see above under Procedures of Analysis) were trained to work with boundary essays marking the bottom limit of categories 3, 5, and 7. Essays were read and rated at the rate of 10 an hour, with at least an hour break between batches. No particular distribution of scores was set or encouraged. Method of computing inter-rater reliability was Winer's adjusted pooled-rater (p. 130).

B. Ideas

2: Essay length. Number of words in essay, excluding title. Word-counting procedure followed Kerek, Daiker, and Morenberg (1980, p. 1121), except that abbreviations, phrasal proper nouns, numbers, and dates were converted to conventional oral expression. So USA was counted as 4 words, 1972 as 3 words, TV as 1 word.

3: Rhetorical mode. Percent of writers in research group who expressed rhetorical evaluation or recommendation, as defined by Eckhardt and Stewart (1979). Evaluation makes an assertion of value (It is sad to think that beauty plays such a major role in all our lives), recommendation an assertion of policy (each person should be allowed to establish their own right and wrong). Both proceed beyond the simpler levels of definition and substantiation, which clarify or support only the factual existence of the subject. An assertion of recommendation of course assumes one of evaluation too. The measure, then, counts those writers who took their rhetorical purpose beyond mere factual reporting to that of value judgment or recommendation for action.

4: Top-level logical pattern. A number from 1 to 14 assigning the essay's

most inclusive, logically coherent organizational scheme to one of the following rank-ordered patterns:

A. Partition:

(1) collection, where categories overlap

(2) classification, where categories are disjunctive

B. Seriation:

(3) degree, simple ranking by one variable

(4) development, chronological staging

(5) comparison, comparison or contrast

C. Consequence:

(6) causation, cause and effect, not a causal train

(7) process, stages toward a planned goal

D. Argument:

(8) inference, enthymemic or inductive, division into premise(s) and conclusion

(9) choice, division into pro/con and choice

(10) solution, division into problem and solution

(11) dialectic, division into conflict (thesis/antithesis) and synthesis.

E. Chaining:

(12) causal chain, where the effect(s) of one causation forms the cause of another causation

(13) sorites, where the conclusion of one inference forms the premise to another inference

(14) sequence, where the final part of one pattern forms the first part of a different pattern.

Other patterns (for instance, chaining of process, choice, solution, and dia-

lectic) are possible, but did not appear in the essays.

An "organizational scheme" is defined as a logically coherent arrangement of ideas in the body of an essay, that is, excluding introductory and concluding material. "Logically coherent" excludes sublogical arrangements such as fantasy, dream, free association, chance, and non-thematic narrative (what D'Angelo calls "non-logical" patterns).

"Top-level pattern" is simply the one unitary pattern (a scheme of support backing a claim or main "point") that embraces the largest number of words in the essay. When two patterns embrace the same words or the same number of words, the higher ranking pattern was selected. This rank order is designed to move from simplest to most complex patterns according to the relative complexity of constituent parts that make up the pattern. For instance, in the inferential patterns, choice entails not only at least one inference but also comparison, solution entails not only at least one choice but also an implied test of the results, and dialectic adds to the notion of solution that of mediation. The taxonomy--which is described fully elsewhere (Haswell, 1986b)--does not essentially challenge existing theories of cognitive growth (Inhelder and Piaget, 1958; Kalusmeier, Ghatala, and Frayer, 1974) or of compositional development (Britton, et alii, 1975; Hillocks, 1982). Existing classifications of logical patterns in writing posit similar, although implicit, hierarchies (Pitkin, 1969; D'Angelo, 1975; Longacre, 1979). For a breakdown of findings, see Table 1 below.

5: Length of top pattern. Number of words contained in the top-level pattern. See Table 2.

6: Second-level pattern. Rank of the organization scheme of the largest logical part of the top-level pattern. Assignment of rank to scheme was

identical to that in M4. The "logical parts" of each pattern are defined in the classification provided above under M4.

7: Length of second-level pattern. Number of words contained in the second-level pattern. See Table 2.

8: Logical depth. The number of times the top-level pattern could be subdivided, continuing with the method of M6. The logical pattern of the largest part of the second-level pattern was determined, if there was one, and the pattern of the largest part of it determined, and so on until no further pattern could be found. This analysis usually continued down to within the T-unit. The following sentence, for instance, expresses a classification: These beauty aids range from make-up, used to change facial appearance, to girdles, used to change body appearance." Yet the largest part of the pattern (words 6 through 11) contains a causation. The essential concept of this measure is of a hierarchy of embedded logical relationships that extends across phrase, clause, sentence, paragraph, and other macrostructural boundaries (Longacre, 1979; Van Dijk). See Table 2.

9: Essay unity. Percent of essays in the research group whose entire word production, excluding introduction and conclusion, is embraced by the top-level pattern of M4.

10: Qualifiers. Expressions of qualification, per word. Four distinct types are here combined:

(a) four modals of probability: can, could, might, may;

(b) five predicates expressing uncertainty: seem, tend, feel, think, believe;

(c) twenty-four adjectival or adverbial expressions of probability, such as likely and on the whole;

(d) T-units asserting a condition of probability--e.g., There are exceptions to this rule.

For statistical breakdown of these four types, see Table 3.

C. Support

11: Exemplification. Instances of logical exemplification, per word. Counted were only examples in support or illustration of a more general assertion. An example did not have to be labelled as such by the writer (for instance). Not counted, however, were definitions, mottos, characteristic detail, arithmetic expressions of frequency, or analytic partition of a subject. To illustrate this last and most difficult distinction, the following sentence about lions was judged to contain no exemplification: Rules have been established to protect the pride, or to respect the elder members of the group.

12: Number of examples. Raw number of examples (as defined above) in an essay. This measure was included on the possibility that authors may have a stronger sense of their support by how many examples they end up providing rather than by how often they provide them.

13: Level-one generalizations. Percent of semantic heads of T-units that are ranked lowest in a four-tier classification of nominal abstraction.

To determine the intended meaning of the T-unit head, both context and syntax were used: expletives were replaced by their true subject, and pronouns by their antecedent, and in all cases restrictive modification taken into account. As the one exception to this procedure, in structures like I believe that . . . and I think that . . . the nominal head analyzed was the subject

of the subsequent that clause.

The four levels of nominal abstraction ranged from most specific or least inclusive (Level 1) to most general or inclusive (Level 4). For example:

Level 4	vegetation	society	physical appearance
Level 3	tree	people	clothing
Level 2	willow	teenagers	shirts
Level 1	the willows in town	my classmates	designer jeans

Levels 1, 2, and 3 respectively correspond to what Eleanor Rosch and her associates (1976) call subordinate, basic, and superordinate classes, or what Brent Berlin (1972) calls specific, generic, and life names. Level 4 is the abstract (as opposed to concrete) conception that subsumes the rest. A preliminary testing of this measure on other essays found that a count which encompassed all heads of clauses did not correlate any better with an independent impressionistic assessment of "support" than did this measure, which counts only T-unit heads.

14: Level-two generalizations. See M13.

15: Level-three generalizations. See M13.

16: Level-four abstractions. See M13.

17: Unique events. Percent of writers in research group who refer at least once in their essay to an event which they describe as occurring once only. The first day of my biology class . . . introduces a unique event, whereas In my biology class we have . . . does not. The event may be either personal or historical, but hypothetical events or episodes from an artistic work were not counted.

18: Allusions. Percent of writers in research group using one or more allusions in their essay: references to a person, place, or event of some historical

or literary fame.

*19: Learned allusion. Whole-group percent of allusions that are learned in nature rather than popular. No matter how contemporary, political figures were categorized as learned.

D. Organization

20: Introduction. Sum of a primary-trait rating of the introductory material in the essay--material which leads up to, states, defines, restricts, or qualifies the main statement of thesis. The four separate assessments of the rating are broken down and described in Table 4. The range of scores was 0 (low) to 11 (high).

21: Length of introduction. Number of words in the introduction.

22: Conclusion. Sum of a primary-trait rating of the concluding material. For a breakdown of the assessment, see Table 4. The range of scores was 0 (low) to 5 (high).

23: Length of conclusion. Number of words in the conclusion.

24: Paragraph linkage. Percent of paragraphs in the essay (disregarding the first paragraph) whose first sentence connects explicity with the previous paragraph. The linkage had to be obvious: a logical connector (So), a demonstrative (These rules), an answer to a previous question, a repetition of a key word or phrase, etc.

25: Paragraph length. Mean number of words per paragraph.

26: Logical indicators. Instances of certain explicit logical connectors, per word. This is not a measure of just inter-T-unit connections, but of connections

between any syntactive unit. For this measure, five different classes of connectors were summed. The classification is adapted from Mary Hiatt (1977):

Illustratives:	for example, for instance, such as, e.g., like, etc.
Illatives:	therefore, thus, so, hence, etc.
Adversatives:	but, however, yet, on the other hand, instead, etc.
Causatives:	because, so that, since, consequently, etc.
Additives:	and, also, moreover, too, not to mention, etc.

Conditional connectors (unless, if . . . then) were counted as causatives, sequential connectors (next, secondly) as additives, but indicators of temporal succession (after, then) were not counted at all, nor were those of concession (of course), analogy (similarly), probability (generally), intensity (even), etc. For a breakdown of the five classes counted, see Table 5.

27: Non-additive indicators. Instances per word of the first four categories of M26, summed.

28: Cohesive ties. Sum of instances of two lexical ties (identicals and synonyms) and two grammatical ties (of reference and substitution), per word. The analysis follows Halliday and Hasan, and cohesion within T-units was not counted. Search for antecedent was pursued only back through the paragraph to the last sentence of the previous paragraph. See M29-32.

