Frequency of Formal Errors in Current College Writing, or Ma and Pa Kettle Do Research

Robert J. Connors and Andrea A. Lunsford

Proem: In Which the Characters Are Introduced

The labyrinthine project of which this research is a part represents an ongoing activity for us, something we engage in because we like to work together, have a long friendship, and share many interests. As we worked on this error research together, however, we started somewhere along the line to feel less and less like the white-coated Researchers of our dreams and more and more like characters we called Ma and Pa Kettle—good-hearted bumblers striving to understand a world whose complexity was more than a little daunting. Being fans of classical rhetoric, *prosopopoeia, letteraturizzazione*, and the like, as well as enthusiasts for intertextuality, *plaisir de texte, differance*, etc., we offer this account of our travails—with apologies to Marjorie Main and Percy Kilbride.

Exordium: The Kettles Smell a Problem

Marking and judging formal and mechanical errors in student papers is one area in which composition studies seems to have a multiple-personality disorder. On the one hand, our mellow, student-centered, process-based selves tend to condemn marking formal errors at all. Doing it represents the Bad Old Days. Ms. Fidditch and Mr. Flutesnoot with sharpened red pencils, spilling innocent blood across the page. Useless detail work. Inhumane, perfectionist standards, making our students feel stupid, wrong, trivial, misunderstood. Joseph Williams has pointed out how arbitrary and context-bound our judgments of formal error are. And certainly our noting of errors on student papers gives no one any great joy; as Peter Elbow says, English is most

Robert J. Connors is an associate professor of English at the University of New Hampshire. The author of many articles on rhetorical history and theory, he was the winner of the 1982 Richard Braddock Award. Andrea A. Lunsford is professor of English and Vice-Chair for Rhetoric and Composition at The Ohio State University. She has published numerous articles and was co-winner, with Lisa Ede, of the 1985 Richard Braddock Award. Bob and Andrea are co-authors of the forthcoming *St. Martin's Handbook*.

often associated *either* with grammar or with high literature—"two things designed to make folks feel most out of it."

Nevertheless, very few of us can deny that an outright comma splice, its/it's error, or misspelled common word distracts us. So our more traditional pedagogical selves feel a touch guilty when we ignore student error patterns altogether, even in the sacrosanct drafting stage of composing. Not even the most liberal of process-oriented teachers completely ignores the problem of mechanical and formal errors. As Mina Shaughnessy put it, errors are "unintentional and unprofitable intrusions upon the consciousness of the reader. . . . They demand energy without giving back any return in meaning" (12). Errors are not merely mechanical, therefore, but rhetorical as well. The world judges a writer by her mastery of conventions, and we all know it. Students, parents, university colleagues, and administrators expect us to deal somehow with those unmet rhetorical expectations, and, like it or not, pointing out errors seems to most of us part of what we do.

Of course, every teacher has his or her ideas of what errors are common and important, but testing those intuitive ideas is something else again. We became interested in error-frequency research as a result of our historical studies, when we realized that no major nationwide analysis of actual college essays had been conducted, to our knowledge, since the late 1930s. As part of the background for a text we were writing and because the research seemed fascinating, we determined to collect a large number of college student essays from the 1980s, analyze them, and determine what the major patterns of formal and mechanical error in current student writing might be.

Narratio: Ma and Pa Visit the Library

Coming to this research as historians rather than as trained experimenters has given us a humility based on several different sources. Since we are not formally trained in research design, we have constantly relied on help from more expert friends and colleagues. Creating a sense of our limitations even more keenly, however, have been our historical studies. No one looking into the history of research on composition errors in this country can emerge very confident about definitions, terms, and preconceptions. In almost no other pedagogical area we have studied do the investigators and writers seem so time-bound, so shackled by their ideas of what errors *are*, so blinkered by the definitions and demarcations that are part of their historical scene. And, ineluctably, we must see ourselves and our study as history-bound as well. Thus we write not as the torchbearers of some new truth, but as two more in the long line of people applying their contemporary perspectives to a numbering and ordering system and hoping for something of use from it.

