ENGLISH AS A SECOND LANGUAGE WRITING REVISITED:
GRADING TIMED ESSAY RESPONSES
FOR OVERALL QUALITY AND GLOBAL ASSETS

By

JEANNA HOWELL OJEDA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004
This dissertation is dedicated
in loving memory

to my mother
Faye Howell Stone

to my father
Edward Leon Howell

and to my husband and children
Fernando, Adriana, and Carlos Ojeda.
ACKNOWLEDGMENTS

This dissertation was completed largely due to the generous assistance of many. My deepest respect and indebtedness go to my committee chair, Dr. Anne Wyatt-Brown, a powerful role model and mentor for me intellectually. A truly skillful communicator, she was always mindful to balance her incisiveness and criticisms with wisdom and inspiration. She taught me volumes about the art of expressing meaning in writing, and about the ever-challenging task of responding to the writing of others.

Gratitude goes to the special committee member, Dr. Willa Wolcott, an invaluable source of knowledge regarding the principles and practices of writing assessment. Her guidance, diligence, and wherewithal in all matters related to the setup and execution of the scoring experiment, and the analysis and reporting of results, were paramount to successes herein.

Profound thanks are offered to the two other committee members, Dr. Diana Boxer and Dr. Joaquim Camps. Their teachings in Linguistics over the years solidly anchored this project. What’s more, their ability to connect research to central themes in second language teaching and learning continues to inspire me to improve my own abilities in these areas.

To the Manager of Student Advisement and Registration Systems in the college where the experimental portion of this study was conducted, I am most thankful. Through his generous permission I was able to gather, from three campus sites of this
college, ESL applicants’ timed essay tests, which comprised the dataset for the two training and scoring sessions.

Thanks also go to two English as a Second Language colleagues for helping me select the essay tests to use as training papers for the scorings. Their expertise in placement and curricular matters, and in evaluating the proficiency level of ESL writing, was an invaluable aid in choosing the necessary representative essays across the pertinent range of levels.

I owe special gratitude to the six college English faculty members and the former English chair—the participants in the trainings, scorings, and post-scoring discussions of the experiment. Their graciously volunteering their time and expertise to the reading, evaluating, and reflecting upon timed ESL essay tests was central to this undertaking.

Without the ever-present love and support of my husband, Fernando Ojeda, this project would not have come to fruition. He contributed in a thousand large and small ways, steadfast in his involvement in all household, childrearing, and professional matters. His uncanny ability to perform many tasks at once was an inestimable help.

Heartfelt thanks go to my children, Adriana and Carlos, for their patience and understanding throughout this research work. Watching them each discover ways to turn their own curiosities into projects, which they then tackled and shared with the family, offered continual and refreshing reprieves. Their love and energy were ever sustaining.

Completion of this work was due in large part to the generosity of my parents-in-law. Ana and Carlos Ojeda have been dependable and tireless in helping us care for our children for years, and through the drudgery of moving our place of residence four times. To them I owe a tremendous debt.
Others to whom I owe thanks are the members of my own birth family and their significant others. Despite the physical distance separating us, the connectedness these individuals maintained always comforted and strengthened me in the most apprehensive of times. Particular thanks go to my mother, Faye Stone, may she rest in peace; my father, Ed Howell; and my sisters, C.C. Winslow, Janice O’Dell, and Elaine Hodges.

Deserving special mention is Michele Houser, a trusty friend whose enthusiasm and confidence in my abilities helped me overcome various obstacles. And I thank my good friends Keith and Suzanne Lindley, whose belief in me was parallel. Enduring their own personal hardships to achieve high goals, all three of these individuals strengthened me with their words and by example, and reminded me to keep focused on the process at hand and the gains to come.

Finally, thanks go to all my English as a Second Language students, to whom my dedication is eternal. My hope is that this project helps me and other ESL teachers better prepare these learners to write timed essays that will meet the expectations of non-ESL college teachers.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................................................. iv

LIST OF TABLES ............................................................................................................. x

LIST OF FIGURES ......................................................................................................... xi

ABSTRACT ................................................................................................................... xiii

CHAPTER

1 INTRODUCTION ...................................................................................................... 1
   The Nature of the Problem ........................................................................................ 1
   Purpose of the Study ............................................................................................... 2
   Background and Rationale for the Study .............................................................. 2
   English as a Second Language Writing: Focusing on Errors and Failures .......... 2
   The Need for New Pedagogical Approaches ....................................................... 3
   Rationale .................................................................................................................. 3
   Significance of the Study ....................................................................................... 4
   Definitions, Terms, and Procedures .................................................................... 4
   Holistic and Analytic Scoring .............................................................................. 4
   The Chief Scorer and the Graders: Non-ESL English Teachers ......................... 6
   The Scoring Experiment ...................................................................................... 6
   Post-Scoring Debriefing ....................................................................................... 7
   Plan of Chapters ................................................................................................... 7

2 REVIEW OF LITERATURE ....................................................................................... 9
   Second Language Acquisition ............................................................................ 9
   Input: The Language the SL Learner Receives ............................................... 10
   Two Distinct Registers of English .................................................................... 11
   Classroom Instruction .......................................................................................... 14
   An undeniable aid in the acquisition of CALP and difficult TL aspects .......... 15
   How noticing register differences develops CALP ........................................... 16
   Writing Assessment: Tests, Training, Scoring, and Expectations ...................... 18
   College-Entry Writing Tests: A Brief History .................................................... 19
   Writing Test Prompts ......................................................................................... 20
   Training and the Use of Holistic Scoring ........................................................... 20
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Scoring criteria for the college basic skills exit test</td>
<td>23</td>
</tr>
<tr>
<td>3-1</td>
<td>Holistic scoring guide—descriptions of ratings for the ESL writings</td>
<td>44</td>
</tr>
<tr>
<td>3-2</td>
<td>Discrepant holistic score pairs—splits requiring resolution</td>
<td>52</td>
</tr>
<tr>
<td>3-3</td>
<td>Acceptable holistic score pairs that need no resolution</td>
<td>53</td>
</tr>
<tr>
<td>3-4</td>
<td>Analytic scoring guide—indicators of global writing competence in an ESL timed written response</td>
<td>55</td>
</tr>
<tr>
<td>3-5</td>
<td>Discordant analytic scores—considered in light of holistic scores</td>
<td>59</td>
</tr>
<tr>
<td>4-1</td>
<td>Eleven holistic 4s (essays scored college-ready)</td>
<td>64</td>
</tr>
<tr>
<td>4-2</td>
<td>Mean holistic and analytic scores—for each grader and for all graders</td>
<td>76</td>
</tr>
<tr>
<td>4-3</td>
<td>Accord on papers graders gave one or two holistic 4s (includes splits)</td>
<td>79</td>
</tr>
<tr>
<td>4-4</td>
<td>Comparison of a holistic 4 (ROYY) and a lower-scored paper (IIKI)</td>
<td>92</td>
</tr>
<tr>
<td>4-5</td>
<td>Discord among six grader scores—for all pre-scored level-3 samples</td>
<td>107</td>
</tr>
<tr>
<td>B-1</td>
<td>Forty-nine essays earning holistic scores of 3 / 2 / 1</td>
<td>132</td>
</tr>
<tr>
<td>C-1</td>
<td>Method for identifying and resolving discrepant holistic score-pairs</td>
<td>134</td>
</tr>
<tr>
<td>D-1</td>
<td>Descriptions of the placement and exit paragraph ratings</td>
<td>135</td>
</tr>
<tr>
<td>E-1</td>
<td>Fifteen splits—and their rectified score-pairs</td>
<td>137</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1</td>
<td>The beginning and end of essay ZLDA (given holistic 4s by R2 and R4)</td>
<td>65</td>
</tr>
<tr>
<td>4-2</td>
<td>Essay KAKE (given holistic 4s by R1 and R4)</td>
<td>66</td>
</tr>
<tr>
<td>4-3</td>
<td>Focus-analytic scores and the first two sentences (of the 11 holistic 4 papers)</td>
<td>67</td>
</tr>
<tr>
<td>4-4</td>
<td>Essay HOSE, a split resolved to a holistic 4 (given a 4 by R1, a 3 by R6)</td>
<td>70</td>
</tr>
<tr>
<td>4-5</td>
<td>Essays KKMA and KAOA (given holistic 4s by R1, R6, R2, R5)</td>
<td>72</td>
</tr>
<tr>
<td>4-6</td>
<td>Essay HIAR, a split resolved to a holistic 4 (given a 3 by R3, a 4 by R6)</td>
<td>75</td>
</tr>
<tr>
<td>4-7</td>
<td>The six graders’ holistic-rank scores for their 20 essays (of the 60)</td>
<td>78</td>
</tr>
<tr>
<td>4-8</td>
<td>Focus-analytic scores from grader R5 (and first two sentences)</td>
<td>83</td>
</tr>
<tr>
<td>4-9</td>
<td>Essay IIKI (given holistic 3s by R3 and R5)</td>
<td>91</td>
</tr>
<tr>
<td>4-10</td>
<td>Essay ROYY (given holistic 4s by R1 and R6)</td>
<td>91</td>
</tr>
<tr>
<td>4-11</td>
<td>Essay ENNT (given a holistic 3 by R2, a holistic 2 by R4)</td>
<td>94</td>
</tr>
<tr>
<td>4-12</td>
<td>Training essays KAKA and PAII (split-scored, resolved to holistic 3s)</td>
<td>95</td>
</tr>
<tr>
<td>4-13</td>
<td>Essays OAPA and GACS, splits resolved to holistic 3s (OAPA given a 4 by R5, a 3 by R2; GACS given a 4 by R6, a 3 by R1)</td>
<td>98</td>
</tr>
<tr>
<td>4-14</td>
<td>Essays LUHA and PYMA, splits resolved to holistic 2s (LUHA given a 4 by R6, a 2 by R1; PYMA given a 4 by R6, a 2 by R3)</td>
<td>99</td>
</tr>
<tr>
<td>4-15</td>
<td>Essay TNGA, a split resolved to a holistic 3 / 2 (given a 4 by R5, a 3 by R2)</td>
<td>101</td>
</tr>
<tr>
<td>4-16</td>
<td>Essay HZJE (given holistic 2s, but analytic 4s and 3s by graders R2 and R4)</td>
<td>102</td>
</tr>
<tr>
<td>4-17</td>
<td>Four additional beginning-ESL level essays</td>
<td>104</td>
</tr>
<tr>
<td>4-18</td>
<td>Training 197-word essay PAMA (pre-scored holistic 4)</td>
<td>106</td>
</tr>
<tr>
<td>A-1</td>
<td>Essay CUKS, 267 words, a best-scoring essay, scored by R2 and R5</td>
<td>126</td>
</tr>
</tbody>
</table>
A-2  Essay ZLDA, 361 words, a best-scoring essay, scored by R2 and R4 ..................127
A-3  Essay KNGK, 487 words, a best-scoring essay, scored by R2 and R5 ............128
A-4  Essay ROYY, 159 words, a best-scoring essay, scored by R1 and R6 ............129
A-5  Essay SALO, 198 words, a best-scoring essay, scored by R1 and R4 ..........129
A-6  Essay PTGE, a 219-word second best-scoring essay, split by R3 and R6 ..130
A-7  Training 275-word essay GKUA, pre-scored holistic 4 ..............................130
A-8  Training 195-word essay OIVM, pre-scored holistic 4 ..............................131
Recent research about second language acquisition, writing assessment, and second language writing offers much that teachers of college-bound English as a second language (ESL) learners can use. This study first presented selections of this research to illuminate current wisdom about one of the greatest challenges ahead for these learners. To enter a college-credit degree-seeking program, they will have to write well enough, under a time limit and test conditions, to convince non-ESL graders of their college readiness.

Next, this study investigated a two-pronged, seldom-studied phenomenon—overall quality and global strengths in college-ready ESL timed writings. To obtain results, an experienced chief scorer first trained six college English professors to score 60 ESL timed essays upon a holistic reading, ranking at the top all that resembled entry-level college freshman composition drafts. Holistic scores ended in 45 agreements and 15 disagreements, an acceptable .75 interrater reliability. After disagreements were
resolved, the 60 essays were divided into four writing proficiency levels at these rank scores: 11 essays at the top score of 4—college-ready; and 49 essays at 3, 2, and 1—advanced, intermediate, and beginning ESL, respectively. Second, the chief trained the graders to rank score 4 global areas—focus, development, organization, and clarity. The graders read the 60 essays more deliberately, considered the 4 global areas in turn, and assigned each an analytic rank score. Discoveries about the global strengths of the 11 essays that these professors scored college-ready suggest that teachers can guide ESL learners to achievable goals. A subsequent post-scoring question-and-answer session enriched the scoring results by contributing these English teachers’ reactions to ESL writing in general, and to some of these timed essays in particular.