29: Identical ties. Instances of identical ties, per word. Function words were not counted as identicals. Changes of form within other grammatical classes were counted, as in number for nouns (boy ↔ boys) or tense for verbs (play ↔ played).

30: Synonym ties. Instances of synonym ties, per word. Function words were excluded from the count. Synonymity included change from one grammatical class to another, as from noun to verb (a change ↔ changed) or adjective to noun (hard work ↔ work ethic).

31: Reference ties. Instances of reference ties, per word. The tie had to be anaphoric (with antecedent previous in the text to the referent) and endophoric (with antecedent established by the text). This second rule excludes most instances of the first and second person pronoun. Included are the other pronominals (he, its, theirs, etc.), demonstratives (this, those, that, then, etc.), the definite article the when used demonstratively, and comparatives (less, such, else, additional, etc.).

32: Substitution ties. Percent of essays in research group with at least one instance of a substitution tie. Substitutions are surrogate expressions for syntactic units: one, so, likewise, do, etc. E.g., And we do so [follow the fashion].

*33: Original title. Percent of essays in research group with a title not taken totally verbatim from the topic. For percent of essays with a title at all, see Table 4.

*34: Paragraphing for logical organization. Percent of times the beginning of a new logical part coincides with a new paragraph. The major boundaries are between body and introduction and conclusion, and between parts of the top-level logical pattern as determined by M4. This is a whole-group rate.

*35: Cohesion and logical organization. Percent of times a new logical part (see *M34) contains an explicit link to the previous part within its first T-unit. Linkage was defined as in M24. Whole-group rate.

*36: Five-paragraph theme. Percent of essays in research group that follows the often-taught format of five paragraphs consisting of introduction, first point, second point, third point, summary.

E. Diction

37: PLOM words. Percent of words in essay that are word types identified by Thorndyke and Lorge as occurring in print 10 times or less per million words. Seven words so identified but not counted were fad, teen, teenager, commercial, pants, television, and make-up. Repeats (tokens) and proper names were not counted. PLOM words do not need to be rarae aves. Examples from our essays are assets, baggy, capitalize, discomfort, emphasis, flare, gathering, hedonistic, intricate, juvenile, kernel, learner, morality, nuisance.

38: Long words. Percent of words in essay 9 or more letters long. The segments of hyphenated compounds were treated separately.

39: Noun adjuncts. Percent of pre-nominal adjectives that are noun adjuncts. The American Heritage Dictionary (1969) was consulted to distinguish between true noun adjuncts (dog collar), compound nouns (dog-ear), and simple nouns (dogfight). The last two were not counted here.

40: Elegant variation. Percent of identical and synonymous ties together that are synonymous ties. See M29 and 30.

41: Noun modification/predicate ratio. Percent of instances of nominal modification (M67) and of predicates together that are instances of nominal modification. Predicates counted were finite verbs, gerunds, present and past participles used adverbally, and infinitives except when used to modify a noun (a day to remember). This measure was constructed in the spirit of

D. P. Boder's well-known "adjective-verb quotient," but the method of counting adjectivization is quite distinct.

42: Passives. Percent of predicates which have been transformed to the passive, either with agent expressed (outstripped by the knowledge) or excised (we are told that). The following kinds of predicates were counted as subject to the passive transformation: main verb, gerund verbal complement (keeps being laughed at), nominal gerund (from being misunderstood), adverbial participles (when pressured by advertisements), post-nominal adjectival particples (to believe the ideas passed to us), infinitives in any form.

43: Process verbs. Percent of predicates that are process (non-stative). Counted were both finite and nonfinite predicates, but not participles functioning as adjectives in a non-complementary position. The test for stativeness was the traditional one of whether the verb in its context will take readily the progressive form or, in ambiguous cases, the test of Miller and Johnson-Laird (1976) of whether the verb would answer the question, "What happened?"

44: Pronominalization. Percent of references to humans or to human groups that are expressed as pronouns. Included under such references are nouns like people, advertisers, John Kennedy, the author, and elliptical expressions thereof like many [women], and pronouns referring to them like we, I, hers, them, one. Not included were the level-four abstractions of M16 like society and corporations, and pronouns referring to them like it or their.

45: First-person singular. Percent of essays in research group with first-person singular pronouns forming 25% or more of all human pronouns.

46: Second-person plural. Percent of essays in research group with the

second-person generic you forming 25% or more of all human pronouns.

*47: Longest words. Percent of words in essay 11 or more letters long. Segments of hyphenated compounds were analyzed separately.

*48: Predicates. Structures of predication, both finite and infinite, per word. Counted were the same verbal structures as in M41. This method of counting disregards verbal expansion for tense, mood, and voice.

*49: First-person singular subjects. Percent of T-units that have I as the grammatical subject.

*50: Psychological state. Percent of stative verbs (see M43) that describe psychological state rather than physical state. E.g., psychological: seem, feel, hope, consider, is [happy]; physical: range, differ, match, relate, is [there]. Whole-group rate.

*51: Possessives. Possessives, formed correctly or not, per word. Whole-group rate.

*52: Contractions. Contractions, formed correctly or not, per word. Whole-group rate.

F. Syntax

53: Sentence size. Mean length of sentences in words, as punctuated by the writer.

54: T-unit size. Mean length of T-units in words, as defined by Hunt (1965, pp. 20-21). A syntactically dependent structure punctuated by the writer as a separate sentence was counted as part of the T-unit to which it is obviously attached. A T-unit is a main clause along with any structure syntactically dependent on it. The end of a T-unit is that point at which a full stop could

be put legitimately (the "T" in Hunt's term stands for "minimal terminable").

55: Clause size. Mean length of clauses in words, as measured by Hunt (1965, p. 15). Sentence fragments were treated as in M54.

56: Clause/T-unit ratio. Number of clauses divided by number of T-units in the essay. See Hunt (1965, pp. 33-34).

57: Coordination of T-units. Percent of sentences, as punctuated by the writer, that are compounded of two or more T-units. This measure does not convert to Hunt's ratio of T-unit to sentence (1965, pp. 37-38). Here intra-sentence T-unit coordination was counted directly, not assumed from sentence and T-unit data. With Hunt's measure, every sentence fragment reduces T-unit coordination by one, whereas here that debilitation of the coordination rate is avoided.

58: Predicate coordination. Number of times predicates (finite and infinite) were coordinated, per clause. The measure indicates how often a writer elects to expand the predication of a clause by coordinating the predicate and retaining the same subject of the predicate (Let the codes be set down and maintained). Coordination of auxiliary verbs (might and should protest) was not counted. Each member of a coordinative after the first added one to the count.

59: Nominal coordination. Number of times noun structures were coordinated, per clause. Coordination of pre-nominal adjectives was not considered. Verbals like clausal and infinitival nominalizations were treated as noun structures and counted: What this code is, and why it exists will be discussed.

60: Subordinate clauses. Percent of total words occurring in subordinate clauses.

61: Relative clauses. Percent of total words occurring in relative clauses.

62: WH clauses. Percent of words occurring in noun clauses formed from ques-

tions or exclamations. The classification follows that of Marcella Frank, pp. 293-299.

63. That clauses. Percent of words occurring in noun clauses formed from statements (see Frank, pp. 286-292).

64: Base clauses. Percent of words occurring in main clauses of multi-clause T-units. Attachable sentence fragments were analyzed as part of the structure to which they depended syntactically.

65: Mono-clauses. Percent of words occurring in single-clause T-units.

66: Independent clauses. Percent of words occurring in main clauses. The figure merely sums M64 and M65.

67: Nominal modification. Percent of total words occurring in structures modifying nominal heads. Modification may be free or bound, but the count stopped with the initiation of a new T-unit. Clausal and verbal nominalizations were included as nominal heads, but the base verbal structure itself was not counted. In the following examples, material in brackets was counted as nominal modification:

Gerund: Lowering the [legal drinking] age [in the state of Washington] has been . . .

That nominal: That a [young] man could die [for his country] is . . .

WH nominal: What happens [when the two talk] is . . .

As these examples of nominalizations show, when adverbial modification occurs within a noun structure, it is counted as nominal modification. The phrase a trend [that changes annually] received a word count of 3. This method of analysis remains true to the spirit of the measure, which is semantic rather than strictly syntactic. The base requirement for an instance of nominal modification is that it restrict logically the class identified by the head nominal.

68: Nominal complexity. Number of different orders of complexity in strings of nominal modification. The measure adopts Hunt's term and basically his method of analysis (1965, pp. 114-115). Each nominal head was given a modification "complexity" count. A count of 1 was added for every instance of class restriction: expensive clothes received a count of 1, very expensive clothes a count of 2, expensive clothes discarded without thought a count of 3. Coordination of modifiers did not add to the count: men's intercollegiate athletics scored 2, but men's and women's athletics scored 1. Unlike in Hunt, nominalized verb and noun clauses were not awarded an initial count of 1. So the four examples cited in M67 received, respectively, counts of 4, 2 (his merely repeats young man), and 1. Note that the final measure was simply the number of different sized complexity counts. A writer with 17 strings of 2, 6 of 3, and 1 of 5 received a final score of 3. See Table 6 for a breakdown by size of string.

69: Post-nominal modification. Percent of total words occuring in modifying structures after the noun head. M69 subtracted from M68 gives the percent of words in the essay in pre-nominal modification.

70: Pre/post-nominal modification ratio. Percent of words in nominal modification that occur after the noun head.