The tradition of research into error patterns is as old as composition teaching, of course, but before the growth of the social-science model in education

it was carried on informally. Teachers had "the list" of serious and common errors in their heads, and their lists were probably substantially similar (although "serious" and "common" were not necessarily overlapping categories).[1] Beginning around 1910, however, teachers and educational researchers began trying to taxonomize errors and chart their frequency. The great heyday of error-frequency seems to have occurred between 1915 and 1935. During those two decades, no fewer than thirty studies of error frequency were conducted.[2] Unfortunately, most of these studies were flawed in some way: too small a data sample, too regional a data sample, different definitions of errors, faulty methodologies (Harap 440). Most early error research is hard to understand today because the researchers used terms widely understood at the time but now incomprehensible or at best strange. Some of the studies were very seriously conducted, however, and deserve further discussion later in this paper.

After the middle 1930s, error-frequency research waned as the progressive-education movement gained strength and the "experience curriculum" in English replaced older correctness-based methods. Our historical research indicates that the last large-scale research into student patterns of formal error was conducted in 1938-39 by John C. Hodges, author of the *Harbrace College Handbook*. Hodges collected 20,000 student papers that had been marked by 16 different teachers, mainly from the University of Tennessee at Knoxville. He analyzed these papers and created a taxonomy of errors, using his findings to inform the 34-part organization of his *Harbrace Handbook*, a text which quickly became and remains today the most popular college handbook of writing.

However Hodges may have constructed his study, his results fifty years later seem problematic at best. Small-scale studies of changes in student writing over the past thirty years have shown that formal error patterns have shifted radically even since the 1950s. The kinds and quantities of formal errors revealed in Mina Shaughnessy's work with basic writers in the 1970s were new and shocking to many teachers of writing. We sensed that the time had come for a study that would attempt to answer two questions: (1) what are the most common patterns of student writing errors being made in the 1980s in the United States?, and (2) which of these patterns are marked most consistently by American teachers?

Confirmatio I: The Kettles Get Cracking

The first task we faced was gathering data. We needed teacher-marked papers from American college freshmen and sophomores in a representative range of different kinds of schools and a representative range of geographic areas. We did not want to try to gather the isolated sample of timed examination-style writing that is often studied, although such a sample would probably have been easier to obtain than the actual marked papers we sought. We wanted "themes in the raw," the actual commerce of writing courses all across Amer-

ica. We wanted papers that had been personally marked or graded, filled with every uncontrolled and uncontrollable sign of both student and teacher personalities.

Gathering these papers presented a number of obstacles. In terms of ideal methodology, the data-gathering would be untouched by self-selection among teachers, and we could randomly choose our sources. After worrying about this problem, we finally could conceive of no way to gather upwards of 20,000 papers (the number of papers Hodges had looked at) without appealing to teachers who had marked them. We could think of no way to go directly to students, and, though some departments stockpile student themes, we did not wish to weight our study toward any one school or department. We had to ask composition teachers for help.

And help us they did. In response to a direct mail appeal to more than 1,500 teachers who had used or expressed interest in handbooks, we had received by September 1985 more than 21,500 papers from 300 teachers all across America.[3]

To say that the variety in the papers we were sent was striking is a serious understatement. They ranged in length from a partial page to over 20 pages. About 30% were typed, the rest handwritten. Some were annotated marginally until they looked like the Book of Kells, while others merely sported a few scrawled words and a grade. Some were pathologically neat, and others look dashed off on the jog between classes. Some were formally perfect, while others approximated Mina Shaughnessy's more extreme examples of basic writing. Altogether, the 21,500+ papers, each one carefully stamped by paper number and batch number, filled approximately 30 feet of hastily-installed shelving. It was an imposing mass.

We had originally been enthusiastic (and naive) enough to believe that with help we might somehow look over and analyze 20,000 papers. Wrong. Examining an average paper even for mechanical lapses, we soon realized, took at the very least ten busy minutes; to examine all of them would require over 3,000 Ma-and-Pa-hours. We simply could not do it. But we could analyze a carefully stratified sample of 3,000 randomly chosen papers. Such an analysis would give us data that were very reliable. Relieved that we would not have to try to look at 20,000 papers, we went to work on the stratification.[4] After stratifying our batches of papers by region, size of school, and type of school, we used the table of random numbers and the numbers that had been stamped on each paper as it came in to pull 3,000 papers from our tonnage of papers. Thus we had our randomized, stratified sample, ready for analysis.