Results suggest that if ESL teachers used holistic and analytic scoring methods, they could give their learners more frequent writing opportunities and more useful feedback about overall quality and global strengths and weaknesses. Findings also suggest that more studies obtaining college English teachers’ reactions to ESL timed writings for college entry would help build and illustrate definitions of the top and lower proficiency levels of such writing, from a non-ESL perspective.
CHAPTER 1
INTRODUCTION

The Nature of the Problem

Large numbers of struggling college preparatory English as a second language (ESL) students, whose writings fail, feel that the writing instruction they receive falls short of their needs. In turn, teachers desperately need writing pedagogy that works. Not uncommonly, a teacher of ESL students can feel frustrated or discouraged by the complex task of guiding and responding to their writing. Frequent or voluminous writing can leave her feeling overwhelmed by the assortment and number of weaknesses and errors. Many will grapple and experiment for years with how much and what kind of writing guidance and feedback to give. And the occasional, more traditionally-minded teacher, convinced—by her students or otherwise—that every error on every paper has to be marked or coded, eventually realizes that this is not only labor-intensive and mind-numbing, but ultimately, an unproductive use of her time.

Simply put, the dilemma for ESL teachers of college-bound learners is how to guide and respond to student writings—without concentrating on error correction—so that these learners can meet English teacher and test grader expectations for longer, more sophisticated essays. Approaches need to be effective and not time-intensive.

Beleaguered teachers lack the time to search for solutions. Yet equipped with more knowledge, they could better help students prepare for the imminent hurdles. An informed teacher could give students practice in responding to moderately-specified prompts—the most typical in many large-scale assessments. Similarly, she could train
students to anticipate and write to meet key rater expectations—organization and wording that conform to rhetorical patterns, and framing and content that establish and fulfill the purpose their essay thesis sets out.

**Purpose of the Study**

The present study sets out to answer one primary question: To what extent is passing caliber entry-level college writing determined by the specific global qualities of focus, development, organization, and clarity, in a timed essay written by an ESL learner? Findings come from 2 different scorings of 60 ESL timed writings by 6 non-ESL college English teachers. In 1 scoring, graders attend to the overall quality of the writing. In the other, they attend to strengths and weaknesses in the 4 global areas.

This research first provides a backdrop of practical information useful for teachers of ESL college-bound learners, in 3 areas: second language acquisition (SLA) with respect to essay writing, writing assessment, and second language (SL) writing pedagogy. Recommendations from this study are most relevant for said teachers, many of whom need more help in equipping SL learners to produce passing-level, or better, writing for non-ESL college-credit courses.

**Background and Rationale for the Study**

**English as a Second Language Writing: Focusing on Errors and Failures**

An increased number of ESL applicants during the 1970s forced universities to institute essay exams for admissions. Since that time, ESL errors have captured and held our attention (Santos, 1998; Silva, 1993). However, the unfortunate reality, now conclusively established by research studies on SL and first language (L1) writing alike, is that the approach that many writing practitioners have taken for years, stressing error correction, does not help improve subsequent essays (Janopoulos, 1993; Leki, 1990;
The Need for New Pedagogical Approaches

Rather than criticize surface features, these experts argue that we need to attend to the development of composing skills. Other compositionists echo the call for teachers to attend less to error correction. While he instills in his student writers a respect for correctness and requires that important final drafts be polished, Elbow insists that learners start writing at the beginning of the course without a concern for mechanics, and contends that all writers can best assure that their composing skills will develop by practicing 2 diametrically opposed cognitive abilities, at different instances (1995, p. 397-399). They must “generate copiously” and “criticize cuttingly,” until they develop the ability to produce “writing worth reading.” To criticize cuttingly entails “reseeing, reshaping—not just editing.” In essence, Elbow believes that learning to write is experiential, and a student’s progress largely depends on how much she writes and refines her writings. Furthermore, during neither the writing nor the refining should the student be primarily concerned with surface features.

Rationale

Consensus thus holds that errors, per se, do not constitute the central issue. Yet ESL instructors and students continue to over-attend to grammatically and mechanically correct language at the expense of satisfactory overall composing. While these issues need to be dealt with, the place to do it is not usually on a first draft, many of which also have weaknesses in focus, development, and organization.

Two recent decades of research have borne out what experienced mainstream and ESL college professors have suspected all along. To L1 scorers, adult SL writing is
fundamentally weaker—shorter and less sophisticated—than L1, aside from being more full of errors (Silva, 1993). English as a second language students have considerable difficulty with writing tests (Thompson, 1990), and high numbers fail college entrance writing exams (Song and Caruso, 1996).

Significance of the Study

College preparatory ESL writing teachers lack specifics about the degree to which their learners’ essays meet the expectations non-ESL college English teachers hold. Results of the present study’s scorings of 60 ESL timed essays by 6 non-ESL college writing teachers effectively help meet this need. Rankings of the essays based on their overall quality offer representative samples at different ESL writing proficiency levels. Scorings of the global areas of focus, organization, development, and clarity provide insights about strengths and weaknesses.

Findings suggest that ESL teachers who applied less time-intensive methods for responding to student drafts would be able to attend more to global composing elements in early drafts. The present research embodies the view of the widely accepted process approach, that the task of error correction is better reserved for later drafts. Teachers of SL learners could also benefit from pertinent information provided by this study’s literature review: 1) connections between SLA and the writing of timed essays; 2) particulars and practices of college-entry writing tests; and 3) practices in SL writing pedagogy to improve learner writing.

Definitions, Terms, and Procedures

Holistic and Analytic Scoring

Holistic and analytic scoring—both widely accepted methods—have been used together in this study so as to produce useful findings about the organization and content
of the writings, by avoiding an unproductive and undue emphasis on surface errors. By
the holistic scoring approach, graders read to evaluate an essay as a whole—without
focusing on or overemphasizing any 1 particular part or aspect. Despite disagreements
that arise about the scorings of some papers, this has become the standard way of scoring
gatekeeping writing tests in large-scale writing assessments (White, 1994). And White
points out that the careful training of raters to score holistically, on national writing
assessments, brings the group of evaluators to a consensus in scoring, which contributes
to scores held largely consistent and accurate.

When grading holistically in this study, a faculty evaluator independently and
quickly read a paper once from beginning to end, determined and scored its overall
quality. This method was also appealing for its familiarity. In essence, papers ending up
with the same holistic scores became representations—each a unique case—of a given
writing proficiency level. College-ready level plus 3 lower groups—advanced,
intermediate, and beginning ESL—constitute the 4 rank levels.

Primarily because it is more cost-effective, holistic scoring is the preferred grading
method for large-scale writing assessments. Yet my limiting the scoring of the ESL
essays to the holistic method would have given no indication of features successfully
meeting grader expectations. The careful, if time consuming, analytic method provides
as reliable scores on direct writing tests as does the holistic. It further provides
indications of points of quality and weakness in papers, so its resulting scores can clarify
the more general holistic results. Analytic scoring was thus also included to provide
some specifics from graders about noticeable global strengths and weaknesses.
When grading analytically, an evaluator independently but more deliberately read and scored the quality of 4 global skills—focus, organization, development, and clarity of expression. The higher the perceived frequency of a global skill—“all of the time,” “most of the time,” “some of the time,” “not at all”—the higher the score that skill merited—4, 3, 2, 1. This method was especially useful in that the resulting global-skill scores provided particulars about midlevel papers—supporting details beyond the single-point overall score those papers had received in the holistic grading. These global scores also offered a means for comparing papers on the 4 global aspects.

The Chief Scorer and the Graders: Non-ESL English Teachers

Graders were selected from among 1 junior college’s English professors with experience grading college entrance writings. Each of 6 teachers, as well as a chief scorer, all cleared through the following screening criteria: 1) had experience teaching college English composition; 2) had training and experience in scoring the writing test for entry to college English; 3) had neither taught nor taken a class in ESL / linguistics; and 4) did not speak, read, or write fluently in any language other than English. The evaluations from this group are thus practical and realistic determinations of the college readiness of the ESL timed writings in the dataset, for this college environment and for similar environments.

The Scoring Experiment

The experimental portion of this research study was designed to determine which of 60 ESL timed writings for college placement are of college-ready caliber. The project sought particulars in the form of global qualities, evident to the graders, in this subset of papers. These may contribute to students passing college-entry writing tests.
The experiment was organized as follows. For about 3 ½ hours, the chief reader trained and supervised the 6 college English writing teachers to score the essays in 2 passes, by using 2 widely accepted scoring approaches. First, they scored holistically, to determine which writings embodied an overall passing quality. Second, they scored analytically, to determine the level of quality in each paper, for 4 global skills—focus, organization, development, and clarity of expression. After the 2 scorings, the chief led the group of teachers, all together, through a verbal question-and-answer session, for about 30 minutes.

Of particular note, the chief is a retired community college English professor and department chair with a great deal of training and experience in scoring a statewide mandated college-level essay test. Having him as their guide and leader guaranteed that the scorers adhered to customary evaluating techniques, and that scorer evaluations and commentary typified those of English professors and scorers in similar environments.

Post-Scoring Debriefing

The question-and-answer debriefing session was included to obtain these English teachers’ opinions and perceptions about ESL writing in general, and about this study’s selected writings in particular. This commentary provides additional substance to enrich the scoring data, and reveals both corroborations and conflicts—in grader reports of their views about ESL writing, as compared to their reactions in the form of holistic and analytic scores for the ESL writings they encountered in this study.

Plan of Chapters

Chapter 2 provides a review of the literature: SLA with regards to the writing of academic essays, writing assessment, and SL writing pedagogy. Chapter 3 specifies materials and methods. In Chapter 4, results of the experiment are presented and
summarized. Chapter 5 gives a discussion of the results, along with conclusions, implications, and recommendations.
CHAPTER 2
REVIEW OF LITERATURE

This review of literature was narrowed chiefly to include research most helpful in preparing college-bound ESL learners for college-level timed essay writing. The purpose is to establish a practical base for the experimental portion of the present research study. The 3 areas addressed by this review are SLA with regards to writing academic essays, the assessment of writing, and SL writing pedagogy.

Second Language Acquisition

Clearly the most fundamental of the 3 areas in this literature review is SLA as it regards writing essays, which is thus the most natural starting point. Second language acquisition has many variables, and formal classroom instruction is only a single contributing factor. The first language, the linguistic environment, reading, noticing register differences, and communicative interaction, among other elements, combine to play a role in the learner’s developing SL system. It is well established by now that learners acquire the SL in a series of stages, and formal instruction is powerless to enable a learner to skip a stage in this progression (Lightbown, 1998), though such instruction can help to speed acquisition.

Studies have shown that, in general, regarding overall target language (TL) proficiency and accuracy at any given level, the performance of adult English as a foreign language (EFL) learners\(^1\) tends to be more accurate—more grammatically correct—while

\(^1\) EFL learners study English while living in predominantly non-English speaking surroundings.
that of ESL learners\(^2\) tends to be more communicative—more fluent (Ellis, 1994). English as a foreign language teachers emphasize grammatical accuracy, but ESL teachers and native language surroundings supply a richer environment, which contributes to the more fluent SL learner output (Sharwood Smith, 1986). This fluent language, however, may not be grammatically correct.

**Input: The Language the SL Learner Receives**

Inside the classroom and out, the linguistic input—the language that SL learners are exposed to—is surprisingly distinct from what linguists and grammarians, and even native speaker intuitions, long held (Williams, 1989). It is comprised of myriad forms, including incomplete, redundant, simplified and reduced stretches of English containing grammatical errors. Significantly, a parallel can be drawn between the writing and speaking of SL learners and the fluent variety of L1 input the learner receives. Both include stops and restarts, reformulations, overlap, and ungrammaticalities (Ellis, 1994).

Second language acquisition experts widely agree: Even when it includes the less grammatical unplanned variety just described, more input tends to facilitate acquisition. Of great benefit to learners, then, are simplifications and expansions of complex structures and vocabulary, which they often receive in the classroom and out. These permit them to use their mental processing capacities toward accessing and understanding communicative meanings, rather than for deciphering and unraveling what is beyond their grasp. Not uncommonly, teachers and native speakers modify language for beginning and intermediate learners to aid their comprehension and communication. For instance, we regularize forms—produce full forms over contractions—and we elaborate

---

\(^2\) ESL learners live and study English in an environment wherein the primary language for communicating is English.
to clarify meanings—add vocabulary and cultural items to enrich the context. English as a second language textbooks also provide comprehensible input, in text formulated to accommodate the language level of the learner.

For several decades now, SLA experts have urged instructors to offer learners diversified input, and to help them integrate all aspects of language use in a communicatively complex way—guidance that contrasts sharply with traditional instruction focused on linguistic forms (DeKeyser, 1998; Sharwood Smith, 1986). The view is that when real language is provided in a whole language approach, SLA is enhanced. Through collaboratively communicating with others in the classroom, students can actively arrive at and learn information and ideas (Rigg, 1993).