71: Free modification. Percent of total words occurring in free modification. Punctuation of writer was followed unless it contradicted the sense of the passage. Then analysis followed context. In ambiguous instances, punctuation determined the analysis. In all major ways, the method of analysis followed Francis Christensen (1968).

72: Initial free modification. Percent of words occurring in free modification positioned before the grammatical subject of the T-unit

73: Medial free modification. Percent of words occurring in free modification positioned between the grammatical subject and the end of the T-unit.

74: Final free modification. Percent of words occurring in free modification occupying the last words of the T-unit.

75: Adverbs. Percent of total words occurring in structures modifying verbs or whole clauses.

76: Initial adverbs. Percent of words in adverbial structures placed before the grammatical subject of the T-unit.

77: Medial adverbs. Percent of words in adverbial structures placed anywhere after the subject and before the end of the T-unit.

78: Final adverbs. Percent of words in adverbial structures occurring as the final words of the T-unit.

79: Prepositions. Percent of total words in prepositional phrases. Prepositions inseparable from the verb (particles) were not counted, nor prepositions without objects used adverbially ("They want out"). Word count began with the preposition and ended with a new preposition or an unbound structure.

80: Prepositional strings. Percent of total words in prepositions composed of strings of two or more prepositional phrases. A prepositional string is defined as two or more contingent but non-coordinated prepositions. By complaining and in words counted as two prepositional phrases but not as a string. By taking a candy bar without paying is a string. The boundary between elements of a string may be free or loose syntactically.

81: Infinitives. Percent of total words in infinitival structures--adjectival,

adverbial, nominalized, or complementary to the predicate. Crypto-infinitives after verbs like help and make were also counted. Word count proceeded as with prepositions.

Adjectival: This never-ending compulsion [to be beautiful] . . .

Adverbial: [To be physically attractive], a man or woman must look . . .

Nominalized: It's alright for him [to drive while drunk].

Verb complement: No individual wants [to be out of fashion], not today.

For a breakdown of these four types by age-group, see Table 7.

82: Adverbial participles. Percent of research group using one or more participial constructions modifying clauses or predicates, either present participles (In [considering the question of beauty], we have to remember that . . .) or past participles (. . . no idea what to do [when presented with a tempting opportunity to deviate from ethical standards]).

83: Nominalized participles. Percent of research group using one or more participial constructions functioning either as verbal complements (The beautiful young must be seen [smoking the correct cigarettes]) or as gerunds (a good method for [learning what is right or wrong]).

84: Appositives. Percent of writers in research group using one or more appositives.

85: Syntactic parallelism. Incidence of intra-clausal syntactic parallelism, per word. The analysis follows Mary Hiatt (1975). Triple constructions, such as polysyndeton, were counted as one instance. Appositives were not counted. For further breakdown, see Table 10.

86: Dependent sentence openers. Percent of all types of sentence openers that are syntactically dependent on the grammatical subject. The measure shows what percent of openers are not noun structures, pronouns, nominalizations, or expletives. It discards from analysis sentences beginning with coordinate

conjunctions. See Table 8.

87: T-unit variance. Standard deviation of word-length of T-units.

*88: Subject coordination. Percent of nominal coordination that occurs before the main verb of a T-unit. Nominal coordination as a whole was measured by M59.

*89: Incidence of free modification. The frequency of free modifiers as the percent of independent clauses and free modifiers together. This is Variable A as devised by Broadhead, Berlin, and Broadhead. It gives a rate of free modification by instances rather than by number of words, as in M71-74. In counting instances of free modification, Broadhead, Berlin, and Broadhead's discrimination among 12 kinds of free modifiers was followed (p. 228). Whole-group rate.

*90: Size of dependent clauses. Difference between the mean independent clause length in words and the mean T-unit length in words. This is Broadhead, Berlin, and Broadhead's Variable Q, which they say "proves to be the best index of the amount, length, variability, and variety of free modifiers" (p. 237).

*91: T-unit range. Difference between the shortest and longest T-unit in words, in an essay.

G. Mechanics

92: Correct spelling. Percent of words correctly spelled. Alternate spellings given by The American Heritage Dictionary (1969) were accepted as correct. Every error was counted, even repeats, apostrophe mistakes, and apparent lapsus calami. Hyphenation and capitalization were not considered.

93: Correct possessives. Percent of writers in research group making no errors in the formation of possessives.

94: Correct contractions. Percent of writers in research group making no errors in the formation of contractions.

95: Correct predication. Percent of main verbs that express reasonable semantic agreement with their subject or complement. Faulty predication of dependent and non-finite verbs was not counted.

96: Pronoun agreement. Percent of writers in research group making no errors in agreement between a human pronoun and its antecedent (see M44). Shift from the neuter singular to plural (e.g., everybody . . . they) was counted as an error, as was shift from he/she or similar construction to he, she, or they. With relative pronouns, both shift in gender and case was counted as error.

97: Correct parallelism. Percent of instances of syntactic parallelism where the coordinated elements were of the same grammatical class.

98: Standard sentences. Percent of writers in research group forming no sentence fragments. Only dependent structures punctuated as a complete sentence and syntactically attachable to a previous or (rarely) following sentence were counted as fragments.

99: Punctuation of main clauses. Percent of writers in research group separating independent clauses with standard punctuation. Two independent clauses, even though short and parallel, separated by a comma or no punctuation, was counted as non-standard.

*100: Lapsus calami. Percent of misspelling resulting in the correct spelling of a word not intended, as in to be more confident in other says. Whole-group rate.

*101: Dropped suffixes. Percent of misspelling caused by omission of suffix. Whole-group rate.

*102: Possessive opportunity. Percent of possessives written that are formed correctly. Whole-group rate.

*103: Contraction opportunity. Percent of contractions written that are formed

correctly. Whole-group rate.

*104: Pronoun agreement opportunity. Percent of human pronouns (M44) that agree in number and gender with antecedent. Whole-group rate.

*105: Punctuation of initial free modifiers. Percent of initial free modifiers conventionally punctuated. Lack of terminal punctuation was not counted as an error if the modifier was five words or less in length. Whole-group rate.

*106: Punctuation of medial free modifiers. Percent of medial free modifiers conventionally punctuated. An instance was counted as error-free only if it had preceding and following punctuation of the same kind. Whole-group rate.

*107: Punctuation of final free modifiers. Percent of final free modifiers conventionally punctuated. Heading the modifier with a comma or a dash or enclosing it in parentheses was deemed conventional. Whole-group rate.

VI. Statistical procedures

STATISTICAL PROCEDURES. The basic test of differences in group means was a 5 by 2 by 2 fixed-effects ANOVA, with main effects for Group, Sex, and Topic. Confidence level of differences between groups was determined by Newman-Keuls multiple comparison (Kirk, 1968). Although the abnormal distributions that frequency data often produce have been shown to have little effect on fixed-effects models with N over 30 (Glass, Peckham, Sander), measures producing more than 25% zeros in any one group were calculated as a group-usage performance (that is, how many writers in the group used the form) and analyzed by chi-square tests for group and sex. On other rate data, arcsin transformations were performed when deemed appropriate (Lindman, pp.

31-34). Non-parametric measures were tested by Kruskal-Wallis one-way ANOVA separately for group and for sex. All of these computations were computer-run, formated by the Statistical Package for Social Sciences (see Nie, Hull, Jenkins, Steinbrenner, and Bent, 1975; Hull and Nie, 1979).

VII. Findings

FINDINGS. In the main table of findings, tests of significance with their probability figures for differences between the age-group means (group N = 32, df = 4) are indicated as follows: an F score records the 3-way parametric analysis of variance; KWχ^2 the chi-square statistic of the Kruskal-Wallis one-way analysis of variance by ranks; χ^2 the chi-square for independent samples. Probability estimates for posteriori comparisons between groups were made only on the parametric ANOVA. The same indications of statistical test hold for the gender and topic differences (group N = 80, df = 1), wher group means are given only where statistical significance ($p < .05$) was achieved. All differences between topics and all interactions on the 3-way ANOVA were non-significant except for the few recorded at the end of this main table.

Standard deviations, where appropriate, are in parentheses below the mean.

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
A. Overall Quality													
1 Holistic evaluation	.92	27.7 (11.6)	29.5 (12.0)	29.7 (9.5)	39.9 (11.3)	34.5 (10.3)	F 6.94 <.001	MA/FR,SO,JU <.001 EF/FR,MA <.05	30.0 (12.0)	34.3 (9.5)	F 4.86 <.05		
B. Ideas													
2 Essay length	.99	316.9 (118.9)	359.4 (136.4)	364.4 (131.2)	433.3 (152.3)	429.5 (120.6)	F 4.58 <.01	FR/EF,MA <.01	358.7 (120.2)	404.5 (129.7)	F 4.30 <.05		
3 Rhetorical mode	.89	93.8%	87.5%	90.6%	87.5%	87.5%	NS				NS		
4 Top-level logical pattern	.82	5.22 (2.96)	6.63 (3.73)	7.69 (4.28)	9.44 (3.65)	5.88 (3.70)	KWχ^2 20.8 <.001				NS		
5 Length of top pattern		208.8 (87.0)	257.9 (125.4)	266.6 (119.7)	333.5 (144.8)	308.1 (109.3)	F 5.39 <.001	FR/EF,MA <.001 SO/MA <.05			NS		
6 Second-level pattern		4.97 (2.80)	5.91 (3.50)	5.53 (3.40)	5.59 (3.82)	4.44 (3.59)	NS						
7 Length of 2nd-level pattern		102.2 (59.0)	132.5 (95.4)	128.5 (68.2)	225.6 (113.4)	195.6 (96.2)	F 10.64 <.001	MA/FR,SO,JU <.001			NS		
8 Logical depth		4.25 (1.44)	4.94 (1.60)	5.00 (1.10)	6.63 (1.72)	6.06 (1.72)	KWχ^2 35.2 <.001				NS		
9 Essay unity		65.6%	93.7%	90.6%	100%	96.9%	χ^2 40.6 <.001				NS		
10 Qualifiers	.91-.95	.015 (.010)	.017 (.011)	.017 (.009)	.012 (.010)	.014 (.009)	NS				NS		
C. Support													
11 Exemplification		.021 (.011)	.016 (.009)	.016 (.010)	.020 (.012)	.024 (.012)	F 3.82 <.01	EF/JU <.01 EF/SO <.05			NS	See below	
12 Number of Examples		6.9 (5.6)	5.5 (3.5)	5.5 (4.1)	8.7 (6.9)	10.2 (5.5)	F 5.12 <.01	EF/SO,JU <.01 FR/EF <.05			NS	See below	