Confutatio: Ma and Pa Suck Eggs

But—analyzed using what? From very early on in the research, we realized that trying to introduce strict "scientific" definitions into an area so essentially

values-driven as formal error marking would be a foolhardy mistake. We accepted Joe Williams' contention that it is "necessary to shift our attention from error treated strictly as an isolated item on a page, to error perceived as a flawed verbal transaction between a writer and a reader" (153). Williams' thoughtful article on "The Phenomenology of Error" had, in fact, persuaded us that some sort of reader-response treatment of errors would be far more useful than an attempt to standardize error patterns in a pseudo-scientific fashion based on Hodges' or any other handbook.

We were made even more distrustful of any absolutist claims by our further examination of previous error-frequency research. Looking into the history of this kind of research showed us clearly how teachers' ideas about error definition and classification have always been absolute products of their times and cultures. What seem to us the most common and permanent of terms and definitions are likely to be newer and far more transient than we know. Errors like "stringy sentences" and "use of *would* for simple past tense forms" seemed obvious and serious to teachers in 1925 or 1917 but obscure to us today.[5]

While phenomena and adaptable definitions do continue from decade to decade, we knew that any system we might adopt, however defensible or linguistically sound it might seem to us, would someday represent one more historical curiosity. "Comma splice?" some researcher in the future will murmur, "What a strange term for Connors and Lunsford to use. Where could it have come from?"[6] Teachers have always marked different phenomena as errors, called them different things, given them different weights. Error-pattern study is essentially the examination of an ever-shifting pattern of skills judged by an ever-shifting pattern of prejudices. We wanted to try looking at this situation as it existed in the 1980s, but clearly the instrument we needed could not be algorithmic and would not be historically stable.

We settled, finally, on several general understandings. First, examining what teachers had marked on these papers was as important as trying to ascertain what was "really there" in terms of formal error patterns. Second, we could only analyze for a limited number of error patterns—perhaps twenty in all. And finally, we had no taxonomy of errors we felt we could trust. We would have to generate our own, then, using our own culture- and time-bound definitions and perceptions as best we could.

Confirmatio II: Ma and Pa Hit the Road

Producing that taxonomy meant looking closely at the papers. Using the random number tables again, we pulled 300 papers from the remaining piles. Each of us took 150, and we set out inductively to note every formal error pattern we could discover in the two piles of papers. During this incredibly boring and nauseating part of the study, we tried to ignore any elements of paper content or organization except as they were necessary to identify errors. Every

error marked by teachers was included in our listing, of course, but we found many that had not been marked at all, and some that were not even easily definable. What follows is the list of errors and the numbers of errors we discovered in that first careful scrutiny of 300 papers:

Error or Error Pattern	# in 300 Papers
Spelling	450
No comma after introductory element	138
Comma splice	124
Wrong word	102
Lack of possessive apostrophe	99
Vague pronoun reference	90
No comma in compound sentence	87
Pronoun agreement	83
Sentence fragment	82
No comma in non-restrictive phrase	75
Subject-verb agreement	59
Unnecessary comma with restrictive phrase	50
Unnecessary words/style rewrite	49
Wrong tense	46
Dangling or misplaced modifier	42
Run-on sentence	39
Wrong or missing preposition	38
Lack of comma in series	35
Its/it's error	34
Tense shift	31
Pronoun shift/point of view shift	31
Wrong/missing inflected endings	31
Comma with quotation marks error	28
Missing words	27
Capitalization	24
"Which/that" for "who/whom"	21
Unidiomatic word use	17
Comma between subject and verb	14
Unnecessary apostrophe after "s"	11
Unnecessary comma in complex sentence	11
Hyphenation errors	9
Comma before direct object	6
Unidiomatic sentence pattern	6
Title underlining	6
Garbled sentence	4
Adjectival for adverbial form—"ly"	4

In addition, the following errors appeared fewer than 4 times in 300 papers:

Wrong pronoun
Wrong use of dashes
Confusion of a/an
Missing articles (the)
Missing question mark
Wrong verb form

Lack of transition
Missing/incorrect quotation marks
Incorrect comma use with parentheses
Use of comma instead of "that"
Missing comma before "etc."
Incorrect semicolon use
Repetition of words
Unclear gerund modifier
Double negative
Missing apostrophe in contraction
Colon misuse
Lack of parallelism

As expected, many old favorites appear on these lists. To our surprise, however, some errors we were used to thinking of as very common and serious proved to be at least not so common as we had thought. Others, which were not thought of as serious (or even, in some cases, as actual errors), seemed very common.

Our next step was to calibrate our readings, making certain we were both counting apples as apples, and to determine the cutoff point in this list, the errors we would actually count in the 3,000 papers. Since spelling errors predominated by a factor of 300% (which in itself was a surprising margin), we chose not to deal further with spelling in this analysis, but to develop a separate line of research on spelling. Below spelling, we decided to go arbitrarily with the top twenty error patterns, cutting off below "wrong inflected ending." These were the twenty error patterns we would train our analysts to tote up.

Now we had a sample and we had an instrument, however rough. Next we needed to gather a group of representative teachers who could do the actual analysis. Fifty teaching assistants, instructors, and professors from the Ohio State University English Department volunteered to help us with the analysis. The usual question of inter-rater reliability did not seem pressing to us, because what we were looking for seemed so essentially charged with social conditioning and personal predilection. Since we did not think that we could always "scientifically" determine what was real error and what was style or usage variation, our best idea was to rationalize the arbitrariness inherent in the project by spreading out the analytical decisions.

On a Friday afternoon in January 1986 we worked with the fifty raters, going over the definitions and examples we had come up with for the "top twenty," as we were by then calling them. It was a grueling Friday and Saturday. We trained raters to recognize error patterns all Friday afternoon in the dusty, stuffy old English Library at OSU—the air of which Thurber must have breathed, and probably the very same air, considering how hard the windows were to open. On returning to our hotel that night, we found it occupied by the Ohio chapter of the Pentecostal Youth, who had been given permission to run around the hotel giggling and shouting until 3:30 a.m. In

despair, we turned our TV volumes all the way up on white-noise stations that had gone off the air. They sounded like the Reichenbach Falls and almost drowned out the hoo-raw in the hallway. After 3:30 it did indeed quiet down some, and we fell into troublous sleep. The next day the Pentecostal Youth had vanished, and Ma & Pa had research to do.

Amplificatio: Ma and Pa Hunker Down

The following day, rating began at 9:00 a.m. and, with a short lunch break, we had completed the last paper by 5:00 p.m. We paused occasionally to calibrate our ratings, to redefine some term, or to share some irresistible piece of student prose. (Top prize went to the notorious "One Night," one student's response to an assignment asking for "analysis." This essay's abstract announced it as "an analysis of the realm of different feelings experienced in one night by a man and wife in love."[7]) The rating sheets and papers were reordered and bundled up, and we all went out for dinner.[8]

The results of this exercise became real for us when we totaled up the numbers on all of the raters' sheets. Here was the information we had been seeking, what all our efforts had been directed toward. It was exciting to finally see in black and white what we had been wondering about. What we found appears in Table 1.

Peroratio: The Kettles Say, "Aw, Shucks"

The results of this research by no means represent a final word on any question involving formal errors or teacher marking patterns. We can, however, draw several intriguing, if tentative, generalizations.

First, teachers' ideas about what constitutes a serious, markable error vary widely. As most of us may have expected, some teachers pounce on every "very unique" as a pet peeve, some rail at "Every student . . . their . . ." The most prevalent "error," failure to place a comma after an introductory word or phrase, was a *bête noire* for some teachers but was ignored by many more. Papers marked by the same teacher might at different times evince different patterns of formal marking. Teachers' reasons for marking specific errors and patterns of error in their students' papers are complex, and in many cases they are no doubt guided by the perceived needs of the student writing the paper and by the stage of the composing process the paper has achieved.