**Two Distinct Registers of English**

It is important, here, to define 2 distinct registers of English, with corresponding proficiencies we commonly refer to as BICS and CALP. The former is the more rapidly acquired “basic interpersonal communicative skills” in English—everyday language and communicative contexts, also called “oral fluency” (Blanton, 1993; Cummins, 1985). Basic interpersonal communicative skills are situated within a context-rich communication—full of paralinguistic and situational cues. Included among the more informal manifestations of BICS are fragmented sound bites—such as those commonly used in advertising—and units containing ungrammaticalities and slang, often found in music lyrics.

On the other hand, CALP is the less rapidly acquired variety of English, the so-called “cognitive / academic language proficiency.” Cognitive / academic language proficiency is based on lingual knowledge and cues in print alone (Cummins, 1985). It is the concatenation of conceptual, linguistic, and academic processes required of college
students—not “intelligence” or “oral fluency” (Blanton, 1993), but the likelihood for college success which institutional placement tests are designed to measure (Larsen-Freeman and Long, 1991a). Cognitive / academic language proficiency comprises formal academic prose, characterized by grammatically and syntactically rich sentences organized into paragraphs and essays, and containing complex verbal units and formal polysyllabic vocabulary.

Inadequate CALP—marked by test scores falling below a predetermined cutoff score—indicates a deficiency in reading, writing, or both, and warrants college preparatory coursework. Interestingly, it has been shown that CALP proficiency correlates strongly across languages (Bachman and Cohen, 1998). This is to say that learners with a strong CALP in their L1 are more likely to attain greater competency in CALP in English. Evidence for this assertion is the relative rapidity with which college-educated SL learners acquire English CALP, as compared to their counterparts without a college background.

Second language learners can, and generally do, acquire the BICS and CALP language proficiencies differently. As for BICS, native speakers of English come equipped with this competency, which may be why mainstream college preparatory writing teachers are sometimes heard advising learners with writer’s block to simply “write as if they were talking to someone.” This advice would be impractical, of course, for the SL learners not yet competent in BICS. And yet, “teachers, psychologists and policy-makers frequently consider [that SL learners] can develop a relatively high degree of English” BICS within about 2 years of exposure to “English-speaking peers, television and schooling” (Cummins, 1985, pp. 132-133). Even without formal classroom
instruction, learners may acquire BICS (Larsen-Freeman and Long, 1991a). Indeed, learners have been shown to proceed along the same path of development in their production of a number of morphemes and other language structures—i.e., negatives, interrogatives, and pronouns—with or without formal instruction (Ellis, 1994; Fathman, 1978; Larsen-Freeman and Long, 1991b).

While one can acquire BICS through moderate efforts, one must work diligently, in comparison, to gain competency in CALP. Reading is repeatedly shown to be vital in the development of CALP. Second language acquisition experts recommend teachers give much instructional time and attention to this skill area, contending that unassisted reading is difficult for learners to sustain. And of the learners who are capable, few tend to read extensively outside of class. Studies that Krashen cites show that reading consistently outperforms direct instruction in developing literacy, reading comprehension, writing style, word-meaning connections, grammar, and spelling (1993). Also of note, strong positive correlations exist between SL writing proficiency and SL pleasure reading (Janopoulous, 1986).

Emphasizing the importance of reading to college success, Blanton claims that students who do not read and write well have no chance for academic success (1993). She adds that instructors can teach students to interact with written texts by decoding the information to connect themselves to the text by interacting with it. By recalling one’s own experiences, ideas, and words, one can validate what one reads, and this yields pleasure and long-term benefits. “Comprehension then becomes a cognitive / intellectual interaction between the reader and the text; the deeper the involvement, the greater the comprehension” (p. 238). In fact, she reports that studies have shown that interaction
aids the acquisition of CALP. And others even consider an instructional inattention to forms risky—it might deprive SL learners of the developed grammar and morphology needed for CALP (Swain, 1985; Zalewski, 1993).

**Classroom Instruction**

Empirical studies have shown that to achieve advanced levels of target language, learners may, indeed, need formal instruction (Fotos, 1993). For instance, instruction is reportedly necessary in helping learners avoid simplifying the written system in mimicry of the spoken—omitting past tense inflections and articles, for instance—and a lack of instruction can even result in fossilization (Selinker, 1972). It has even been claimed that full native speaker competence may be impossible to attain without instruction (Lightbown, 1998; Long, 1988; Zalewski, 1993). All experienced teachers have encountered cases of fossilized learners who, despite years of living and communicating in the SL environment, never acquired redundant features such as the plural –s on nouns or the third person singular –s on third person present tense verbs.

Second language acquisition research has also shown that when a learner pays attention to input, it becomes intake—a subset of information that gets noticed then stored in memory, available for subsequent cognitive processing (Doughty, 1991; Schmidt, 1993; Swain, 1993; VanPatten and Cadierno, 1993). Interlanguage (IL) theory, a widely accepted cognitive view of SL acquisition, proposes that learners receive, process, and produce language based largely on their hypothesis testing—that is, based on the constructing and reconstructing of the developing language system (Ellis, 1994; Sharwood Smith, 1986).

Instructors can capitalize on this, by helping the learner to grasp meanings and formulate proper rules of difficult aspects in their developing SL system. Without rich
supplies of meaningful input, attention to language features, and practice in hypothesis forming and testing, SL development may terminate at a basic communicative level, with morphology and syntax that are non-targetlike (Swain, 1998; Zalewski, 1993). But when SL teachers help adult learners grasp the connections between grammatical features and meanings, they are essentially providing learners the crucial access to the language they need.

An undeniable aid in the acquisition of CALP and difficult TL aspects

To reiterate an important point, mere exposure to the spoken language will not assure the acquisition of CALP. Formal education is considered necessary, as demonstrated by the fact that some native speakers also need instruction in academic language. Evidence of the distinctness of English oral fluency from CALP is the existence of developmental classes “full of fluent (native) speakers of English” (Blanton, 1993, p. 237). A teacher can help students understand the particular meaning nuances of academic vocabulary. She can explain complex verbal units—auxiliary verbal forms or modals with main verbs, inflected to signal time—past, present, future—and aspect—progressive, perfective. She can emphasize the importance of developing some mastery of the grammar inflections, complex syntax, and formal vocabulary of CALP.

Optimally, formal instruction offers the learner greater access and exposure to conversational and academic English vocabulary, syntax, and grammar, all of which helps students acquire better language skills. And SL experts and practitioners tend to agree that adults rely on their L1 grammar, the aspects of which interfere with acquiring correct forms of the target language, so attention to some grammatical elements is beneficial. This is particularly true for cases of difficult TL features, such as “problematic overgeneralization,” where the L1 is more general—completely regular in
pattern, without variations—than the SL (Ellis, 1994, p. 657; Schmidt, 1993). Examples are irregular or marked forms—e.g., plurals—and nonexistent aspects in the mother tongue—e.g., articles.

Proficiency in such forms requires emphasis in the classroom, as well as a great deal of the learner’s time, attention, and practice. To learn these features, students need to have them pointed out (Schmidt, 1993; Zalewski, 1993). Most learners need help making sense of the complex English article system. The difficulty of acquisition of articles will be greatest for speakers of native languages without them—most Asian\(^3\) and Slavic\(^4\) languages and many African\(^5\) languages (Celce-Murcia and Larsen-Freeman, 1999; Master, 1995). Learners need help in recognizing patterns of article usage. And unless some rules and meanings are made clear, many learners abandon efforts to use articles. A teacher who increases the meaningfulness of articles improves a learner’s ability to acquire them (Zalewski, 1993). If classroom instruction emphasizes the use of articles through the discussion of errors and the integration of feedback with writing, some learners can produce fewer article errors in due time (Master, 1995).

**How noticing register differences develops CALP**

One important and effective means for developing learners’ CALP is through emphasizing that they mindfully attend to register differences in the English to which they are exposed. Crucially, one who wishes to succeed in college needs to be able to recognize and understand formal academic English. Otherwise, the need to express

---

\(^{3}\) Asian languages include Korean, Vietnamese, and Thai, among others.

\(^{4}\) Slavic languages include Russian, Ukrainian, and Serbian, among others.

\(^{5}\) African languages include Ethiopian, Yoruba, and Somali, among others.
thoughts and opinions in this style, when necessary, will be nearly impossible. Students can be asked to informally paraphrase formally expressed ideas, and vice versa. Formal and informal passages can be presented in class for students to classify by register and defend by pointing out identifying features.

Certainly, if teachers emphasize typical features of formal register, they can improve the ability of their learners to incorporate such features into their own academic writing. But what are typical features? Analyzing the range of verbal productions of native English speaking professors or graduate students at the University of California at Berkeley or the State University of New York at Albany, researchers found clear manifestations of formal register (Chafe and Danielewicz, 1987). Reviewing 4 modes—conversations, lectures, letters, and academic written text—these researchers found the predominance of formal register in only academic written text. English teachers were advised to emphasize the 3 most frequently employed formal features—prepositional phrases, nominalizations, and attributive adjectives (-ed or -ing forms before or after nouns, as modifiers). Student writing that contains frequent and accurate usage of formal features better adheres to academic style, and thus arguably earns higher approval ratings from faculty graders.

Elbow suggests one way to give learners practice in accurately distinguishing and using both the formal and informal registers; teachers and students collaborate to read texts never before seen, pausing periodically to discuss and write perceptions and reactions (1996a). Such spontaneous sharing blends the 2 registers to allow a rich combination of formal and natural input in the classroom, all the while integrating reading, writing, and speaking. Moreover, using all language modes is considered to be
more natural than using one mode in an unnatural separation (Rigg, 1993). Student speech or writing that includes inappropriately formal or inappropriately informal wordings can be addressed in the classroom. Understanding the misuse of register, and knowing how to formulate repairs are key skills most learners need help developing.

**Writing Assessment: Tests, Training, Scoring, and Expectations**

The second area in this review of literature is writing assessment—college-entry writing tests, scoring, training, and expectations—information about a challenging bar held up for SL students, which many fail to clear on one or more attempts. It has been suggested by some that a main cause for ESL writing test failures is examinee ignorance of both the format of the typical college-entry writing test, and the expectations—grammatical, syntactic, rhetorical—of the typical test grader. As Hamp-Lyons (1991a, p. 57) puts it, ESL learners writing for these tests are being asked to “demonstrate readiness for membership in a community with which they may be wholly unfamiliar.”

We all know of cases, as this study suggests, in which teachers of English to L1 and SL learners—whether in high school, in college preparatory, or in ESL courses—are failing to prepare their charges sufficiently for the relatively daunting tasks of timed writing tests and college-level essay writing. Sadly, many are not wholly aware of the expectations themselves. These reports from Hamp-Lyons came mostly from L1 writing research (1991a). Yet, the inopportune lack of L1 learner preparation about college-entry tests and grader expectations suggests that such practical information could “be at least as unfamiliar, and thus more threatening, to second language students” (p. 51).

There exists a great deal of helpful research with informative findings for participants on both sides of the writing assessment issue: 1) for teachers and examinees—detailed discourse expectations inherent in writing tests; and 2) for test
designers, test writers, and test readers—full descriptions of student examinees (Hamp-Lyons, 1991a). Becoming more informed about the examinees has helped the test designers, examiners, and scorers improve assessment instruments, procedures, and scorings in recent years. It follows that college preparatory teachers who become more informed about topic design, testing, and scoring could better guide their students to success in composing passing freshman level writing.

This second section in my literature review provides information about college entry placement testing, as contextualized by several elements: 1) writing test prompts; 2) training and scoring methods; and 3) rater expectations. But to first build a base, a brief history of college-entry writing tests will now be presented.

**College-Entry Writing Tests: A Brief History**

U. S. colleges and universities have always used tests to determine readiness for entry. In the late 1800s, test instruments were added to measure the ability of the applicant to express logical reasoning skills in an acceptable written form.

Hamp-Lyons details the fluctuating use of writing tests in the process of college-entry testing, and our improved ability to effectively test writing directly (1991b). After World War I, standardized multiple-choice tests purporting to measure “writing” supplanted direct writing tests nationwide, despite some educators’ apprehension about a potentially parallel de-emphasis of writing in our college classrooms. Then in the 1970s, colleges and universities reintroduced direct writing tests, and again examined the actual writing of applicants.

As research on direct writing tests and assessment since the 1970s has shown, certain precautionary measures—tested writing prompts, established scoring systems, carefully trained raters—have helped make direct writing test scores a more dependable
component in college-entry and course placement decisions (Enginarlar, 1991; White, 1994; Wolcott, 1998). Now in the 21st century, it is widely accepted and becoming standard practice, nationally, to use direct writing test results in 2 standard examinations for college entry—the SAT (for U. S. high school students) and the TOEFL (for foreign students).