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
13 Level-one generalizations	.86	10.0% (1.4)	7.8% (.9)	10.1% (1.1)	20.1% (2.1)	12.3% (1.7)	F 4.08 <.01	MA/SO <.01 MA/FR,EF,JU <.05			NS		
14 Level-two generalizations	.69	31.8% (1.9)	21.7% (1.5)	23.8% (1.5)	19.0% (1.5)	32.7% (1.8)	F 5.07 <.01	MA/FR,EF; EF,SO <.01 FR/SO; EF/JU <.05			NS	See below	
15 Level-three generalizations	.72	45.3% (1.7)	53.0% (2.2)	49.7% (1.8)	44.3% (2.1)	40.7% (1.9)	NS				NS		
16 Level-four generalizations	.79	12.9% (1.3)	17.5% (1.5)	16.5% (1.3)	16.6% (1.4)	14.2% (1.5)	NS				NS		
17 Unique events		6.3%	3.1%	3.1%	15.6%	12.5%	NS				NS		
18 Allusions		28.1%	25.0%	21.9%	37.5%	31.3%	NS				NS		
*19 Learned allusions		31.0%	46.7%	58.3%	66.7%	47.6%							
D. Organization													
20 Introduction	.78-.97	4.97 (2.40)	6.66 (2.38)	6.72 (2.61)	7.63 (2.55)	7.72 (1.78)	KW$\chi^2$25.2 <.001		6.39 (2.49)	7.09 (2.22)	KW$\chi^2$3.77 <.05		
21 Length of introduction		48.8 (45.2)	61.5 (34.4)	57.5 (34.1)	64.4 (49.7)	75.5 (47.1)	NS				NS		
22 Conclusion	.81-.99	2.5 (1.6)	3.5 (1.2)	3.3 (1.4)	3.6 (1.5)	3.2 (1.6)	KW$\chi^2$9.60 <.05				NS		
23 Length of conclusion		35.9 (26.0)	41.9 (20.8)	41.6 (27.0)	49.7 (36.3)	45.6 (28.2)	NS		38.7 (26.9)	47.2 (28.9)	F 4.20 <.05		
24 Paragraph linkage	.78-.86	56.7% (3.6)	60.9% (3.1)	56.4% (3.5)	95.4% (3.2)	68.6% (4.2)	F 7.78 <.001				NS		
25 Paragraph length		78.5 (30.4)	76.2 (36.9)	76.7 (23.7)	86.2 (30.8)	87.0 (28.5)	NS				NS		
26 Logical indicators		.064 (.017)	.060 (.016)	.055 (.012)	.056 (.015)	.057 (.014)	NS				NS		
27 Non-additive indicators		.030 (.014)	.029 (.015)	.025 (.012)	.021 (.010)	.027 (.009)	F 3.02 <.05	MA/FR,SO; FR/JU .05	.028 (.013)	.025 (.011)	F 4.11 <.05		

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
28 Cohesive ties		.124 (.030)	.122 (.030)	.125 (.031)	.104 (.033)	.130 (.027)	F 3.76 <.01	MA/FR,EF,SO,JU <.01	.125 (.026)	.116 (.028)	F 4.64 <.05		See below
29 Identical ties		.063 (.031)	.058 (.025)	.057 (.021)	.051 (.026)	.061 (.027)	NS		.063 (.025)	.053 (.026)	F 6.28 <.05		
30 Synonym ties		.024 (.014)	.029 (.016)	.031 (.014)	.022 (.010)	.026 (.012)	NS				NS		
31 Reference ties		.035 (.019)	.033 (.014)	.033 (.014)	.025 (.012)	.038 (.013)	F 4.17 <.01	MA/FR,EF,SO,JU <.01					See below
32 Substitution ties		34.4%	43.8%	56.3%	34.4%	50.0%	NS				NS		
*33 Original title		9.4%	34.4%	12.5%	25.0%	12.5%							
*34 Paragraphing for log. org.		65.4%	81.6%	79.0%	68.0%	81.7%							
*35 Cohesion and logical org.		40.2%	55.0%	50.5%	55.0%	45.9%							
*36 Five-paragraph theme		9.4%	21.9%	6.3%	3.1%	25.0%							
E. Diction													
37 PLOM words		2.0% (1.5)	3.2% (2.2)	2.6% (1.5)	4.4% (2.7)	2.2% (.9)	F 8.79 <.001	MA/FR,EF,SO,JU <.001 FR/SO <.05			NS		
38 Long words		6.2% (2.1)	7.9% (3.5)	7.5% (2.7)	9.9% (3.5)	11.9% (2.8)	F 19.29 <.001	EF/MA <.001 FR/SO <.05 EF MA/FR,SO,JU <.05			NS		
39 Noun adjuncts		13.2% (14.2)	12.5% (11.0)	17.9% (14.2)	21.5% (17.6)	12.2% (7.1)	F 3.07 <.05	MA/FR,EF,SO <.05			NS		
40 Elegant variation		29.1% (12.0)	33.8% (15.0)	35.5% (11.9)	32.5% (13.9)	31.7% (13.2)	NS		29.7% (13.3)	34.9% (12.9)	F 7.03 <.01		
41 Noun modif./ predicate ratio		55.9% (32.6)	58.0% (28.8)	59.6% (29.6)	62.0% (26.6)	55.0% (26.4)	F 3.46 <.01	MA/FR <.01 MA/EF <.05			NS		

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
42 Passives	.93-.97	10.4% (5.7)	15.4% (9.4)	11.1% (7.9)	13.5% (8.3)	13.9% (8.6)	NS				NS	See below	
43 Process verbs		55.1% (11.3)	51.9% (9.2)	55.4% (12.2)	53.5% (9.2)	46.6% (13.3)	F 4.15 <.05	EF/FR,SO,JU,MA <.05			NS	See below	
44 Pronominalization		49.6% (7.5)	51.2% (1.9)	50.8% (3.5)	57.3% (17.3)	52.9% (7.5)	F 2.57 <.05	MA/FR,SO <.05			NS		
45 First-person singular		9.4%	3.1%	25.0%	25.0%	9.4%	χ^2 10.46 <.05				NS		
46 Second-person plural		28.1%	6.3%	12.5%	3.1%	3.1%	χ^2 14.86 <.01				NS		
*47 Longest words		1.4%	2.1%	2 0%	2.8%	4.7%							
*48 Predicates		.147	.143	.144	.135	.145							
*49 First-person sing. subjects		5.4%	4.0%	8.1%	10.5%	4.1%							
*50 Psychological state		27.8%	25.3%	28.6%	26.4%	25.3%							
*51 Possessives		.005	.005	.005	.005	.004							
*52 Cont actions		.009	.005	.010	.004	.007							
F. Syntax													
53 Sentence size		17.2 (3.3)	16.6 (3.2)	18.6 (4.3)	19.8 (3.6)	16.8 (2.5)	F 5.07 <.001	MA/FR,SO,EF <.001 EF/JU; SO/JU .05			NS		
54 T-unit size		15.6 (3.5)	15.1 (2.7)	15.9 (2.7)	18.3 (3.8)	14.3 (2.3)	F 7.97 <.001	MA/FR,EF,SO <.001 MA/JU <.05			NS		
55 Clause size		9.1 (1.8)	9.4 (1.7)	9.1 (1.3)	10.2 (1.5)	8.7 (1.5)	F 4.04 <.01	MA/EF <.01 MA/FR,SO,JU <.05	9.6 (1.8)	9.0 (1.3)	F 5.57 <.05		