Second, teachers do not seem to mark as many errors as we often think they do. On average, college English teachers mark only 43% of the most serious errors in the papers they evaluate. In contrast to the popular picture of English teachers mad to mark up every error, our results show that even the most-often marked errors are only marked two-thirds of the time. The less-marked patterns (and remember, these are the Top Twenty error patterns overall) are

Table 1

Error or Error Pattern	# found in 3000 papers	% of total errors	# found marked by teacher	% marked by teacher	rank by # of errors marked by teacher
1. No comma after introductory element	3,299	11.5%	995	30%	2
2. Vague pronoun reference	2,809	9.8%	892	32%	4
3. No comma in compound sentence	2,446	8.6%	719	29%	7
4. Wrong word	2,217	7.8%	1,114	50%	1
5. No comma in non-restrictive element	1,864	6.5%	580	31%	10
6. Wrong/missing inflected endings	1,679	5.9%	857	51%	5
7. Wrong or missing preposition	1,580	5.5%	679	43%	8
8. Comma splice	1,565	5.5%	850	54%	6
9. Possessive apostrophe error	1,458	5.1%	906	62%	3
10. Tense shift	1,453	5.1%	484	33%	12
11. Unnecessary shift in person	1,347	4.7%	410	30%	14
12. Sentence fragment	1,217	4.2%	671	55%	9
13. Wrong tense or verb form	952	3.3%	465	49%	13
14. Subject-verb agreement	909	3.2%	534	58%	11
15. Lack of comma in series	781	2.7%	184	24%	19
16. Pronoun agreement error	752	2.6%	365	48%	15
17. Unnecessary comma with restrictive element	693	2.4%	239	34%	17
18. Run-on or fused sentence	681	2.4%	308	45%	16
19. Dangling or misplaced modifier	577	2.0%	167	29%	20
20. Its/it's error	292	1.0%	188	64%	18

marked only once for every four times they appear. The number of errors found compared to the number of errors marked suggests a fascinating possibility for future research: detailed observation of teacher marking, accompanied by talk-aloud protocols. Such research seems to us a natural follow-up to the findings presented here.[9]

Third, the reasons teachers mark any given error seem to result from a complex formula that takes into account at least two factors: how serious or annoying the error is perceived to be at a given time for both teacher and student, and how difficult it is to mark or explain. As Table 1 shows, the errors marked by the original teachers on our papers produce a different (although not completely dissimilar) ranking of errors than the formal count we asked our raters to do. Some of the lesser-marked errors we studied are clearly felt to be more stylistic than substantive. Certain of the comma errors seem simply not to bother teachers very much. Others, like wrong words or missing inflections, are much more frequently marked, and might be said to have a high "response quotient" for teachers. In addition, we sensed that in many cases errors went unmarked not because the teacher failed to see them, but because they were not germane to the lessons at hand. A teacher working very hard to help a student master subject-verb agreement with third-person singular nouns, for instance, might well ignore most other errors in a given paper.

Teachers' perceptions of the seriousness of a given error pattern seem, however, to be only part of the reason for marking an error. The sheer difficulty of explanation presented by some error patterns is another factor. Jotting "WW" in the margin to tip a student off to a diction problem is one thing; explaining a subtle shift in point of view in that same marginal space is quite another. Sentence fragments, comma splices, and wrong tenses, to name three classic "serious" errors, are all marked less often than possessive apostrophes. This is, we think, not due to teachers' perception that apostrophe errors are worse than sentence-boundary or tense problems, but to their quickness and ease of indication. The its/it's error and the possessive apostrophe, the two highest-marked patterns, are also two of the easiest errors to mark. This is, of course, not laziness; many composition teachers are so chronically overworked that we should not wonder that the errors most marked are those most quickly indicated.

Fourth, error patterns in student writing are shifting in certain ways, at least partially as a result of changing media trends within the culture. Conclusions must be especially tentative here, because the time-bound nature of studies of error makes comparisons difficult and definitions of errors counted in earlier research are hard to correlate. Our research turned up several earlier lists of serious errors in freshman composition, however, whose order is rather different from the order we discovered.