**Writing Test Prompts**

As for the writing prompts of college-entry tests, from years of trial, error, and review, it has been found that moderately specified writing topics (in audience, purpose, content, and voice), not minimally or highly specific ones, result in higher mean scores on student performances (Hamp-Lyons, 1991a). As a rule then, large-scale writing tests designed by professional test writers tend toward moderately specified topics. These appear to provide respondents both guidance and flexibility, allowing them to formulate a response that they value, by way of support and guidance provided within the wording of the prompt. Minimally specified tasks are less effective because they lend themselves well to only those whose personal experience or knowledge, or both, provides them enough relevant background to formulate a response. Highly specified ones present the opposite problem for the examinee. They so rigidly define the writing task that one responding may have little freedom to express a valued point of view or relevant experience—as these appear to be outside of the specifications.

**Training and the Use of Holistic Scoring**

Most national college-entry writing tests are scored holistically, and in every well-managed holistic scoring of writing tests, before the evaluating of student papers commences, there is a training session during which a chief reader guides the raters as they score a set of pre-scored samples (White, 1994). This process serves to effectively
lead the raters into a consensus about how each score point, on the scoring guide at hand, accords with particular sample papers. In theory, training lessens the risk that raters score idiosyncratically, and thus increases the number of uniform evaluations. In practice, rater training is a successful mechanism for building rater consensus, regarding precisely what, in the scoring method being applied to the papers, each score point represents.

Graders are instructed to evaluate a written response as a whole, upon 1 reading, rather than by considering discriminate parts as manifestations of separate skills. They are to read each paper from start to finish in 1 holistic pass, and to attend to meanings and to what the student writer has done well overall—not to the learner’s errors. Trainers sometimes explicitly ask raters not to keep a count of errors as they read; the writing is to be seen as a work in progress, a snapshot of a process—not as a final product.

In preparation for the training, a selection committee of 2 or more chief readers—who will direct the training and scoring—carefully chooses a number of sample papers. The pre-scored samples are student writings—representative of the actual papers the graders will encounter—to be scored and discussed before the scoring of actual papers begins, to build a scoring consensus within the group. They then score them by the guidelines in which the raters will be trained.

The procedure for training holistic scorers is generally the same. Having enough clean copies of the pre-scored samples to distribute to all the assembled raters, the chief begins by having the raters rank order a small number of pre-scored samples. They might start with 6 to 10 papers. They then read the scoring guidelines, and independently apply the scoring method to these samples, delivering their judgment on overall quality in the form of a single point score, on a scale from 1 to 6, for instance.
While they are reading each paper once through and determining appropriate point scores for overall quality for each, the chief is readying the chart on which scoring results will be tallied, on a whiteboard or projected image at the front of the room. This chart has the identifying code for each paper down the left side, and all the possible point scores across the top, from the highest—generally 6 or 4—to the lowest—1. Empty boxes connecting the columns and rows will be filled with counts of rater scores for each paper, at each point score.

When the raters all finish scoring the samples, the chief calls for their scores. By repeated shows of hands, the chief ascertains and records, for each sample, how many raters gave that paper each possible point score. In trainings with many raters, the chief simply puts a check mark inside the box corresponding to the score point that most raters assigned the given paper. In gradings with few raters, the chief counts hands and writes the actual number of raters that assigned each point score in each box. Open discussion ensues, and is especially helpful for resolving issues on problematic papers—those for which no scoring consensus exists, or on which the majority scored too high or too low, in comparison to the pre-scorings done by the selection committee.

The scale of scoring guidelines, also called a scoring rubric, consists of a table of scores, from the highest presented at the top, to the lowest at the bottom. Beside each point score is a paragraph, detailing particular manifestations of proficiency that a writing at that point score tends to exhibit, in the way of organization, development, language control, and so on. The scoring rubric quite often features a 6-point scale. Graders might be trained to consider the upper half scores from 4 through 6 to be passes, (with 6 representing writing of a superior caliber), and scores from 1 through 3, non-passes.
Variations upon this design do occur, however. For instance, although they use a 6-point rubric (Table 2-1), trainers at the junior college where my experimental study was held demarcate passing and failing on this scale differently. Entry to freshman composition is allowed to any learner whose writing receives grades of 3, 4, 5, or 6. Entry is blocked if the writing receives lower scores—1 or 2.

Table 2-1. Scoring criteria for the college basic skills exit test

**Score of 6** The paper has a clearly established main idea that the writer fully develops with specific details and examples. Organization is notably logical and coherent. Vocabulary and sentence structure are varied and effective. Errors in sentence structure, usage, and mechanics are few and insignificant.

**Score of 5** The paper has a clearly established main idea that is adequately developed and recognizable through specific details and / or examples. Organization follows a logical and coherent pattern. Vocabulary and sentence structure are mostly varied and effective. Occasional errors in sentence structure, usage and mechanics do not interfere with the writer’s ability to communicate.

**Score of 4** The paper has an adequately stated main idea that is developed with some specific details and examples. Supporting ideas are presented in a mostly logical and coherent manner. Vocabulary and sentence structure are somewhat varied and effective. Occasional errors in sentence structure, usage, and mechanics may interfere with the writer’s ability to communicate.

**Score of 3** The paper states a main idea that is developed with generalizations or lists. The paper may contain occasional lapses in logic and coherence and is mechanical. Vocabulary and sentence structure are repetitious and often ineffective. A variety of errors in sentence structure, usage, and mechanics sometimes interferes with the writer’s ability to communicate.

**Score of 2** The paper presents an incomplete or ambiguous main idea. Support is developed with generalizations and lists. Organization is mechanical. The paper contains occasional lapses in logic and coherence. Word choice is simplistic, and sentence structure is disjointed. Errors in sentence structure, usage, and mechanics frequently interfere with the writer’s ability to communicate.

**Score of 1** The paper has no evident main idea. Development is inadequate and / or irrelevant. Organization is illogical and / or incoherent. Vocabulary and sentence structure are garbled and confusing. Significant and numerous errors in sentence structure, usage, and mechanics interfere with the writer’s ability to communicate.

---

6 This guide has been used for several years at this junior college, to evaluate the 45-minute timed writing tests of students in advanced ESL and developmental basic writing.

Thus, only with a pair of scores both within the passing range of 6, 5, 4, or 3 would a learner be granted entry to freshman composition. What distinguishes a 3-paper from a 2-paper is the following. The writer of a 3-paper demonstrates the ability to meet 5 expectations: 1) to state a main idea; 2) to convey coherence; 3) to use words above the basic level; 4) to correctly use English sentence structures; and 5) to avoid frequent errors that interfere with clarity of expression. The writer of a 2-paper, on the other hand, fails to demonstrate 1 or more of these abilities.

To reiterate, holistic scoring is the commonly preferred method for evaluating large-scale writing tests. And yet, while it is widely accepted as the most appropriate tool for this purpose, many writing practitioners and administrators reportedly “view the holistic technique with reservation” (Enginarlar, 1991, p. 39). Some may hold this view as a result of a single unfavorable holistic scoring experience. Hamp-Lyons (1996, p. 234) objects to the single point score resulting from holistic grading, which she claims “obscures a pattern of consistent overemphasis or underemphasis on basic language control.” Another claim is that a higher share of midlevel papers seem to end up with split scores—scores 2 points apart that must be resolved by a third grader—so holistic scoring may be less effective for evaluating midlevel performances than an analytic method might be (Elbow, 1996b).

Such objections suggest a trepidation that holistic graders might idiosyncratically assign unduly low scores. For instance, an ESL essay could be primarily weak in rhetorical style. And if it were, a low holistic score might seem unjust, given that other strengths existed in the essay. We have evidence to show that even after years away from
their L1 cultures, SL writers generate rhetorical patterns of prose native to their cultures (Land and Whitley, 1989).

For numerous reasons, the holistic scoring technique is solidly established as the best device for use in large-scale scorings. However, for other purposes, the analytic scoring method can be useful. Most importantly, it can help meet the need that many ESL learners who fail writing tests have, for feedback on both the deficiencies and the qualities in their writing (Kroll, 1990). Let us now turn to the analytic scoring technique.

The Analytic Scoring Approach

While it is no panacea for the scoring issues that the problematic midlevel papers present, analytic scoring can at least give some insight into the particular strengths and weaknesses of such essays. To analytically score, a rater reads a composition methodically, and generally more than 1 time through, to evaluate a number of discriminate features, such as language use, development, organization, and vocabulary. Thus, rather than evaluating the whole writing with a single point score, the rater judges its individual parts, and assigns each a separate score. In this way, the ESL papers most difficult to evaluate can be more methodically evaluated. For instance, if language use were weak, but the other 3 areas were strong, a grader could assign a low score for language use and higher scores for the other areas, notwithstanding language use errors.

One such analytic scoring method, from a research study by Song and Caruso (1996), contained a set of 10 discrete features—6 rhetorical and 4 linguistic—and was applied by 32 English and 30 ESL faculty graders to score 4 written responses. Nonnative speakers had written 2 of the responses, native speakers, the other 2. Among the 6 rhetorical features were a central idea, supports, and transitions. Constituting the 4 linguistic features were fluency, sentencing, diction, and grammar. All 10 features were
rated, individually, on a scale from 1—poor—to 6—superior. Each of the 4 essays contained numerous errors typical to the population of its writer, and also, each was predetermined to be moderately strong in rhetorical features and content. Results showed that no matter the type of errors, ESL and English graders scored these midlevel essays similarly—with no significant differences—on the 10 analytic features. Findings suggest that analytic scoring could result in consistent grading of midlevel essays, and would offer learners useful feedback on the strengths and weaknesses in their writings, besides.

Another analytic scoring method was created by a team of assessment specialists—Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981)—with SL learners in mind. This one is comprised of 5 features to be analytically scored, for a combined total of 100 points—a paper can earn up to 50 points for rhetorical features, up to 50 for linguistic. Feature score maximums are as follows: content, 30; organization, 20; vocabulary, 20; language use, 25; and mechanics, 5. Each feature is subdivided into 4 scoring ranges, at 4 levels of quality. A set of adjectives or short phrases, or both, describes each range. Thus, for example, beside “content” one sees 4 scoring ranges—the top being “30-27,” and the bottom being “16-13.” The top range for content is labeled “EXCELLENT TO VERY GOOD” and described. The bottom range for content is labeled “VERY POOR” and described. This method has been used successfully for some years at one major university I know of, for scoring timed writing tests of nonnative graduate learners.

As anyone who has participated in an analytic group grading, or who has used an analytic scoring rubric to grade the writings of his learners will attest to, analytic scoring is more time consuming than holistic. It is also less straightforward, since there are generally more discrete points to keep in mind when evaluating a paper. Consider the 2
sample analytic methods described above. Requiring that graders score 5 or more
features of every paper, both designs suggest that graders do 1 or more rereadings of
every essay. Nonetheless, the primary advantage of these and other such analytic scoring
methods is that the resulting scores provide learners helpful feedback about strong and
weak aspects in their writings.

One Comparison: Holistic vs. Analytic Scoring

Enginarlar’s comparison study of the holistic and the analytic scoring approaches
showed that, as a measure of overall writing proficiency, the holistic ranked slightly
higher than the analytic in validity, reliability, and practicality (1991). To measure
validity, the holistic and analytic scoring results were examined in light of averaged
classroom-generated test scores for 5 months—for the same learners. In validity, the
correlation of .73 for the holistic scoring was a bit higher than that of .71 for the analytic
scoring. In reliability, the interrater coefficients—marking consensus, or agreement
among scorers—for the analytic method and the holistic approach were nearly identical
(.77 and .78 respectively), and .75 and over is generally considered acceptable. In
practicality, the holistic scoring came out ahead, in that it required less scorer time—no
more than 1 reading of each paper—and less scorer effort—no writing beyond the
recording of a single point score for each paper graded.

---

8 Two types of statistical analyses provided answers—the multiple regression analysis, and the Pearson
product movement correlation coefficients. Results showed that both the holistic and the analytic methods
correlated significantly with the monthly test averages (the p value of each, at .001, was significant at the
.05 level).

9 Scorers were 6 experienced English as a foreign language (EFL) teachers, and essay test examinees were
intermediate level Turkish graduate students. The scorers were not given training prior to the holistic
scoring process; thus, their reliability estimates might have been higher still.
Rater Expectations

Direct tests of writing are widely considered, by now, the most accurate way to measure one’s ability to write. Of great importance with regard to rater expectations—and intuitively sensible—student writers whose texts demonstrate an attention to the topic and an ability to gain ownership of it yield higher mean scores (Hamp-Lyons, 1991a). But many second language learners are unable, unfortunately, to demonstrate such writing finesse under time constraints. In one particular study Hamp-Lyons cites here, discourse-level characteristics were predictive for SL and L1 holistic essay scores, while syntax and lexis figured somewhat more in the judgments of the SL essays. It has become painfully evident to SL learners and their teachers that even when an ESL response adequately frames and expounds the topic, it will probably be failed if evaluators overemphasize mechanics and usage.