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
56 Clause/T-unit ratio		1.72 (.27)	1.62 (.25)	1.76 (.25)	1.82 (.39)	1.67 (.27)	NS				NS		
57 Coordination of T-units		14.5% (16.2)	14.4% (13.7)	18.2% (19.6)	13.2% (11.0)	21.1% (14.6)	NS				NS		
58 Predicate coordination		.059 (.063)	.065 (.074)	.060 (.049)	.056 (.057)	.054 (.046)	NS				NS		
59 Nominal coordination		.110 (.084)	.089 (.073)	.061 (.070)	.184 (.104)	.099 (.052)	F 10.83 <.001	MA/FR,EF,SO,JU <.001 FR/JU <.05			NS		
60 Subordinate clauses		13.4% (7.9)	12.2% (7.8)	12.6% (6.5)	11.3% (7.0)	12.3% (5.5)	NS				NS		
61 Relative clauses		7.6% (4.9)	7.5% (5.0)	10.1% (7.1)	10.0% (5.9)	6.7% (4.4)	F 2.58 <.05	EF/JU <.05			NS		
62 WH clauses		3.6% (3.2)	2.5% (2.3)	3.2% (3.7)	4.6% (4.5)	5.1% (4.5)	F 2.48 <.05	EF/SO <.05			NS		
63 That clauses		8.8% (5.9)	8.0% (5.6)	8.4% (6.0)	10.3% (8.9)	7.9% (6.0)	NS				NS		
64 Base clauses		27.7% (8.8)	24.9% (11.0)	29.1% (10.2)	25.6% (9.2)	27.3% (9.2)	NS		25.8% (9.5)	28.0% (9.0)	F 4.58 <.05		See below
65 Mono-clauses		39.0% (15.0)	45.0% (19.0)	36.6% (16.6)	38.3% (19.7)	40.8% (12.6)	NS		45.2% (16.9)	37.7% (15.4)	F 4.63 <.05		See below
66 Independent clauses		66.7% (9.2)	69.8% (10.5)	65.8% (10.2)	63.9% (14.6)	68.2% (7.7)	NS				NS		See below
67 Nominal modification		31.9% (10.0)	34.6% (9.4)	37.6% (10.3)	41.7% (8.3)	31.0% (7.2)	F 7.48 <.001	MA/FR,EF,SO <.001 JU/FR,EF <.05			NS		See below
68 Nominal complexity		4.63 (1.29)	5.50 (1.63)	5.53 (1.27)	6.59 (1.46)	5.94 (1.23)	KWχ^2 30.0 <.001				NS		
69 Post-nominal modification		22.4% (9.4)	25.5% (9.0)	29.0% (10.7)	33.3% (8.6)	22.0% (6.9)	F 8.93 <.001	MA/FR,EF,SO; FR/JU <.01 JU/EF,SO <.05			NS		

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
70 Pre/Post-nom. modif. ratio		68.3% (12.2)	72.5% (10.0)	75.4% (9.4)	79.1% (6.8)	70.0% (8.9)	F 6.79 <.001	MA/FR,EF <.001 MA/SO; FR/JU <.05	71.1% (9.9)	75.1% (8.9)	F 7.37 <.01	See below	
71 Free modification		19.2% (7.6)	21.9% (7.1)	19.0% (7.9)	26.2% (8.0)	16.9% (5.6)	F 7.39 <.001	MA/FR,EF,SO,JU <.001			NS		
72 Initial free modification		11.5% (7.2)	10.9% (6.2)	8.6% (4.8)	10.8% (5.7)	9.0% (4.8)	NS				NS		
73 Medial free modification		2.2% (2.2)	2.0% (2.4)	2.2% (1.9)	4.6% (3.5)	2.2% (2.0)	F 6.01 <.001	MA/FR,EF,SO,JU <.001			NS		
74 Final free modification		5.5% (5.7)	9.0% (4.9)	8.2% (5.9)	10.8% (6.2)	5.7% (4.0)	F 5.38 <.001	MA/FR,EF <.001 SO/FR,EF <.05			NS		
75 Adverbs		39.8% (6.7)	41.3% (7.5)	43.1% (7.9)	40.1% (8.1)	37.3% (8.1)	F 2.48 <.05	EF/JU <.05			NS		
76 Initial adverbs		11.2% (6.8)	10.8% (6.2)	8.6% (4.9)	10.1% (5.7)	10.3% (4.9)	NS				NS		
77 Medial adverbs		7.5% (4.3)	8.5% (3.7)	8.2% (4.0)	6.4% (3.0)	6.0% (2.0)	F 2.95 <.05	MA/SO,JU; EF/SO,JU <.05			NS		
78 Final adverbs		21.1% (7.5)	22.1% (6.3)	26.3% (8.4)	23.6% (6.8)	21.0% (5.2)	F 3.12 <.05	JU/FR,EF,SO <.05			NS		
79 Prepositions		32.4% (10.1)	30.9% (8.4)	33.8% (6.9)	39.8% (7.5)	37.6% (9.3)	F 6.42 <.001	MA/FR,SO,JU <.001 EF/SO <.01	36.4% (8.4)	33.4% (8.5)	F 5.28 <.05		
80 Prepositional strings		11.5% (8.4)	11.1% (6.2)	13.1% (6.9)	20.5% (7.4)	12.6% (5.1)	F 10.91 <.001	MA/FR,EF,SO,JU <.001	15.5% (7.1)	12.0% (6.2)	F 10.74 <.001		
81 Infinitives		12.2% (6.1)	11.9% (6.2)	12.0% (6.4)	11.1% (5.3)	10.4% (5.4)	NS		11.0% (5.1)	12.8% (6.6)	F 4.18 <.05		
82 Adverbial participles		59.4%	75.0%	71.9%	87.5%	71.9%	NS				NS		
83 Nominalized participles		69.7%	90.6%	87.5%	81.2%	81.2%	NS				NS		

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
84 Appositives		21.9%	40.6%	40.6%	75.0%	46.9%	χ^2 19.01 <.01				NS		
85 Syntactic parallelism	.83-.92	.040 (.015)	.039 (.017)	.033 (.012)	.041 (.011)	.034 (.007)	F 2.58 <.05	FR/JU <.05			NS	See below	
86 Dependent sent. openers		35.9% (4.3)	40.5% (9.2)	30.8% (6.4)	37.4% (10.6)	35.0% (5.1)	NS				NS		
87 T-unit variance		7.40 (1.82)	7.07 (1.94)	6.81 (1.79)	9.21 (2.77)	6.54 (1.68)	F 6.14 <.001	MA/FR,EF,SO,JU <.001			NS		
*88 Subject coordination		46.6%	55.6%	54.9%	28.8%	16.0%							
*89 Incidence of free modific.		34.4%	35.6%	34.4%	39.6%	32.8%							
*90 Size of dependent clauses		5.34 (2.26)	4.72 (2.12)	5.58 (2.01)	7.09 (3.98)	4.37 (2.02							
*91 T-unit range		27.22 (8.33)	27.66 (9.02)	26.03 (7.70)	34.63 (10.09)	26.59 (8.68							
G. Mechanics													
92 Correct spelling		97.9% (2.1)	98.2% (2.1)	98.3% (1.7)	99.4% (0.8)	98.3% (1.4)	F 3.76 <.01	MA/FR,EF,SO,JU <.01	98.0% (1.8)	98.8% (1.4)	F 8.71 <.01		
93 Correct possessives		59.4%	59.4%	50.0%	78.1%	53.1%	NS				NS		
94 Correct contractions		84.4%	87.5%	78.1%	96.9%	84.4%	NS				NS		
95 Correct predication	.57-.78	93.6% (4.9)	92.7% (5.1)	93.5% (5.0)	97.9% (2.7)	92.1% (5.2)	F 7.80 <.001	MA/FR,EF,SO,JU <.001			NS		
96 Pronoun agreement		37.5%	41.6%	47.0%	65.6%	37.5%	NS				NS		
97 Correct parallelism	.66-.94	90.1% (10.0)	87.6% (14.2)	90.6% (8.3)	98.6% (3.6)	91.8% (8.0)	F 5.93 <.001	MA/FR,EF,SO,JU <.001			NS		

DEPENDENT VARIABLE		GROUP							SEX			TOPIC	INTER-ACTION
Measure	Reliability	FR mean	SO mean	JU mean	MA mean	EF mean	Test of signif.	Posteriori comparisons	Male mean	Female mean	Test of signif.		
98 Standard sentences		59.4%	62.5%	81.2%	84.4%	65.6%	NS				NS		
99 Punctuation of main clauses		75.0%	59.4%	71.9%	78.1%	53.1%	NS				NS		
*100 Lapsus calami		22.2%	34.4%	36.5%	49.3%	17.3%							
*101 Dropped suffixes		5.9%	7.3%	11.3%	24.6%	8.9%							
*102 Possessive opportunity		66.7%	54.4%	54.5%	81.2%	66.7%							
*103 Contraction opportunity		94.4%	88.7%	87.8%	98.2%	85.3%							
*104 Pronoun agreement opport.		93.9%	93.0%	95.4%	97.7%	93.8%							
*105 Punctuation of initial free mod.		83.9%	82.1%	90.9%	92.1%	89.5%							
*106 Punctuation of medial free mod.		95.8%	75.5%	82.4%	85.3%	89.5%							
*107 Punctuation of final free mod.		64.1%	62.1%	64.9%	77.8%	60.6%							

Significant effects for Topic (Topic A, "Standards of beauty," N = 80; Topic B, "Codes of conduct," N = 80):

M11, Exemplification: F 4.51, <.05; A, .021; B, .018
M12, Number of examples: F 3.94, <.05; A, 8.0; B, 6.6
M14, Level-two generalizations: F 5.87, <.05; A, 28.5%; B, 22.4%
M42, Passives: F 8.00, <.001; A, 10.8%; B, 15.0%
M43, Process verbs: F 10.63, <.001; A, 20.1%; B, 7.1%
M70, Pre/Post-nominal modification ratio: F 7.37, <.01; A, 71.1%; B, 75.1%
M85, Syntactic parallelism: F 4.30, <.05; A, .040; B, .035

Significant interactions (N of cells = 16):

Sex and Group:

M28	Cohesive ties	F 2.84 <.05	Male:	FR .121	SO .133	JU .140	MA .103	EF .130
			Female:	FR .127	SO .110	JU .109	MA .104	EF .128
M31	Reference ties	F 4.91 <.001	Male:	FR .029	SO .036	JU .041	MA .024	EF .039
			Female:	FR .041	SO .029	JU .026	MA .026	EF .037
M64	Base clauses	F 4.47 <.01	Male:	FR 25.5%	SO 19.8%	JU 30.4%	MA 25.7%	EF 27.6%
			Female:	FR 29.9%	SO 29.9%	JU 27.9%	MA 25.4%	EF 27.0%
M65	Mono-clauses	F 5.58 <.001	Male:	FR 42.7%	SO 56.2%	JU 37.1%	MA 34.5%	EF 40.4%
			Female:	FR 35.3%	SO 33.5%	JU 36.2%	MA 42.2%	EF 41.3%
M66	Independent clauses	F 3.96 <.01	Male:	FR 68.1%	SO 76.1%	JU 67.5%	MA 60.2%	EF 68.0%
			Female:	FR 65.2%	SO 63.4%	JU 64.0%	MA 67.6%	EF 68.3%

Topic and Group:

M31	Reference ties	F 3.09 <.05	Topic A:	FR .029	SO .034	JU .035	MA .039	EF .036
			Topic B:	FR .043	SO .031	JU .032	MA .022	EF .040
M87	T-unit variance	F 3.74 <.01	Topic A:	FR 7.93	SO 7.33	JU 6.87	MA 8.27	EF 7.37
			Topic B:	FR 6.64	SO 6.81	JU 6.73	MA 10.15	EF 5.89
M67	Nominal Modific.	F 2.78 <.05	Topic A:	FR 34.9%	SO 34.0%	JU 36.2%	MA 38.4%	EF 31.3%
			Topic B:	FR 27.6%	SO 35.1%	JU 39.3%	MA 44.9%	EF 30.7%

TABLE 1

Use of Top-Level Organizational Patterns

(Complexity = rank assigned by Measure 4; % = percent of group using pattern; $\bar{X}$ = mean words embraced by pattern).

Pattern	Complexity	FR		SO		JU		MA		EF	
		%	$\bar{X}$	%	$\bar{X}$	%	$\bar{X}$	%	$\bar{X}$	%	$\bar{X}$
PARTITION		25.0%	204	21.9%	333	21.9%	357	3.1%	211	21.9%	313
Collection	1	12.5%	248	9.4%	377	9.4%	419	0.0%	---	21.9%	313
Classification	2	12.5%	160	12.5%	301	12.5%	311	3.1%	211	0.0%	---
SERIATION		28.1	245	31.3%	235	12.5%	383	15.6%	324	28.1%	347
Degree	3	12.5%	315	3.1%	184	0.0%	---	3.1%	393	3.1%	353
Development	4	6.3%	180	12.5%	261	12.5%	383	6.3%	313	15.6%	370
Comparison	5	9.4%	195	15.6%	225	0.0%	---	6.3%	301	9.4%	306
CONSEQUENCE		21.9%	239	9.4%	311	9.4%	190	9.4%	240	21.9%	227
Causation	6	21.9%	239	9.4%	311	9.4%	190	9.4%	240	18.8%	237
Process	7	0.0%	---	0.0%	---	0.0%	---	0.0%	---	3.1%	166
ARGUMENT		25.0%	147	21.9%	204	34.4%	206	37.5%	351	18.8%	371
Inference	8	9.4%	166	3.1%	73	6.3%	114	15.6%	340	3.1%	439
Choice	9	12.5%	142	9.4%	254	12.5%	195	6.3%	303	6.3%	379
Solution	10	0.0%	---	0.0%	---	6.3%	198	9.4%	381	9.4%	342
Dialectic	11	3.1%	110	9.4%	197	9.4%	288	6.3%	382	0.0%	---
CHAINING		0.0%	---	15.6%	241	21.9%	238	34.4%	355	9.4%	246
Causal Chain	12	0.0%	---	0.0%	---	6.3%	247	6.3%	210	3.1%	208
Sorites	13	0.0%	---	6.3%	286	3.1%	211	6.3%	240	3.1%	209
Sequence	14	0.0%	---	9.4%	211	12.5%	240	21.9%	429	3.1%	320
		100.0%	209	100.0%	258	100.0%	267	100.0%	334	100.0%	308

TABLE 2

Use of Lower-Level Logical Patterns

(Complexity = rank assigned by Measure 4; % = percent of all logical patterns beneath the top-level patterns, as determined by Measure 8)

Pattern	Complexity	FR	SO	JU	MA	EF
PARTITION		23.0%	17.9%	31.2%	29.0%	31.5%
Collection	1	6.7%	6.5%	11.7%	14.2%	13.6%
Classification	2	16.3%	11.4%	19.5%	14.8%	17.9%
SERIATION		31.7%	36.6%	32.0%	25.6%	30.2%
Degree	3	6.7%	4.1%	3.9%	2.7%	3.7%
Development	4	1.9%	6.5%	4.7%	3.8%	4.9%
Comparison	5	23.1%	26.0%	23.4%	19.1%	21.6%
CONSEQUENCE		27.0%	17.9%	18.0%	23.0%	18.6%
Causation	6	26.0%	17.9%	18.0%	21.9%	16.7%
Process	7	1.0%	0.0%	00.0%	1.1%	1.9%
ARGUMENT		9.7%	18.8%	12.5%	12.0%	14.9%
Inference	8	8.7%	9.8%	9.4%	7.1%	8.0%
Choice	9	1.0%	4.9%	2.3%	1.1%	1.9%
Solution	10	0.0%	0.0%	0.0%	2.2%	3.1%
Dialectic	11	0.0%	4.1%	0.8%	1.6%	1.9%
CHAINING		8.7%	8.9%	6.3%	10.4%	4.9%
Causal Chain	12	7.7%	7.3%	4.7%	7.7%	3.7%
Sorites	13	0.0%	0.0%	0.0%	0.0%	0.0%
Sequence	14	1.0%	1.6%	1.6%	2.7%	1.2%
		100.0%	100.0%	100.0%	100.0%	100.0%

TABLE 3

Use of Four Kinds of Qualifiers
Distinguished by Measure 10
Expressed as Incidents Per Word

Qualifiers	FR	SO	JU	MA	EF
(a) Modals	.0064	.0089	.0052	.0043	.0055
(b) Predicates	.0032	.0020	.0039	.0033	.0028
(c) Adjectives and Adverbs	.0041	.0055	.0067	.0038	.0057
(d) T-units	.0011	.0006	.0007	.0006	.0003
Summed rate (Measure 10)	.0148	.0170	0165	.0120	.0143

Table 4

Rubric for Primary-Trait Scoring
Of Introductions (Measure 20) and Conclusions (Measure 22)
With Number of Essays Earning Each Score

Criterion	Score	FR	SO	JU	MA	EF
A. Success of the introduction in conveying the scope of the main body of the essay:						
--No introduction	0	3	1	1	1	0
--Incorrect scope shown: the scope indicated by the introduction does not overlap with that shown by the main body	1	5	4	2	0	1
--Partial scope shown: the introduction indicates more or less than the body actually covers	2	16	14	15	9	6
--Full scope accurately shown by the introduction	3	8	13	14	22	25
		32	32	32	32	32
B. Success of the introduction in conveying the purpose for writing:						
--No introduction	0	3	1	1	1	0
--Taking of the purpose already provided by the examination topic, merely repeating in a non-committal way the topic's wording	1	13	11	8	5	3
--Indication of a self-generated purpose, but without emphasis	2	11	10	7	9	16
--Indication of a self-generated purpose and a divergence of opinion or timeliness; presence of a reasonable rhetorical question or phrases such as "strange to think," "main problem," "controversial"	3	5	10	16	17	13
		32	32	32	32	32
C. Coherence of the introduction						
--No introduction, or absolutely no direction apparent (e.g., one sentence without subordination)	0	10	3	5	4	3
--Connection of ideas, but little indication of direction	1	6	5	5	7	4
--Provision of direction, but with transitional gaps, stumbling blocks	2	13	12	12	5	14
--Clear and directional connection of ideas	3	3	12	10	16	13
		32	32	32	32	32

Table 4 (cont.)

Criterion	Score	FR	SO	JU	MA	EF
D. Introductory purpose of the title						
--No title	0	26	19	22	14	21
--The title indicates the subject only	1	6	9	5	15	8
--The title indicates the thesis	2	0	4	5	3	3
		32	32	32	32	32
E. Success of the conclusion						
--No conclusion	0	4	1	2	4	3
--Mere repetition or restatement of the introduction	1	6	2	1	0	4
--Incomplete summary of the main body of the essay	2	5	1	4	0	2
--Successful summary of the main body	3	6	9	9	7	4
--Summary with additional implications or new assessment	4	8	15	10	12	14
--Conclusion stands as the last part of a top-level logical pattern of Argument or Chaining (see Measure 4). If the essay is not unified by one logical pattern (see Measure 9), it cannot score 5	5	3	4	6	9	5
		32	32	32	32	32

TABLE 5

Use of Five Categories of Logical Indicators
Distinguished by Measure 26
Expressed as Incidents Per Word

Indicators	FR	SO	JU	MA	EF
Illustratives	.0025	.0025	.0021	.0014	.0033
Illatives	.0029	.0019	.0013	.0030	.0027
Adversatives	.0137	.0164	.0114	.0087	.0126
Causatives	.0105	.0078	.0102	.0077	.0081
Additives	.0345	.0315	.0297	.0354	.0299
Summed rate (Measure 26)	.0641	.0601	.0547	.0562	.0566

TABLE 6

Use of Nominal Modification by Levels of Complexity
Determined by Measure 68
Expressed as Percent of All Nominal Strings
(Pre-college Data Derived from Hunt, 1965, p. 117)