Roy Ivan Johnson, writing in 1917, reported on 198 papers written by 66 freshmen, and his list of the top ten error patterns in his study is as follows (wherever possible, we have translated his terms into ours):

1. Spelling
2. Capitalization
3. Punctuation (mostly comma errors)
4. Careless omission or repetition
5. Apostrophe errors
6. Pronoun agreement
7. Verb tense errors and agreement
8. Ungrammatical sentence structure (fragments and run-ons)
9. Mistakes in the use of adjectives and adverbs
10. Mistakes in the use of prepositions and conjunctions

In 1930, Paul Witty and Roberta Green analyzed 170 papers written in a timed situation by freshmen. Here is their top ten list, translated into our terms where possible:

1. Faulty connectives
2. Vague pronoun reference
3. Use of "would" for simple past tense forms
4. Confusion of forms from similarity of sound or meaning
5. Misplaced modifiers
6. Pronoun agreement
7. Fragments
8. Unclassified errors
9. Dangling modifiers
10. Wrong tense

As we mentioned earlier, the largest-scale analysis of errors was done by John C. Hodges in the late 1930s. Unfortunately, we know very little about Hodges' research. He never published any results in contemporary journals, and thus it is difficult to know his methods or even very much about his findings, because we can see them only as they are reflected in the *Harbrace Handbook*, which today still uses the exact arrangement that Hodges gave it in its first edition in 1941. In the "To the Instructor" preface of his first edition, Hodges says that his 20,000 themes "have been tabulated according to the corrections marked by sixteen instructors," which suggests that his raters looked only for teacher-marked errors (Hodges iii). In a footnote on the same page, Hodges gives the only version of his top-ten list ever published:

1. Comma
2. Spelling
3. Exactness
4. Agreement
5. Superfluous commas
6. Reference of pronouns
7. Apostrophe
8. Omission of words
9. Wordiness
10. Good use

That is all we know of Hodges' findings, but it does not seem unreasonable to assume that he reports them in order of frequency.

In terms of how patterns of error have changed, our findings are, of course, extremely tentative. Assuming that Hodges' *Harbrace* list constitutes some version of the error patterns he found in 1939, however, we note some distinct changes. In general, our list shows a proliferation of error patterns that seem to suggest declining familiarity with the visual look of a written page. Most strikingly, spelling errors have gone from second on the list to first by a factor of three. Spelling is the most obvious example of this lack of visual memory of printed pages seen, but the growth of other error patterns supports it as well.[10]

Some of the error patterns that seem to suggest this visual-memory problem were not found or listed in earlier studies but have come to light in ours. The many wrong word errors, the missing inflected endings, the wrong prepositions, even the its/it's errors—all suggest that students today may be less familiar with the visible aspects of written forms. These findings confirm the contrastive analysis between 2,000 papers from the 1950s and 2,000 papers from the 1970s that was carried out by Gary Sloan in 1979. Sloan determined that many elements of formal writing convention broke down severely between the fifties and seventies, including spelling, homophones, sentence structure elements, inflected endings, and others (157-59). Sloan notes that the effects of an oral—and we would stress, an *electronic*—culture on literacy skills are subversive. Students who do not read the "texts" of our culture will continue to come to school without the tacit visual knowledge of written conventions that "text-wise" writers carry with them effortlessly. Such changes in literate behavior have and will continue to affect us in multiple ways, including the ways we perceive, categorize, and judge "errors."

Finally, we feel we can report some good news. One very telling fact emerging from our research is our realization that college students are *not* making more formal errors in writing than they used to. The numbers of errors made by students in earlier studies and the numbers we found in the 1980s agree remarkably. Our findings chart out as follows:[11]

Study	*Year*	*Average Paper Length*	*Errors per Paper*	*Errors per 100 words*
Johnson	1917	162 words	3.42	2.11
Witty & Green	1930	231 words	5.18	2.24
Ma & Pa	1986	422 words	9.52	2.26

The consistency of these numbers seems to us extraordinary. It suggests that although the length of the average paper demanded in freshman composition has been steadily rising, the formal skills of students have not declined precipitously.