Particularly problematic to ESL writers can be the appropriate usage of CALP features, especially in the complex system of suffixing to represent the various grammatical forms of English vocabulary. Having knowledge of BICS and CALP equivalents, native English speakers can write using the BICS forms—e.g., “get,” in place of “acquire”—and avoid spelling errors. On the other hand, some SL learners rely primarily on academic English vocabulary, which may reduce their chances of passing essay tests; characteristics of ESL writing include the inevitability of wrong or missing endings on polysyllabic English words, and errors in the use of subordination.

In one study, morphological errors were higher in number in all the papers, relative to global / syntactic errors—word order, deletion within SVO, sentence combining error, and continuity of tense across clauses, which more often lead to miscommunications (Bardovi-Harlig and Bofman, 1989). Morphological errors include grammatical functors
(e.g., articles, prepositions), inflectional morphology (e.g., nominal plurals, possessives, tense markings), derivational morphology (constituted by prefixed and suffixed units), and lexical-idiomatic errors (word choice).

This study compared the timed written essay tests of 2 groups of advanced learners: 1) learners who were let into academic level English; and 2) learners whose entry was blocked (Bardovi-Harlig and Bofman, 1989). Analysis across 6 first language groups (Arabic, Chinese, Korean, Malay and Spanish) clearly indicated that any given essay contained moderately strong syntax but incomplete and variable morphology, lexis, and grammatical functors (p. 26). The fail group's errors were virtually identical in type to those of the pass group—roughly 17% syntactical, 27% morphological, and 56% lexical-idiomatic. What actually separated the pass group from the fail group was the total volume of errors—not the type. Accordingly, ESL students need to be guided to edit their work and reduce as many of their primary errors as they can—word choice, articles and prepositions, noun plurals and possessives, tense markings and affixes.

Unfortunately for SL learners who write for non-ESL teachers, curricular “standards” sometimes presuppose native-like proficiency. And when they do, they will be nearly impossible to meet. Traditional assessment methods sometimes render exclusionary status quo policies and practices as acceptable (Freire, 1970). For example, writing that is adequate, save occasional errors in prepositions and articles, will inevitably be evaluated by some raters as non-proficient and mediocre, when it in fact deserves to be rewarded for its various qualities (Zalewski, 1993).

Second Language Writing Pedagogy

The third area in this review of literature is SL writing pedagogy—focused mostly on improving timed essay writing. Pedagogical advice and response to SL writing in the
formal ESL classroom setting can be helpful to varying degrees, or at worst, discouraging and unhelpful. To develop their skills in writing, students need regular guidance, practice, and reinforcement. Recommended tasks include: instruction, modeling, and review of sound writing practices; student-teacher conferencing about written works; and the use of checklists to help students revise and edit.

**Written Discourse and Pedagogy**

Second language teachers are advised to draw learner attention to global form-function relationships, from high beginning levels of ESL upward. It becomes increasingly important, as SL students advance in their course levels, to deal with meaning and grammar within units of discourse longer than the phrase, clause, and sentence (Zalewski, 1993). One suggested exercise is to have students cooperate to reorganize lists of sentences into a paragraph, or paragraphs into an essay (Rutherford, 1987b). Through cooperative human interactions, an individual’s “learning awakens a variety of internal developmental processes” that can become a part of this person’s “independent developmental achievement” (Vygotsky, 1978, p. 90). Thus, communicative sharing can lead one’s developmental processes to higher realms.

But since college-level courses measure students’ “intellectual development . . . by their ability to solve problems unassisted,” SL instruction “must prepare them to solve problems on their own” (Zou, 1998). Instruction and practice in paraphrasing and summarizing texts can be used, followed up by such homework. This practice challenges students to demonstrate that they can express understanding of contextualized meanings. Processing the meanings in new vocabulary has also been shown to leave “memory traces” and increase retention (Brown and Perry, 1991).
A similarly useful activity, for emphasizing global form-function relationships, is for the instructor to select an academic written discourse, or an academic lecture, to use for teaching chunks of meaningful ideas. This can be done with a reading passage students have already seen, or a lecture they have heard. While students listen, but do not read, she produces meaningful units, or chunks of information, asking students to mimic them. Afterwards, they can discuss and paraphrase the units. This activity is enjoyed by learners for its challenging nature, and helps in the development of their SL fluency.

**Input Enhancement with Reading**

Instructors who enhance forms in the passages their learners receive are effectively drawing their attention to one or more linguistic features—by emphasizing them (Fotos, 1993). Schmidt argues, “the subjective experience of ‘noticing’ is the necessary and sufficient condition for the conversion of input to intake . . . what must be attended to and noticed is not just the input in a global sense but whatever features of the input . . . [are] relevant for the target system” (1993, p. 209).

Investigating the effect of text modification on the reader’s retention of vocabulary, Watanabe compared 3 text formats: unmodified texts, texts plus appositives that defined words, and texts with specific words glossed in the margins (1997). Students presented with the glosses substantially outperformed the others on tests, indicating that when students are presented texts that incorporate structured aid, they may better learn vocabulary. A simple and useful input enhancement method is to read texts together, with instructional aids. It can help to project an image of a text (in large font) onto a screen and use a laser pointer to call attention to particular units, or just follow along while it is being read aloud.
Pedagogical Advice about Writing

One of the most complex skills for SL learners to develop is the ability to produce acceptable academic prose. Rhetorical patterns, lexis, and grammar and syntax all figure into grader judgments of overall quality. Thus, the more an instructor can do to ensure that her learners internalize the ability to check their essays for overall acceptability, the more likely it is that these learners’ revising skills will improve.

Although it has been shown that learners can notice problems without receiving feedback (Sommers, 1980; Swain and Lapkin, 1995), giving both positive and negative (corrective) instrumental feedback will increase the chances that learners’ revisions incorporate substantive improvements (Corder, 1967). Giving positive and negative feedback may also help some develop the ability to recognize whether their writing is competent, and make it conform to expectations. Some understandable device can be used to train them to review a whole essay to recognize its serious flaws as well as its assets. Checklists are useful for this.

Through the use of checklists, learners can come to write better responses to essay tests. Instructors who provide, explain, and guide students in how to methodically apply scoring methods will not only help them internalize a scoring approach, but will also involve them in the assessment process, thereby helping eliminate the mystery that often leads to feelings of powerlessness and uncertainty (White, 1994). Giving them access to a scoring approach and explaining its application enables them to form a basis for writing and revising to meet grader expectations. Meanwhile, it reinforces the proper way of communicating with members of the academic discourse community. As an example, if students were exposed to a method for grading in 2 passes—first for framing,
organization, and development; and second for sentencing, grammar, and word choice—they would see that only papers judged good in both passes would earn high scores.

A distressing reality, and one of the most difficult aspects of teaching academic English writing to non-native speakers, is that students must acquire the ability to write in grammatically correct English, even though correct grammar clearly does not predominate in the ESL world—or in any other, for that matter. Herein lies the prime dilemma for SL students writing under time constraints. Which response do graders of ESL timed writings for college entry tend to consider more up-to-standard—a less developed one that is grammatically correct, or a more developed one with grammar errors? It would help teachers to know the answer.

More empirical research in this area—as the present experimental study provides—can help inform us. Time and again in our classes we see conscientious ESL learners who work diligently, taking whatever time necessary, to deliver the caliber of writing we expect. We set high standards, which many students meet, but when called upon to write an impromptu response under a time limit, they are again forced to make the difficult choice. Do they sacrifice grammaticality to write a longer answer—even though in doing so, they will surely incorporate some of the ungrammatical quirks of fluent language? Their other choice is to forego fuller development, in order to make what they write more grammatically correct, which carries its own risk.

One successful teaching strategy is to make more efficient use of the outside environment, to encourage noticing of the differences between grammatically correct English and English in the real world. Second language learners need to develop a keen awareness of these register differences in English. English as a second language learners
need informing that unless they develop this awareness, they face the risk of composing in the spoken style (Halpern, 1984). And doing so may result in their essays being judged by academic readers as “illogical, fragmented, or inappropriately discursive” (Schleppegrell, 1996, p. 273). This will reduce their chances for success on timed tests and in college courses. Through modeling, a teacher can help students notice register differences. Success at this will depend in part on how consistently one attends to noticing the words and structures in the context. The ability to recognize differences will improve one’s skill at writing acceptable academic English.

Also, while students are prewriting and drafting their essays, teachers can use facilitating and coaching strategies to consistently commend good composing behaviors, good style, and good writing, and they should encourage ESL students to find meaning in the writing tasks (Emig, 1977; Shaughnessy, 1970; White, 1995). Coaching is good; students like it and respond well to it. What’s more, working together with peers in class is a good way to improve reading or writing skill development. For example, those with particular skills and strategies take leadership roles, helping the less skilled or competent in those areas (Spaulding, 1992).

Teacher Response to Student Essay Drafts

Recent research has also done much to inform us about good and bad teacher response to student writing; the written responses shown to be bad for L1 writers are equally unproductive for SL writers, and vice versa (Leki, 1992). For instance, papers on which a grade appears leave most students indifferent to the written commentary also present. Moreover, students have difficulties locating stated weaknesses, have trouble making requested revisions, and feel discouraged by indecipherable teacher handwriting.
Results of one informative cross-sectional study put grammar errors into a broader context of exploring whether writing quality might develop over time (Haswell, 1988). Haswell compared 4 levels of 32 writing samples—freshman, sophomore, and junior university levels; and writings done by post-college employees. He measured 8 kinds of error, which “experienced readers consider damaging and that raters could measure reliably” (p. 485). The error types were the following: misformed possessives, faulty predication (subject-verb agreement), pronoun reference, syntactic parallelism, punctuation of final free modification, fragments, comma splices, run-ons, and spelling. Also considered was the context; construction types that might lead a writer to make an error—e.g., appositives and words infrequently in print—were factored in, as a means of accounting for writing maturity.

What emerged was that freshman writing was the worst, and writing quality improved along with the college class level (Haswell, 1988). Further, this research revealed that in the development of writing skill, not only did students repeat old mistakes, but they also made new ones. Surprisingly, more mature writers increasingly misspelled common words and dropped suffixes. Haswell speculated about a variety of causes, including an improvement in content, a higher rate of production, and less concern or time, or both, for proofreading.

Haswell’s study effectively helps put errors into a more realistic, and less idealistic, perspective. The fact that L1 student writing at higher class levels is improved, but still contains errors—even basic ones, such as misspellings of easy words and omissions of word endings—is a meaningful finding. As Haswell himself pointed out, while surface errors did not always diminish, mistakes were often accompanied by “measurable growth
in that same writing toward mature competence” (p. 495). This is a warning to composition instructors that “squelch[ing] writing mistakes . . . [could equate to] . . . squelching the growth in writing that precipitated the mistakes” (p. 495). It is certainly feasible that SL learner writings would similarly contain new errors and past errors while they advance in college class levels. While learners’ composing skills advance, they experiment with more advanced, complex language forms, which seems to be partly to blame for the pervasiveness of errors. This study underscores the reality that teacher time spent marking errors in papers is drudgery for her and unhelpful for her students.

Of course, writing teachers must not ignore the errors in student compositions. Current wisdom, however, about how to respond to the errors may seem, at first, counterintuitive. Both L1 and SL composition pedagogy now advise postponing editing until an essay exhibits purpose, substance, and organization. Of course, the most serious grammatical, syntactic, and rhetorical errors should be discussed and dealt with (Shaughnessy, 1970; Thompson, 1990; White, 1995). Thus, the formerly acceptable approach of concentrating primarily on surface-level issues when responding to drafts is of no noticeable benefit to learners. We now know that attending to the smaller problems thwarts attending to the more serious meaning-related problems. Students who receive responses of the former type come to view writing as an exercise in language use, not as a process of articulating meaningful intentions to an audience (Zamel, 1985).

While it has been shown that a scorer who encounters numerous errors throughout a response will probably fail it (Bardovi-Harlig and Bofman, 1989), most ESL timed writings contain serious and less serious errors. The serious confuse or block understanding, while the less serious seldom hamper the comprehension of meaning.
Tomiyana’s study (1980) is illustrative; 2 types of errors—articles and connectors—were considered on 3 levels: 1) omission, 2) insertion, and 3) wrong choice. In general, a clause connector error—a coordinate or subordinate conjunction, or a relative word—was harder for a native English speaker to repair. The levels of difficulty for native speaker repairs were, from the most to the least difficult: 1) omission, 2) wrong choice, and 3) insertion. Article errors were the easier type for native speakers to repair, unless they were article insertion errors.

Tomiyana equates a lack of comprehensibility with a degree of “communication breakdown” (1980). Hence, the inability of native English speakers to repair ungrammaticalities and inaccuracies in the written text determines the degree of seriousness of the communication breakdown. Little breakdown occurs where a native English speaker is capable of supplying the misinformation with repaired information that matches the intention of the writer. Judgments of comprehensibility would be useful as a guide to determining which errors to focus instructional attention on most.