Level of Nominal Complexity	Grade 4	Grade 8	Grade 12	FP	EF	SO	JU	MA
1	84.7%	72.0%	71.4%	58.8%	56.3%	52.8%	52.7%	46.5%
2	12.6%	19.6%	19.3%	22.6%	25.3%	23.8%	23.6%	25.0%
3	2.1%	6.1%	6.3%	8.7%	10.1%	11.6%	10.9%	12.2%
4	0.5%	1.5%	1.5%	4.2%	4.7%	5.3%	5.4%	6.5%
5	0.2%	0.5%	0.7%	3.3%	1.7%	3.3%	3.5%	4.0%
6		0.1%	0.5%	1.5%	0.8%	1.7%	1.9%	2.0%
7			0.1%	0.5%	0.6%	0.6%	1.1%	1.1%
8			0.1%	0.1%	0.3%	0.4%	0.3%	1.5%
9			0.1%	0.1%	0.1%	0.3%	0.7%	0.2%
10						0.3%		0.4%
11				0.1%	0.1%		0.1%	0.3%
12					0.1%			0.1%
13								0.2%
14				0.1%	0.1%			0.1%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

TABLE 7

Use of Infinitive Structures
Distinguished by Measure 81
Expressed by Percent of Words in Structure
(And by Percent of All Infinitive Incidents)

	FR		SO		JU		MA		EF	
Structure	Word	Incident	Word	Incident	Word	Incident	Word	Incident	Word	Incident
Adjectival	2.8%	(26.3%)	3.1%	(28.5%)	4.2%	(35.6%)	2.0%	(20.7%)	1.9%	(24.6%)
Adverbial	1.5%	(10.4%)	0.9%	(10.6%)	0.5%	(04.6%)	2.2%	(20.0%)	0.6%	(8.0%)
Nominalized	1.3%	(11.9%)	1.2%	(8.3%)	0.7%	(5.3%)	1.4%	(9.9%)	1.8%	(13.3%)
Verbal complement	6.6%	(51.4%)	6.7%	(52.7%)	6.6%	(54.5%)	5.5%	(49.4%)	6.1%	(54.3%)
		(100.0%)		(100.0%)		(100.0%)		(100.0%)		(100.0%)
Summed rate (Measure 81)	12.2%		11.9%		12.0%		11.1%		10.4%	

TABLE 8

Use of Sentence Openers
Expressed as Percent of All Sentences
(Whole-group Rate)

Syntactic Function	FR	SO	JU	MA	EF
Independent					
Noun phrases and nominalizations	36.7%	35.4%	38.1%	37.5%	39.7%
Pronouns	19.2%	18.5%	23.1%	17.7%	19.5%
Questions	4.4%	5.3%	5.0%	6.1%	5.9%
Expletives and reversals	2.8%	0.2%	1.4%	3.6%	1.4%
Dependent					
Prepositional phrases	11.8%	11.9%	8.5%	9.5%	9.1%
Adverbial infinitives and participles	1.2%	4.0%	2.1%	1.4%	1.5%
Other adverbial structures	19.4%	20.3%	17.7%	19.9%	21.0%
Coordinating conjunctions	4.4%	4.4%	4.1%	4.3%	1.9%
	100.0%	100.0%	100.0%	100.0%	100.0%

TABLE 9

Use of Pronouns
Defined by Measure 44 but Excluding Relatives
Expressed as Percent of All Such Pronouns
(Whole-group Rates)

Pronoun Category	FR	SO	JU	MA	EF
	(Male) (Female)	(Male) (Female)	(Male) (Female)	(Male) (Female)	(Male) (Female)
First-person singular	12.3%	4.9%	18.9%	21.5%	10.9%
	(5.0%) (17.0%)	(4.5%) (5.2%)	(20.7%) (17.2%)	(23.4%) (20.3%)	(4.7%) (15.7%)
First-person plural	14.6%	20.6%	15.1%	22.2%	17.6%
	(15.1%) (14.2%)	(11.7%) (26.3%)	(10.7%) (15.0%)	(28.3%) (17.8%)	(19.7%) (15.9%)
Second-person generic "you"	14.0%	6.7%	16.3%	5.9%	5.5%
	(10.5%) (16.3%)	(11.7%) (3.5%)	(8.8%) (22.7%)	(2.9%) (8.0%)	(8.4%) (3.2%)
Third-person singular	25.5%	32.7%	24.8%	31.0%	34.2%
	(34.5%) (19.6%)	(35.1%) (31.4%)	(31.7%) (19.0%)	(21.6%) (37.5%)	(34.4%) (33.5%)
Third-person plural	33.6%	35.0%	25.0%	19.4%	32.2%
	(34.9%) (32.8%)	(36.9%) (33.8%)	(28.1%) (22.2%)	(23.0%) (16.6%)	(32.8%) (31.3%)
	100.0%	100.0%	100.0%	100.0%	100.0%
	(100.0%) (100.0%)	(100.0%) (100.0%)	(100.0%) (100.0%)	(100.0%) (100.0%)	(100.0%) (100.0%)

TABLE 10

Use of Intra-T-unit Syntactic Parallelism
Expressed as Percent of All Instances
(Whole-group Rate)

Type of Parallelism	FR	SO	JU	MA	EF
Coordinative: and, or, etc.					
Series of two elements (e.g., with a curl or a piece of clothing . . .)	80.2%	77.8%	76.7%	77.4%	77.6%
Prenominal series of three or more (e.g., the tall, dark, and handsome type)	1.2%	0.7%	2.3%	3.5%	1.9%
Other series of three (e.g., They comb your hair, dress you neatly, and present you . . .)	7.6%	8.3%	9.5%	9.7%	7.2%
Other series of four or more (e.g., Their interest focusses on sports, hobbies, work, and being with the group)	2.4%	1.9%	1.3%	4.2%	1.3%
Subordinative: not, than, but, etc.					
Comparatives (e.g., to look worse when they come out than they did when they went in)	5.1%	6.9%	8.2%	2.1%	8.2%
Other (e.g., I don't mean a small argument, but a heated debate)	3.4%	4.4%	2.1%	3.1%	3.8%
	100.0%	100.0%	100.0%	100.0%	100.0%

VIII. Conclusions

CONCLUSIONS. As has been mentioned, a lengthy analysis of these findings, assessing their implications for teachers and administrators, has been written (Haswell, 1986a), and that assessment will not be iterated here. I append, however, a few summary points that may help researchers judge the viability of the study and interpret the quantified findings reported above.

When the seven pre-test variables listed above on p. 7 were used as independent variable on separate analyses of variance, no significant differences among the four undergraduate groups was found. The Washington Pre-College Examination verbal composite scores for these four groups match almost exactly, both in mean and in variance, the scores of all students in the state of Washington who take that test ($\bar{X}$ of 50 and S.D. of 10). The sampling for this study, then, generated well matched groups, who represent, as judged at least by the measures of the Washington Pre-College test, a fairly typical entering class of freshman students.

The matching of the two essay prompts also seems to have succeeded fairly well. On only seven measures were significant differences found in the way they surfaced with regard to topic. There was even less effect of topic interacting with group upon the measures, and significant interactions were discovered in only three.

Differences in the way the two sexes wrote, however, appeared somewhat more often than research into gender traits in adult writing may lead one to expect. Significant differences were found in 16 of the 84 pre-set measures. There were five significant interactions between sex and group, indicating

that during college males and females may develop somewhat differently in their writing--a phenomenon that has been seen repeatedly in pre-college language maturation. But interestingly enough, there were no interactions between sex and topic.

The chief discovery of this study is that writing does seem to change during the first two years of college, and to change toward what may be characterized as competent, workaday performance. The writing of the post-college group certainly differs from that of the students; in 44 of the 84 pre-set measures the MA essays differed significantly from one or more of the student sets. Yet from one student group to another there was also movement, and of the 84 measures junior or sophomore essays differ significantly from freshman in 21. More importantly, in 19 of these 21 this performance of the students shifts in the direction of the post-college writing. This is to say that where post-freshman writing changes, it changes largely toward older, working-world competence, not from.

A second look at the 44 measures where post-college writing differs from college reveals an even more sweeping undergraduate shift toward work-world competence. Of these 44 the junior essays shift from the freshman in the direction of the older writers in 35. Here, of course, many of the shifts from FR to JU do not reach statistical significance, and some are minute, but an assumption of random movement from FR to JU would hypothesize around 50% of these differences to favor the JU writing, and instead we get nearly 80%. The odds that such a preponderance might happen by chance are small ($p < .001$ by sign test). Even this hefty affirmation of the maturation of the post-freshman underclass students possibly understates the case. Of

the nine measures where junior writing ebbs back toward freshman rather than toward matured writing, the differences in eight are statistically non-significant. The odds are we are seeing not a backsliding but a plateau. On only one of the 44 (coordination of nouns, M59) does this study provide statistical assurance that sophomores or juniors reverse the trend toward the levels of matured writing set by the MA writers

A final note. The composition teachers who executed the holistic rating of all the essays together (Measure 1) saw no difference in quality among the three first-of-the-semester student essays (FR, SO, JU) and also saw post-college writing as substantially better than student. Yet the analytic measures found these very students progressing toward that very post-college performance. The implication is that college composition teachers do not have a very good eye for those particular writing traits where students are maturing. This blindness may help account for the common belief among teachers that college writing does not get better and perhaps even gets worse (e.g., Hammond 1984). The argument advanced by this study, that the belief is wrong, has obvious and major implications for the instruction of writing in college.