In the light of the "Johnny Can't Write" furor of the 1970s and the sometimes hysterical claims of educational decline oft heard today, these results are striking—and heartening. They suggest that in some ways we *are* doing a better job than we might have known. The number of errors has not gone down,

but neither has it risen in the past five decades. In spite of open admissions, in spite of radical shifts in the demographics of college students, in spite of the huge escalation in the population percentage as well as in sheer numbers of people attending American colleges, freshmen are still committing approximately the same number of formal errors per 100 words they were before World War One. In this case, not losing means that we are winning.

Epilogos

Our foray into the highways of research and the byways of the Pentecostal Youth are over for a time, and we are back on the farm. From our vantage point here on the porch, we can see that this labor has raised more questions than it has answered. Where, for instance, *do* our specific notions of error come from? Can we identify more precisely the relationship among error patterns in written student discourse and other forms of discourse, especially the mass media? Could we identify regional or other variations in error patterns? How might certain error patterns correlate with other patterns—say age, gender, habits of reading, etc.? How might they correlate with measures of writing apprehension, or the "ethos," the ideology of a specific curriculum? Most provocatively, could we derive a contemporary theory of error which would account for the written behaviors of all our students as well as the marking behavior of teachers? These are a few of the problems we'd like to fret over if and when we decide to take to the research road again.

Notes

1. As an example of shifting perceptions of student error patterns, it is worth noting that Charles T. Copeland and Henry M. Rideout, writing in 1901, identified the most serious and common grammatical error in Harvard freshman papers as a confusion of the rules for use of "shall" and "will" to express futurity (71n).

2. For a list of most of these studies, see Harap 444-46.

3. We wish here to express our gratitude to the College Division of St. Martin's Press, which graciously offered respondents a choice from the St. Martin's trade book list in exchange for 30 or more teacher-marked student papers or xeroxes of student papers. We are especially grateful to Nancy Perry, Marilyn Moller, and Susan Manning, without whose help this research could never have been accomplished. From assistance with mailings to the considerable tasks of paper stacking, stamping, sorting, and filing, they made the task possible. Their support, both institutional and personal, is deeply appreciated.

The demographics of the papers we were sent were interesting, as we found when examining them for our stratified sample. After pulling all the papers that were illegible, or were not undergraduate papers, were too short to be useful, or were clearly papers from ESL courses, we were left with 19,615 papers. We divided up the U.S. into seven fairly standard geographical regions:

(1) Northeast
(2) Southeast
(3) Midwest
(4) Mid-South

(5) Plains States
(6) Southwest (including Hawaii)
(7) Northwest (including Alaska)

Here are the raw numbers of how the papers were distributed as they came in to us:

Region	1	2	3	4	5	6	7	Total
Total number of papers	3,652	3,478	3,099	4,974	1,229	2,292	891	19,615
Total number of teachers	61	51	54	55	18	47	14	300
Total number of 4-year schools	47	35	40	39	14	24	7	206
Total number of 2-year schools	14	16	14	16	4	23	7	94
Total number of state schools	44	49	48	48	18	44	13	264
Total number of private schools	17	2	6	7	0	3	1	36
Number of schools with total enrollment under 1,000	2	2	0	1	1	1	1	8
Enrollment 1-3,000	9	13	7	11	3	5	4	52
Enrollment 3-5,000	13	5	5	14	2	7	2	48
Enrollment 5-10,000	19	9	16	10	6	7	4	71
Enrollment 10-20,000	14	9	13	13	1	15	2	67
Enrollment over 20,000	4	13	13	6	5	12	1	54

4. We wanted to find out whether the sample of papers we had received mirrored the demographic realities of American higher education. If it did not, we would have to adjust it to represent the student and teacher populations that were really out there.

When we looked at *The Digest of Education Statistics*, we found that some of our numbers approximated educational statistics closely enough not to need adjustment. The breakdown between 4-year colleges and 2-year colleges, for instance, is 71%/29% in the statistical tables and 69%/31% in our sample. The state schools/private schools ratio is statistically 79%/21%, while our sample ratio was 88%/12%, but the over-representation of state schools did not seem serious enough to worry about for our purposes. In terms of enrollment, we found middle-sized schools slightly over-represented and very small and very large schools slightly under-represented, but in no case was the deviation more than 7% either way:

	% of students nationally	% in sample
Number of schools with total enrollment under 1,000	4	2
Enrollment 1-3,000	11	17
Enrollment 3-5,000	13	16
Enrollment 5-10,000	21	24
Enrollment 10-20,000	25	22
Enrollment over 20,000	25	18

We found the most serious discrepancies in the regional stratification, with some regions over- and others under-represented.