Time permitting, or for the more advanced students, the local errors in student writings should also be dealt with. Lennon (1991) highlighted an important drawback in error analysis, which is that scorer variability results from different scorers identifying errors differently; varying levels of acceptability were assigned to the same instance of code. For some evaluators perhaps what figures most into a determination of failure is not the type of error, but the total number of errors in the sample. Another strategy then might be to deal with frequent errors of a particular type in a given learner’s writing. Helping the learner eliminate the most errors possible with the least effort, or learner confusion, could be a teaching goal.
Nonetheless, composition instructors must not lose sight of the writing task—topic—and how well students express their response to it. Sommers (1980) and Silva (1990) agree: Without exception, teachers need to have learners revise meaningfully, expanding all unclear and insufficient expression in light of the whole text. Reiterations abound: It is crucial to clearly and explicitly relay composing strengths and weaknesses to learners, emphasizing how well students answer the writing task and express intended meanings (Anson, 1999; Hamp-Lyons, 1996; Thompson, 1990).

Furthermore, just as learners must be induced to repair all anomalies in style and form that lead to miscommunications or ambiguities (Zalewski, 1993), teachers need to consider the ability of learners to understand the responses they put on learner drafts. Through teacher-student conferencing, negotiations can lead students to express meanings more comprehensibly, meanwhile helping them better understand the causes of the confusions. Conferencing students individually about their essays is, in fact, one of the best ways to address strengths and weaknesses. Students need direct instruction and practice in repairing portions that stray from the overall purpose—so that they can either improve the relevance or delete them altogether, with absolutely no concern for any embedded errors. Meeting students one-on-one also facilitates their improving textual connections and organization, and better understanding where more development is needed. It can also help them develop the ability to independently recognize strengths and repair weaknesses in their essays. What’s more, it opens the door for greater personalizing of their learning.

Second language acquisition research widely recognizes that to increase the benefits of formal instruction, SL teachers should not put the learner on the defensive
(Krashen, 1982). Besides, “understanding how human beings feel and respond and believe and value is an exceedingly important aspect” in the classroom and out (Brown, 1987). Praises on composing strengths often boost confidence levels and increase the motivation for revising. Students are thus likely to benefit when teachers emphasize the development of skills, good writing practices, and effective student essays. It is less effective for teachers to dwell on weaknesses and errors in student writing, and give primarily negative feedback.
CHAPTER 3
MATERIALS AND METHODS

Six non-ESL college writing teachers used 2 well-known scoring methods to evaluate 60 timed English as a second language writings. Holistic and analytic scoring are widely recognized for determining distinct aspects of quality in writings. Each scoring was preceded by a training session. It was critical that the holistic training and the holistic scoring be fully completed before proceeding with the analytic training and scoring. This precaution was taken to avoid confounding one scoring approach with the other. Since the 2 methods are distinct, using them intermittently risks confusing graders in their decision making, and producing less trustworthy scores. After both evaluations, the chief led the 6 scorers, all together, in a question-and-answer discussion session. This was to inform and enrich the scoring data. Commentary includes thoughts and reactions from the graders and from the chief about the ESL papers evaluated, ESL writing in general, and the effectiveness of my holistic and analytic methods for scoring these essays.

Holistic scoring is essentially effective for achieving grader agreement on overall quality. After the chief’s training (using 14 representative samples), each teacher holistically evaluated 20 papers—thus each of the 60 received 2 independent evaluations. Analytic training and scoring followed (using fresh copies of the same papers). Analytic

---

1 Details about holistic and analytic scoring and a short comparison of the two approaches are given in Chapter 2. The designs of my versions of these 2 methods are given and explained later in this chapter.

2 Training is described in Chapter 2 and further discussed in this chapter.
scoring, while more time-consuming than holistic, is considered effective for determining distinct strong and weak areas. Four global skills were each independently scored for quality—focus, development, organization, and clarity.

**Selection of Data Samples**

The full dataset for this study consisted of 74 timed college-entry test writings composed by ESL applicants (between 1999 and 2000) at 3 campuses of the junior college where this experimental study was held. Fourteen were used in the trainings, 60 in the scorings. The papers were written on 2 basic prompts: 1) My ambition: what I want to do with my life; and 2) What job would you like to have in the future?

While some holistic scorings use a 6-point range, as discussed in Chapter 2, a 4-point range is not uncommon, and was selected for my study. Holistic graders are, customarily, first trained with a subset of like responses across the score range. For proper calibration, the training needed to include papers across all proficiency levels—1, 2, 3, and 4. This would allow distinctions to emerge, enabling graders to more accurately determine rank levels. Thus, 74 stratified\(^3\) samples were drawn—a good number at each rank level—and later divided into 14 training papers and 60 scoring. At 2 campuses, the college’s ESL placement teachers had recorded rank scores to mark the proficiency of each sample—beginning ESL, intermediate ESL, advanced ESL, or college English composition. Different levels of essays were easily drawn from these 2 campuses, by regarding the assigned writing scores.

At the third campus, however, no writing scores had been recorded. In order to draw stratified samples from there, the ESL placement teacher’s course placements were

---

\(^3\) Stratified sampling entails *selectively* drawing samples from the data source, to represent the range of groups studied.
relied upon. This was not problematic, for 2 reasons: these placements were easy to find, and the writing proficiency level was said to have heavily driven placement.

In the period from 1999 to 2000, during which essays were drawn, designated ESL teachers at this same junior college were using a 6-point ESL Holistic Scoring Guide (Appendix D) to evaluate the timed writings of ESL entrants. Generally, 1 trained ESL teacher read each timed response and evaluated its writing proficiency level, giving it a point score on the scale. A middle-range point score indicated that the response was solidly representative of the description given for that level. An up-arrow designation indicated the essay had some strong features for that level, but that it did not qualify for the next higher level. A down-arrow designation marked the essay as being notably weak in some aspects for the level, yet not warranting a drop in proficiency level ranking.

The ESL placement teachers at this college all had access to pre-scored samples to help calibrate and check their evaluating, and in questionable cases, asked for assistance from another ESL placement teacher or from their program chair. Theoretically, a timed written response was considered college-ready—or good enough to exempt a learner from ESL—only if assigned the point score of 6 on this scale. A 6 paper was distinct in that it satisfied 5 expectations: 1) stated a main idea; 2) was coherent and fully developed; 3) used correctly spelled academic words; 4) formed compound and complex sentences appropriately; and 5) contained infrequent errors.

All 74 essays drawn shared 2 main features: 1) each answered 1 of the 2 topic prompts; and 2) each was composed under test conditions on the college premises. Test

---

4 From archived files, by a committee of ESL teachers responsible for the evaluating of ESL writing tests for course placement. Copyright 1999 by St. Petersburg Junior College. Reprinted with permission.
subjects were male and female speakers of various languages—Russian, Vietnamese, Korean, French, Japanese, Spanish, Arabic, Polish, Greek, and Portuguese, among others—new to this college, and with varying degrees of formal instruction in English. The timed ESL essays have been coded to preserve learner anonymity.

Next, the researcher and 2 colleagues, all college ESL placement scorers and ESL teachers at different campus sites, met to select and rank score 14 papers from a subset of those collected from the 3 campuses, for use in the trainings. They all read and independently scored, using my Holistic scoring guide, given in Table 3-1. I created this guide by adapting, with permission, parts from 2 holistic scoring guides I myself have used to grade timed writing tests. One was the ESL Holistic Scoring Guide (Appendix D), and the other was the Scoring criteria for the college basic skills exit test (Table 2-1).\(^5\) Explanation for my holistic design is given later in this chapter. The 14 training papers ultimately selected were considered accurate representations of the 4 proficiency levels: 3 papers at level 4 (college-ready); 4 at level 3 (advanced ESL); 4 at level 2 (intermediate ESL); and 3 at level 1 (beginning ESL).

**The 20 Papers in Each Grader Packet**

English as a second language rankings were relied upon when drawing stratified samples, and then again for preparing grader packets. Accordingly, it must be indicated that the holistic guidelines used by ESL graders (Appendix D) between 1999 and 2000 were more linguistically based than those used for the holistic scoring in this study. Holistic results for the papers in this study may not, therefore, correspond well with the pre-scores for these same papers. In fact, the ESL group tended to give papers with

\(^5\) The ESL holistic scoring guide is discussed earlier in this chapter, and the guide for the college basic skills exit test is described in Chapter 2.
occasional language inaccuracies pre-scores corresponding to a value lower than a 4 on my holistic scale. Corresponding values are discussed shortly, in this same chapter.

Table 3-1. Holistic scoring guide—descriptions of ratings for the ESL writings

<table>
<thead>
<tr>
<th>Preliminaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Please bear in mind how difficult a task an impromptu writing can be for an ESL learner. Moreover, consider how constrictive a time limit 35 minutes is for an ESL learner to compose, revise, and edit a response.</td>
</tr>
<tr>
<td>• Since examinees responded with paragraphs and essays, I will call the writing a “paper”, and I will call the paper’s main point its “purpose statement.”</td>
</tr>
<tr>
<td>• Consider how this writer might fare in Composition I, and rank the overall quality of the paper in light of the 35-minute time limit of this writing test.</td>
</tr>
<tr>
<td>• Please be sure to write your ranking in the upper right-hand corner of the paper. Thank you.</td>
</tr>
</tbody>
</table>

4  College-ready ESL—ready for Comp I The paper’s purpose statement is focused and near the beginning of the paper. The development is very satisfactory for the 35-minute time limit: Major and minor supports are detailed, organized, and connected. The writer demonstrates noticeable facility with word choice, sentence structure, grammar, and mechanics. This paper is quite readable and satisfactorily communicates meaning; it engages the reader.

3  Advanced ESL The paper’s purpose statement is generally focused and near the beginning of the paper. The development is satisfactory for the 35-minute time limit, in some parts: Some major and minor supports are detailed, organized, and connected. The writer demonstrates some facility with word choice, sentence structure, grammar, and mechanics. This paper is readable and communicates meaning, but of a more limited, less interesting sort.

2  Intermediate ESL If a purpose statement exists, it may be a copy or a restatement of the writing prompt. Though organization and development may be evident, the writing is unsatisfactory for the 35-minute time limit. Some support may consist of generalities and / or unconnected ideas or experiences. Readable parts coexist with occasionally inaccurate word choice, sentence structure, grammar, and mechanics, making the communication of meaning unsatisfactory.

1  Beginning ESL If a purpose statement exists, it is most likely a copy of the writing prompt. The lack of development and / or organization is most unsatisfactory for the 35-minute time limit: What support exists often consists of generalities and / or unconnected ideas or experiences, strings of words, or lists. Lack of substance or frequently inaccurate word choice, sentence structure, grammar, and mechanics, result in a basic, incomplete communication of meaning.

The 20-paper packets for the 6 graders were prepared by allocating, to each, nearly equal shares of responses at my study’s 4 ESL writing proficiency levels—college-ready, advanced, intermediate, beginning. Correspondences between the 6-point ESL rubric
(Table D-1) and my 4-point scale (Table 3-1) were made by collapsing the 6-point rankings into a college-ready ESL level plus 3 college-preparatory ESL levels. The rankings were collapsed as follows: 6 became my 4 (college-ready); 5 and 4 became my 3 (advanced); 3 became my 2 (intermediate); 2 and 1 became my 1 (beginning). To make comparing grader scores more principled, their sets included like distributions of papers, according to correspondences to this study’s 4 rank levels:

- R1 / R2 / R4 / R5: four papers at level 1, five at level 2, nine at level 3, two at level 4
- R3 / R6: three papers at level 1, five at level 2, eleven at level 3, one at level 4.

Also worth mentioning, over 50% of each grader packet consisted of papers with upper-half holistic pre-scores (my study’s advanced and college-ready levels).

Participants

Seven current or former English composition teachers participated in this study, and all were given codes. One served as a chief reader; the other 6 were scorers.

The Chief Reader

The chief reader is semi-retired, taught college composition for many years, and served as an English department chair in 2 community colleges. He has extensive experience in scoring writing tests such as a statewide mandated college-level essay test, and is a seasoned trainer for group scorings. He continues to work as a scorer and trainer for direct writing tests. The chief conducted and managed this 4-hour scoring experiment—the holistic training and scoring; the analytic training and scoring; the post-scoring debriefing—and was given the code CR.

The Six Scorers

Six scorers of varying ethnicities were all, at the time of the experiment, employed at 1 campus of the college from which the papers were drawn. They were
given the following codes: grader R1—a white female; grader R2—an African-American female; grader R3—a white female; grader R4—a white female; grader R5—a white male; and grader R6—an American male. Mention should be made here that participation in this experiment was voluntary, for 4 hours on a Saturday. This may have been partly responsible for an observed level of verbal participation from 2 graders. Relative to the other 4, graders R2 and R6 gave less verbal commentary during the trainings and the post-scoring question-and-answer session.

**Procedures**

The experiment took place in February of 2002, in a comfortable classroom at 1 campus site of this junior college. The chief scorer presided, establishing and maintaining an amicable, professional environment. He had previously been given copies of all papers. Each scorer was given papers and scoring guides when needed.