IX. Acknowledgments

ACKNOWLEDGMENTS. Funding for the collecting of research material, for computer resources, and in part for manuscript preparation was supported by the Basil and Eila Gerard Endowment Fund, the Washington State University Humanities Research Center, and the WSU College of Sciences and Arts.

Five individuals in particular deserve a line of credit to themselves: Laura Bloxham, Irene D. Hayes, Bill Krieger, Linda Stairet, and Susan Wyche-Smith.

Central to this study was the collaboration of three generations (year generations) of teacher-students. I owe special obligations to a few, for having devised measures adopted or adapted in this investigation: Kathy Cusack and Paul Meyers, Measure 10; Christina Haas and Jennie Nelson, Measures 13-16; Veronica Franklin-Ellis, Measure 24; Sue Shirley, Measure 26; Steve Simko, Measures 28-32; Jean Heglund, Measure 44; Helen Pournaya-Karydas, Measure 85; Lloyd Winston, Measure 86; Derise Wigand, Measure 92; Linda Stairet, Measure 95.

X. Bibliography of works cited

Bailey, Richard W., Richard T. Brengle, and Edward L. Smith, Jr., "Measuring Student Writing Ability." In Reinventing the Rhetorical Tradition, ed. by Aviva Freedman and Ian Pringle, pp. 137-144. Conway, Arkansas: L & S Books, 1980.

Berlin, Brent. "Speculation on the Growth of Ethnobotanical Nomenclature," Language in Society 1 (1972): 51-86.

Boder, D. P. "The Adjective-verb Quotient: A Contribution in the Psychology of Language," Psychological Record 3 (1940): 309-343.

Britton, James, Tony Burgess, Nancy Martin, Alex McLeod, Harold Rosen. The Development of Writing Abilities (11-18), Schools Council Research Studies. London: MacMillan, 1975.

Broadhead, Glenn J., James A. Berlin, Marlis Manley Broadhead. "Sentence Structure in Academic Prose and Its Implication for College Writing Teachers," Research in the Teaching of English 16, no. 3 (Oct. 1982): 225-240.

Christensen, Francis. "The Problem of Defining a Mature Style," English Journal 57 (April 1968): 574-578.

D'Angelo, Frank J. A Conceptual Theory of Rhetoric. Cambridge, Mass.: Winthrop, 1975.

Davis, Ken. "Significant Improvement in Freshman Composition as Measured by Impromptu Essays: A Large-Scale Experiment," Research in the Teaching of English 13, no. 1 (Feb. 1979): 45-48.

Eckhardt, C. D., and D. H. Stewart. "Towards a Functional Taxonomy of Composition," College Composition and Communication 30 (1979): 338-342.

Eblen, Charlene. "Growth in Writing During College Years: A Preliminary Look" (March, 1981). ERIC Document Reproduction Service No. 217 423.

Faigley, Lester. "The Influence of Generative Rhetoric on the Syntactic Maturity and Writing Effectiveness of College Freshmen," Research in the Teaching of English 13, no. 3 (Oct. 1979): 197-206.

Frank, Marcella. Modern English: A Practical Reference Guide. Englewood Cliffs, NJ: Prentice-Hall, 1972.

Gebhard, Ann O. "Writing Quality and Syntax: A Transformational Analysis of Three Prose Samples," Research in the Teaching of English 12, no. 3 (Oct. 1978): 211-232.

Glass, G. V., P. D. Peckham, and J. R. Sanders. "Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analysis of Variance and Covariance," Review of Educational Research 15 (1959): 182-191.

Gorrell, Donna. "Toward Determining a Minimal Competency Entrance Examination for Freshman Composition," Research in the Teaching of English 17, no. 3 (Oct. 1933): 263-274.

Grabor, Terry, Leo Hines, and Irene Miranda. "Measuring Writing Progress:

An Experiment," College English 35, no. 4 (Jan 1974): 484-485.

Gray, Donald. "Another Year with College English," College English 44 (1982): 385-389.

Hagen, Lyman B. "An Analysis of Transitional Devices in Student Writing," Research in the Teaching of English 5, no. 2 (1971): 190-201.

Halliday, M. A. K., and Ruqaiya Hasan. Cohesion in English. London: Longman, 1976.

Hammond, Eugene. "Freshman Composition--Junior Composition: Does Co-Ordination Mean Sub-Ordination?" College Composition and Communication 35, no. 2 (May 1984): 217-221.

Haswell, Richard H. "Motives and Measures: Studies in the Development of Student Writing During the First Two Years of College." Unpublished manuscript, 288 pages.

_____. "The Organization of Impromptu Essays," forthcoming in College Composition and Communication.

Hiatt, Mary. The Way Women Write. New York: Teachers College Press, 1977.

Hillocks, George Jr. "Inquiry and the Composing Process: Theory and Research," College English 44, no. 7 (Nov. 1982): 659-673.

Hull, C. Hadlai, and Norman H. Nie. SPSS Update: New Procedures and Facilities for Releases 7 and 8. New York: McGraw-Hill, 1979.

Hunt, Kellogg W. Grammatical Structures Written at Three Grade Levels. National Council of Teachers of English Research Report no. 3. Champaign, IL: National Council of Teachers of English, 1965.

Inhelder, Barbel, and Jean Piaget. The Growth of Logical Thinking from Child- to Adolescence: An Essay on the Construction of Formal Operational Structures, translated by Anne Parsons and Stanley Milgram. New York:

Basic Books, 1958.

Jewell, Ross M., John Cowley, and Gordon Rhum. The Effectiveness of College-Level Instruction in Freshman Composition. Final Report, U. S. Office of Education Project No. 2188, Amended. Cedar Falls: University of Northern Iowa, 1969.

Kerek, Andrew, Donald A. Daiker, and Max Morenberg. "Sentence Combining and College Composition," Perceptual and Motor Skills 51 (1980): 1059-1157. Monograph Supplement 1-V51.

Kimmey, John L. "Freshman Composition in the Junior Year," College Composition and Communication 26 (Dec. 1975): 347-349.

Kirk, R. E. Experimental Design for the Behavioral Sciences. Belmont, CA: Brooks/Cole, 1968.

Kitzhaber, Albert R. Themes, Theories and Therapy: The Teaching of Writing in College. New York: McGraw-Hill, 1963.

Klausmeier, H. J., E. S. Ghatala, and D. A. Frayer. Conceptual Learning and Development: A Cognitive View. New York: Academic Press, 1974.

Lindman, Harold R. Analysis of Variance in Complex Experimental Designs. San Francisco, CA: W. H. Freeman, 1974.

Longacre, Robert E. "The Paragraph as a Grammatical Unit." In Discourse and Syntax, ed. Talmy Givon, pp. 115-134. New York: Academic Press, 1979.

Maimon, Elaine P. and Barbara F. Nodine. "Words Enough and Time: Syntax and Error one Year After." In Sentence Combining and the Teaching of Writing, ed. Donald A. Daiker, Andrew Kerek, ar Max Morenberg, pp. 101-108. Akron, OH: University of Akron, 1979.

Meyer, Paul R. "A Study of Sex Differences in the Freshman Composition Course at the University of Texas at Austin" (1982). ERIC Document Reproduction

Service No. ED 214 167.

Miller, George A. and Phillip N. Johnson-Laird. Language and Perception. Cambridge, MA: Harvard University Press, 1976.

Miller, Susan. "Rhetorical Maturity: Definition and Development." In Reinventing the Rhetorical Tradition, ed. Aviva Freedman and Ian Pringle, pp. 119-127. Conway, AK: L & S Books, 1980.

Morenberg, Max, Donald A. Daiker, and Andrew Kerek. "Sentence-Combining at the College Level: An Experimental Study," Research in the Teaching of English 12 (1978): 233-244.

Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karin Stinbrenner, and Dale H. Bent. Statistical Package for the Social Sciences, 2nd ed. New York: McGraw-Hill, 1975.

Noeth, Richard J. A 10-Year Study of Test Scores for College-Bound Students in the State of Washington. Seattle, WA: The Washington Pre-College Testing Program, 1979.

Pitkin, Willis L. Jr. "Discourse Blocs," College Composition and Communication 20 (1969): 138-148.

Purves, Alan C. "In Search of an Internationally-Valid Scheme for Scoring Compositions," College Composition and Communication 35, no. 4 (Dec. 1984): 426-438.

Rest, J. Development in Judging Moral Issues. Minneapolis, MN: University of Minnesota Press, 1979.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Broen. "Basic Objects in Natural Categories," Cognitive Psychology 8 (1976): 382-439.

Stewart, Murray F. "Syntactic Maturity from High School to University: A

First Look," Research in the Teaching of English 12, no. 1 (Feb. 1978): 37-46.

Thorndike, Edward L. and Irving Lorge. The Teacher's Word Book of 30,000 Words. New York: Teacher's College, Columbia University, 1944.

Van Dijk, Teun A. Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition. Hillsdale, NJ: Lawrence Erlbaum, 1980.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 962.

Witte, Stephen P. and Lester Faigley. "A Comparison of Analytic and Synthetic Approaches to the Teaching of College Writing." TWRG Research Report No. 1. Austin, TX: Department of English, University of Texas, 1981. ERIC Document Reproduction Service No. ED 209 677.

Whitla, Dean K. Value Added: Measuring the Outcome of Undergraduate Education. Cambridge MA, Office of Instructional Research and Evaluation: Harvard University, 1977

Wohlvill, Joachim F. The Study of Behavioral Development. New York: Academic P ;s, 1973.