Region	1	2	3	4	5	6	7
% of students nationally	23	12	23	15	4	19	4
% of students in sample	19	18	15	25	6	12	5

On the basis of the regional discrepancy we found, we decided to stratify the sample papers regionally but not in any other way.

For help with the methodological problems we faced, and for advice on establishing a ran-

dom stratified sample of 3,000 papers, many thanks to Charles Cooper. When the going gets tough, the tough go ask Charles for advice.

5. These two examples of old-time error patterns are cited in Pressey and in Johnson.

6. The term "comma fault" was by far the most popular term to describe this error pattern until the ubiquitous *Harbrace* seeded the clouds with its terms in 1941, advancing "comma splice," previously a term of tertiary choice, into a primary position by 1960. See Lunsford, Glenn, and Connors, "Changing Pedagogical Nomenclature," forthcoming when we can all stop panting.

7. This paper, five lovingly-written pages of classic Victorian pornography, was extremely popular with the raters. Example passage: "Tammy's own arousal came with suddenness. Bill's urgent caresses kindled a delicious warmth in her flesh and then a melting trembling heat." We would quote more, but we're prudes, and this is a family magazine. For an original xerox copy of this extremely interesting piece of pedagogical history, send $25.00 and a plain brown self-addressed envelope to the Ma and Pa Kettle Go To Waikiki Fund, c/o this magazine.

The teacher's comment on this paper, incidently, was curt. "This is narration," wrote the teacher, "Sorry you didn't use analysis to explain. Remember the definition of explanatory prose?" Another kick in the teeth for Art.

8. In addition to the error-rating sheets, on which the raters kept track of errors found and errors marked, we asked them to write down on a separate list every misspelled word in every paper they saw. This spelling research is only partially tabulated and will be presented in another study.

9. We were also intrigued to find that of the 3,000 papers examined, only 276 had been marked using the letter-number system of any handbook. Handbooks may be widely used, but fewer than 10% of out papers relied on their systems. The rest had been marked using the common symbols and interlinear notes.

10. With our spelling research partially tabulated at this point, we are struck by the prevalence of homophone errors in the list of the most commonly misspelled words. The growth of *too/to* and *their/there/they're* error patterns strongly suggests the sort of problem with visual familiarity suggested by our list of non-spelling errors.

11. These comparisons are not absolutely exact, of course. Johnson counted spelling errors, while Witty and Green and we did not. The numbers in the chart for Johnson's research were derived by subtracting all spelling errors from his final error total.

Works Cited

Copeland, Charles T., and Henry M. Rideout. *Freshman English and Theme-Correcting at Harvard College*. Boston: Silver, Burdett, 1901.

Elbow, Peter. Unpublished document. English Coalition Conference. July 1987.

Harap, Henry. "The Most Common Grammatical Errors." *English Journal* 19 (June 1930): 440-46.

Hodges, John C. *Harbrace Handbook of English*. New York: Harcourt, Brace & Co., 1941.

Johnson, Roy Ivan. "The Persistency of Error in English Composition." *School Review* 25 (Oct. 1917): 555-80.

Pressey, S. L. "A Statistical Study of Children's Errors in Sentence-Structures." *English Journal* 14 (Sept. 1925): 528-35.

Shaughnessy, Mina P. *Errors and Expectations*. New York: Oxford UP, 1977.

Sloan, Gary. "The Subversive Effects of an Oral Culture on Student Writing." *College Composition and Communication* 30 (May 1979): 156-60.

Snyder, Thomas D. *Digest of Education Statistics 1987*. Washington: Center for Education Statistics, 1987.

Williams, Joseph. "The Phenomenology of Error." *College Composition and Communication* 32 (May 1981): 152-68.

Witty, Paul A., and Roberta La Brant Green. "Composition Errors of College Students." *English Journal* 19 (May 1930): 388-93.