The experiment was completed in 4 hours, as planned. It was composed of 3 parts that took place in the following order: first, a holistic training and scoring; second, an analytic training and scoring; third, a group question-and-answer discussion session.

The only incident was the 10-minute late arrival of grader R4, due to traffic and inclement weather. Her tardiness appeared influential in her reported dissatisfaction with the wording of the descriptors for the highest writing proficiency level in my Holistic scoring guide. Nonetheless, this seemed to have no negative effect on her scoring.

---

6 Given to the chief were the Holistic scoring guide (Table 3-1); the Holistic Refereeing Method (Appendix C); the Analytic scoring guide (Table 3-4); the 14 training papers, pre-scored holistically; the 60 scoring papers, unscored; and new copies of the 14 training papers, margin-marked and pre-scored analytically.

7 Upon grader R4’s late arrival, the chief instructed her briefly in how to score using the Holistic scoring guide (Table 3-1). She then proceeded to rank score the first six training papers along with the rest of the scorers, then continued to score and contribute to discussions in ways consistent with other raters in the group.
At the outset of the experiment, as a baseline for how the teachers were to grade the 60 papers in both scoring instances, the chief instructed the raters to establish in their minds the standards that they ordinarily applied when they evaluated freshman entry-level writings. Essentially, they were to rely on their training for and experience in the college classroom, trusting their knowledge and first impressions as they scored.

My Holistic scoring guide described the level-4 essay, beside the heading “4 College-ready ESL—ready for Comp I” as follows:

The paper’s purpose statement is focused and near the beginning of the paper. The development is very satisfactory for the 35-minute time limit: Major and minor supports are detailed, organized, and connected. The writer demonstrates noticeable facility with word choice, sentence structure, grammar, and mechanics. This paper is quite readable and satisfactorily communicates meaning; it engages the reader.

The researcher and chief reader further described a level-4 paper, early in the holistic training:

R: A paper that you give the score of 4 should indicate that the student writer would be able to survive in Composition I.

CR: We’re not going to put somebody in freshman composition with a 3 / 4 score. It’s going to be either a 4 / 4, in freshman comp, or a 3 / 3, into whatever is below that.

Both my Holistic scoring guide (Table 3-1) and my Analytic scoring guide (Table 3-4) contained wordings intended to keep the scorers from allowing novel ways of scoring to supplant customary ones. My holistic guide instructed them to “consider how this writer might fare in Composition I,” and my analytic guide specified that the scorers “stay concentrated on the ideas.” Furthermore, to encourage typical college teacher evaluations and commentary throughout the experiment, while in attendance to audio record, distribute and collect papers, I refrained from joining in discussions unless called
upon to do so by the chief. Most particularly, I avoided giving evaluative opinions about the ESL samples, so as not to influence scorings.

**Rater Training: Design Considerations and Anticipated Effects**

Rater training is considered helpful in reinforcing and facilitating desired consensus scoring behaviors. It helps clarify—for the graders—the goals of the scoring method; thus, it is a means of enabling the graders to meet those goals.

Immediately preceding each scoring in this study, the chief directed a training session, instructing and guiding the scorers as they ranked a set of pre-scored samples, in accordance with the particular scoring approach. Methods to achieve this end included: exposure to some of the papers, practice in applying the grading methods, and discussions and justifications of scores given.

Besides giving practice in applying the standards uniformly, the trainings presumably helped the non-ESL teachers overcome 2 particular challenges. First, in the holistic scoring and in the analytic scoring, graders needed to adjust their expectations, somewhat, to accommodate for the relatively short 35-minute time limit these examinees had been given. These English faculty graders were accustomed to grading longer responses, written in a total of 45 minutes, so the training should help them make this particular adjustment in their expectations.

The other challenge, in the second—analytic—scoring, was that the graders needed to be able to avoid focusing on the reparable errors in word endings, articles, and prepositions while they read and evaluated. The Analytic scoring guide (Table 3-4—detailed later in this chapter) instructed them to pay less attention to reparable, mechanical, sentence level errors, in order to stay focused on the communication of meaning. This directive would presumably better enable them to score each global
feature in isolation, a primary goal of this study. Training would thus help them avoid an undue focus on reparable errors.

Here is common wording taken from the instructions on both my Holistic scoring guide (Table 3-1) and my Analytic scoring guide (Table 3-4), given as a “Preliminary,” to instruct graders to keep in mind the difficulty of the SL learner’s task:

Please bear in mind how difficult a task an impromptu writing can be for an ESL learner. Moreover, consider how constrictive a time limit 35 minutes is for an ESL learner to compose, revise, and edit a response.

Since my holistic guide included wording at every point score along the range that would have guided graders to consider language issues—“word choice, sentence structure, grammar, and mechanics”—it is expected that graders would have maintained some concern for language issues while scoring holistically.

On the other hand, my analytic guide explicitly directed graders to concern themselves about language only when scoring the global area of clarity. When scoring focus, development, and organization, they were instructed to attend less to errors and omissions at the sentence level—errors affecting 1 idea only—so as to attend to communicative meaning:

In order to keep a global focus, you will need to pay less attention to reparable errors / omissions at the sentence level; do not let minor misspellings or inaccuracies in articles, prepositions, and commas command your attention. Stay concentrated on the ideas.

Method One—The Holistic Training and Scoring

The chief trained and directed the holistic scoring in accordance with the theoretical guidelines and practical procedures discussed in Chapter 2, and as given in my Holistic scoring guide in Table 3-1. My reasons for creating this as an adaptation of 2 holistic
scoring guides used at this college, and limiting it to a 4-point scale, hinged upon familiarity and ease of implementation.

In the first place, not only were these 2 holistic guides tested (making them trustworthy), but also, 1 of the guides had been used by all 6 graders, so the tenets drawn upon would be familiar to them. Furthermore, regarding the 4-point scale selected, it seemed unreasonable to expect these graders to be able to distinguish 6 levels of ESL papers (as a 6-point scale would have implied), when neither their experience nor their training equipped them to make such fine distinctions in ESL writing proficiency levels. The scale was divided, then, into 4 levels, with the maximum score being 4, and the minimum being 1.

To apply the holistic scale, each rater first individually rank ordered the training papers, with no 2 raters paired together at any time. Acceptable accord in designating anchors—model papers for the graders to refer to when grading—was defined as a majority agreement on the score of 4, 3, 2, or 1 for a paper. Thus, when 4 or more of the 6 graders agreed that a paper exemplified a given writing proficiency level, accord was determined satisfactory, and the graders were directed to refer to that paper as a representation of that score level.

These anchors then served as concrete representations, referents for grader rankings better than the more abstract descriptions given in the scoring guide across the range of levels. In the chief’s words:

When you’re evaluating the other papers, don’t go back to the descriptors if you can help it. That is, use the range finder [or anchor] as the sample 2 or the sample 3. If you go back and keep reading those descriptors, you’ll start tripping over things.
In preparation for the scoring, 2 copies each of 60 papers—a total of 120 papers—had been divided into 6 packets. Each scorer was given a packet of 20 papers to evaluate. At the end of the scoring, the 60 papers had each been scored independently twice, once by 2 different teachers—with no collaborating.

The graders’ and the chief’s previous training and experience, including their scorings of college English writing tests, enhanced their ability to determine successfully and expediently which papers were college-ready. A 1-½ hour allotment was made for the holistic training and the scoring of 20-paper packets. This part went smoothly and finished sooner than anticipated. The holistic scoring divided the 60 essays into 45 papers at 4 proficiency levels—4, 3, 2, and 1—and 15 papers with discrepant score pairs.

Resolving Discrepant Holistic Score Pairs

In holistic scorings, 2 raters occasionally assign the same paper 2 different scores, and the pair may not be acceptable for the purposes of the scoring. Because of this inevitability, it had been essential to define unacceptable score pairs, and to establish a method for rectifying them, in advance.

Four unacceptable holistic score pairs—also called splits—were defined for this study’s holistic scoring. The 3 constituting conflicts about an examinee’s readiness for college-entry are these: 4 with 3; 4 with 2; and 4 with 1. That is, in any such pair, 1 rater found an examinee to be college-ready and assigned a holistic score of 4, while a second rater assigned a score of 3, or 2, or 1, marking the need for additional ESL preparatory coursework. The remaining split—3 with 1—was included because it fulfilled the customary definition of a split in holistic scorings—being a pair 2 units apart. However, it was not of particular interest to the purposes of this study. The 4 defined types of holistic splits are presented in Table 3-2.
Table 3-2. Discrepant holistic score pairs—splits requiring resolution

<table>
<thead>
<tr>
<th>Split Type A</th>
<th>Split Type B</th>
<th>Split Type C</th>
<th>Split Type D</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 / 3*</td>
<td>4 / 2*</td>
<td>4 / 1*</td>
<td>3 / 1</td>
</tr>
</tbody>
</table>

* Any score pair constituted by a 4 and any other score is a split of particular interest because it marks a grader disagreement on college readiness.

It must be emphasized here that while congruent scores are not ordinarily seen as splits, the score pair of 4 with 3 was defined as problematic by this research study, whose design required that an undivided rater decision be made any time any rater assigned a college-ready score. Indeed, a college-ready score of 4 together with an ESL score of 3 is an ambiguous combination for my purposes. Theoretically, college-ready ESL writing and college preparatory ESL writing are distinct, and global differences between the 2 are of interest. The former is thought more comprehensible and organized, and developed enough to suggest that the individual’s composing could meet the expectations of an English composition faculty member. The latter is considered weak in 1 or more areas, suggesting a need for 1 or more semesters of college preparatory ESL coursework.

Yet unfortunately, as my own experiences in group scorings have revealed, and as has been shown in one research study of the holistic scorings of midlevel timed writings (Ruetten, 1994), this theoretical expectation is not always realized in practice. When a range of noticeable problems co-exist with rhetorical control or coherent arguments, discrepant holistic scores can result. Undeniably, it is no easy matter to judge borderline cases, holistically. Even proponents of holistic testing express words of caution regarding the effectiveness of this method for judging the problematic ESL writings: “problem papers—such as those containing second-language difficulties—may be included as samples but not usually as range finders” (Wolcott, 1998, p. 77). Wolcott continues that problems with syntax or errors at the sentence-level that seem particularly
troublesome are likely to lead readers to assign an essay a lower score, especially if they obscure communicative meaning.

The Refereeing Method\(^8\) established for rectifying the 15 discordant score pairs embodies the approach used to resolve grader disagreements in some holistic scorings. This method and its application were discussed with the chief in advance, due to the special circumstances of the unconventional \(4 / 3\) split introduced here.

Every paper that received a split score pair, the chief read and evaluated. He gave a score that replaced the most discrepant score in the pair, and retained the other grader’s. All papers receiving splits, of \(4\) with any lower score, were resolved so that they ended up with either 2 ratings of \(4\) or 2 scores below \(4\). The determination—of whether an essay qualified as a holistic \(4\) or not—was primary. That is, first, the best caliber writings needed to be identified.\(^9\) Thereafter, the distinguishing qualities in the best papers could be identified. This was a central goal of this research project.

In contrast to the 4 discrepant holistic score pairs, 6 holistic pairs were defined as acceptable. These included matches—\(4\) with \(4\); \(3\) with \(3\); \(2\) with \(2\); \(1\) with \(1\)—and pairs in which the 2 adjacent scores represent ESL writing at relatively analogous proficiency levels—\(3\) with \(2\); \(2\) with \(1\). All acceptable score pairs are presented in Table 3-3.

<table>
<thead>
<tr>
<th>Pair A</th>
<th>Pair B</th>
<th>Pair C</th>
<th>Pair D</th>
<th>Pair E</th>
<th>Pair F</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 / 4</td>
<td>3 / 3</td>
<td>2 / 2</td>
<td>1 / 1</td>
<td>3 / 2*</td>
<td>2 / 1*</td>
</tr>
</tbody>
</table>

*Contiguous scores can stand together without refereeing, as a rule.

\(^8\) The Refereeing Method used to resolve splits is given in Appendix C.

\(^9\) This junior college regulates the resolution of pass / fail pairs on its barrier test in this same way—any paper receiving one college-ready ranking and any ranking below college-ready—even the adjacent, lower, point score—must be scored by another faculty to replace the division. This reasoning was applicable to my study’s purposes.
Method Two—The Analytic Training and Scoring

Attention must be drawn, at this point, to the theoretical distinctions between holistic and analytic scoring. A holistic score is a single point grade for a whole essay in its entirety. The scorer reads holistically and grades the multifaceted, multi-aspectual nature of the essay. Analytic scores, constituted by a set of independent rank scores (four in this study), each represent a grade for a distinct rhetorical or grammatical aspect within the essay. The scorer reads more deliberately and grades each aspect in turn.

Additionally, when scoring holistically, graders are specifically instructed not to reread the essay. Thus, while a holistic reading of a timed ESL essay in this study could have resulted in a somewhat unclear understanding of meaning, a more deliberate analytic reading of that same essay, later, could have resulted in clearer grader comprehension of intended meanings. Chances of this are heightened by the fact that raters often review or reread parts of an essay when analyzing the various features to analytically score. A scorer might have better understood an intended meaning in an essay in this study, for instance, after analytically reading that essay than he or she had understood after reading that same essay once through earlier, in the holistic scoring.

The readers were instructed to follow and complete my Analytic scoring guide, given in Table 3-4. For the same reasons that led me to my design of the holistic scoring method, I developed my analytic scoring method by adapting, with permission, 4 recurring global themes in 2 holistic scoring guides with which I had experience—the ESL holistic scoring guide (Appendix D) and the Scoring criteria for the college basic skills exit test (Table 2-1). Four fundamental tenets recurring in the descriptions of these guides—focus, development, organization, and clarity of expression—seemed to be primary in determining acceptability in academic writing. Thus, because these 4 themes
Table 3-4. Analytic scoring guide—indicators of global writing competence in an ESL timed written response

A Preliminary

Please bear in mind how difficult a task an impromptu writing can be for an ESL learner. Moreover, consider how constriction a time limit 35 minutes is for an ESL learner to compose, revise, and edit a response. Evidence from research has well established that nonnative-speaker timed writing is less fluent than native-speaker. Thus, parts can come across as 1) off topic; 2) out of position; 3) disconnected; 4) oddly worded; or 5) incomprehensible. In light of all this, as you read each paper, please make an extra effort to grasp the intended meanings, by:

- reading this writing more slowly than you would read native-speaker writing
- granting a writer re-readings of a few stretches of text
- mentally relaxing between papers.

General Directions

Please evaluate this examinee’s global writing competence; indicate performance levels for four global skills. In order to keep a global focus, you will need to pay less attention to reparable errors / omissions at the sentence level; do not let minor misspellings or inaccuracies in articles, prepositions, and commas command your attention. Stay concentrated on the ideas. If, however, despite your extra efforts, a portion of text interferes with meaning more than it communicates meaning, please draw a vertical line in the right-hand margin directly beside the troublesome text. As soon as you finish reading the paper, complete Sections 1 and 2 below.

Please notify me when you have finished scoring all the papers assigned to you.

Section 1

My Reader Code (Roman Numeral) ___
The ESL Writer’s Code (4 Letters) ______

Please verify the Reader Code and the Writer’s Code. Thank you.

Section 2

Notwithstanding the performance in syntax, grammar, and mechanics, I consider this student’s performance in the four Global Skills below as (circle one Level for each Global Skill)

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>high</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 all of the time</td>
<td>3 most of the time</td>
</tr>
</tbody>
</table>

(Optional Comment)

Please assure that one Level of Performance is circled for each Skill. Thank you.

If applicable, finish the following statement
I was impressed by ______________________________
were familiar, I expected that the graders in my study would be able to apply them with ease. Ease of application would, in turn, help alleviate one drawback of analytic scoring—that it is time-intensive (discussed in Chapter 2). I chose a 4-point scale of quality for the 4 features, for the same reason—it would take less grader time than a more complex scale.

To help the graders apply the analytic scoring method, the chief followed the same order of steps as he had in the holistic scoring. The training came first. The 6 raters were instructed to make every effort to disregard reparable mechanical errors and omissions at the sentence level, so as to stay concentrated on the ideas and meanings. It was believed that their attending less to language use would enable their staying focused on the 4 global skills of interest—focus, organization, development, and clarity of expression. This focus would, in turn, presumably enable them to score each skill in isolation, as much as possible.

Whenever a given global skill was demonstrated consistently in a paper, raters presumably gave a score of 4 for that skill. A 4 corresponds to the qualifier “all of the time.” The other scores and their corresponding qualifiers are these: an analytic 3 corresponds to “most of the time,” a 2 to “some of the time,” and a 1 to “not at all.”

In consultation prior to the experiment, the researcher and the chief agreed that analytic scores of 4 and 3, being upper-half values, both represented levels of performance meeting grader expectations. However, a 4 corresponded with skill at the highest expected level, and a 3, with skill at a satisfactory level. And analytic scores of 2 and 1, being lower-half values, both represented levels of performance failing to meet grader expectations. But again, there was a difference in terms of degree. A score of 1
indicated a complete lack of skill, while a 2 marked an existing, but unsatisfactory level of skill.

During the analytic scoring, graders were instructed to mark a paper with a vertical line in the side margins in every place where the wording seemed confusing. Here are the explicit instructions, in the “General Directions” on the Analytic scoring guide:

Please evaluate this examinee’s global writing competence; indicate performance levels for four global skills. In order to keep a global focus, you will need to pay less attention to reparable errors / omissions at the sentence level; do not let minor misspellings or inaccuracies in articles, prepositions, and commas command your attention. Stay concentrated on the ideas. If, however, despite your extra efforts, a portion of text interferes with meaning more than it communicates meaning, please draw a vertical line in the right-hand margin directly beside the troublesome text.

The chief clarified this scoring guide’s instruction for margin marks during the analytic training. He gave an explanation, first, in answer to an inquiry made by grader R1. She asked, “Notwithstanding performance in syntax, grammar, and mechanics . . . throw that out of my brain, and . . . ?,” to which the chief replied:

Well, no, with one exception. . . . Clarity takes care of that. Now what I found, I scored all of these and what I found is I frequently had a higher score as a result of analytical scoring than I did with holistic scoring, because I, for example, if the paper had a clear central idea, I tended to score it a 4, and it really wasn't a [level] 4 paper, but . . . there it was.

Thereafter, this follow-up question was posed by grader R6 about what sort of clarity issues warranted their marking papers in the margins: “Now to evaluate that, is it ESL awkwardness, or Comp I awkwardness, because there's a difference?” The chief clarified:

Well, I would think Comp I awkwardness. Well you'll always know what they mean. It's simply a matter of sometimes it's so difficult to figure it out. If it's really terribly unidiomatic, I think you probably want to mark it, in the margin. . . . I've been a comp teacher long enough to know that what I really get hung up on is the faulty English. The misspelled words . . . they're going to get that in freshman composition. The question is whether or not the language is sufficient, so that
they’re not going to be destroyed by a thousand misspelled words and awkward, you know ‘awk,’ ‘awk,’ ‘awk,’ on the side.

After the analytic training was the analytic scoring. Each grader received a folder of clean copies of the same 20 papers that he or she had previously holistically scored, along with 20 copies of the scoring guide for recording analytic scores for each essay. Scoring papers they had seen before presumably minimized the scoring time and avoided irritating tired graders. It must be pointed out, however, that graders had no access to their packets of holistically scored papers—or to the holistic scores they had already assigned—since the researcher had already collected these.

Discrepant Analytic Scores—Analysis, Not Resolution

Scoring differences are common in analytic evaluations, since there are more points to be disputed. Resolving analytic differences was not planned, due to time constraints. Rather than resolving discrepant analytic scores during the experiment, discordant cases were investigated after the experiment by the researcher. The chief alerted the graders to the expected discord, at the start of the analytic training:

Analytic scoring is . . . different from the holistic scoring. In holistic scoring, the assumption is that everybody is going to agree. . . . Eventually, you come to consensus about what a 2 is . . . imprint that on your mind, and that’s what you’re looking for. . . . In analytical scoring . . . there’s likely to be a little bit of difference.

He clarified “a little bit of difference” by explaining that differences between an analytic 3 and a 4, as between an analytic 1 and a 2 involved degree. That is, though both upper-half scores indicated that expectations were met, and both lower-half scores indicated that they were not met, analytic scores 4 and 1 marked the greatest degrees of satisfaction and dissatisfaction, respectively.

In accord with the theoretical distinctions and explanations given in Chapter 2 about analytic training and scoring, and with further such discussion in this chapter,
papers were expected to have received higher analytic scores, at times, than holistic. Thus, the score combinations in Table 3-5 represent practical guidelines delimiting the exceptional cases—discordant analytic scores, in light of holistic scores. In practice, a paper that got 2 holistic 4s might get the occasional—one or 2—lower-half analytic scores, but it seems discordant for any paper rewarded with holistic 4s to get 3 or more lower-half analytic scores. Likewise, while it is feasible that a paper that got 2 lower-half holistic scores might get some upper-half analytic scores, it seems discordant for any paper given such low holistic scores to get 7 or more upper-half analytic scores.

Table 3-5. Discordant analytic scores—considered in light of holistic scores

<table>
<thead>
<tr>
<th>Holistic Scores (2 scores total, from 2 graders)</th>
<th>Analytic Scores (8 scores total, from 2 graders)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 holistic 4s</td>
<td>3 or more analytic 2s and 1s</td>
</tr>
<tr>
<td>2 lower-half holistic scores (2s or 1s)</td>
<td>7 or more analytic 4s and 3s</td>
</tr>
</tbody>
</table>

Setbacks in the Analytic Scoring

One setback in the analytic scoring occurred on the morning of the scoring. The chief informed me, then, that he would not show the 14 prepared transparencies of analytically-scored and margin-marked samples because he disagreed with some of the scores and marks. While in theory, in a previous meeting, the chief and I had agreed about how to assign analytic scores and how to mark the margins, in practice, we disagreed on some papers. In the training, then, the chief briefly discussed the range of scores and the 4 global skills to be analytically scored, and gave verbal examples of instances of unclear prose alongside which raters were to put marks in the margins. Grader reports of fluctuating confidence levels in their analytic scores owed partly, then, to the limited analytic training delivered, and partly to their comparatively high confidence levels with holistic scoring.
Another setback was that 2 graders—R2 and R6—put few marks on few papers, despite giving numerous essays low scores for the global skill of clarity. Equally problematic, 2 other graders—R1 and R3—wrote specific and abundant comments on many papers, even marking minor errors directly on papers that they considered to have communicated well, as shown by the high scores they gave these on clarity. And the remaining 2 graders—R4 and R5—lightly marked about half the papers, some that they had not scored low for clarity and some that they had. The abbreviated analytic training, the absence of visual models, and the limited practice before scoring, all help explain why grader markings on the papers inconsistently illuminated the particular instances of prose the graders considered to be unclear.

It was especially surprising, also, that R2 repeatedly asked for clarification about what to mark and how to put marks in the margins. The most experienced grader in the group, R2 was near retirement, had the most years teaching college composition, and was the only one who had regularly participated for years in large-scale scorings of a state-mandated timed writing test for entry into junior level university studies. Five times she asked for clarification on making margin marks. Her last request for information was a direct statement of uncertainty: “I don’t know what I’m expected to put in the margins.”

To this, the chief responded with an illustrative example from a training paper:

Ok, let’s assume you’re reading down here, and it says, ‘After I want to return to my country and I want to work there.’ Now, let’s assume you think that’s awkwardly worded or hard to read. Then you put a little mark in the margin. That’s all.

To the chief, this clarification must have seemed sufficiently explicit, for he let discussion move on to other topics. In the analytic scoring, R2 gave many low scores for
clarity to papers in her set, but she made no marks in any margin—or elsewhere—on any paper. Her reluctance clearly rested in her uncertainty about what and how to mark.

Such setbacks probably contributed to the unreliability of some analytic scores. I had planned to provide categorizations of unclear phrasings, by determined levels of seriousness of the breaches in clarity. However, the variability in scorer markings discouraged my analyzing these markings, as I had planned to do.

Post-Scoring Debriefing

Once all the scoring was complete, it was time for the debriefing discussion session. The primary goal of this final portion of the experiment was to gather from the group of 6 graders, all together, 30 minutes of verbal commentary about these essays, about my scoring approaches, and about ESL writing and writers in general. It was believed that 30 minutes of verbal response from the group would provide a sufficient amount of information to round out the scoring data. A secondary purpose of this debriefing was to provide closure to the experiment.

At the outset, the chief encouraged all graders to contribute answers to the group discussion. He then asked all 10 questions in sequence. Some graders participated more frequently and more fully than others, but all gave some responses. Here are the 10 questions:

1. Which global skill(s) seemed strongest in the upper level ESL papers that you saw here today?
2. Which global skill(s) seemed weakest most often? How would you help your own students here?
3. When marking student drafts, which one(s) of these 4 global skills do you tend to emphasize?
4. What could you take away from this session to give better feedback to your ESL writers?
5. Has this session changed the way you feel about ESL writing or ESL student writers? How?

6. What advice would you like to give to ESL teachers preparing students for English composition?

7. What, if anything, irritated you while you were reading these ESL papers? Why is that?

8. Share one of your experiences of an ESL student’s success…a small success or a large one.


10. Was there anything in the Analytic scoring guide with which you disagreed? Explain.
CHAPTER 4
RESULTS

The central question this research posed was this: To what extent is college-ready caliber, in an ESL timed essay, determined by the specific global qualities of focus, development, organization, and clarity?

Holistic scoring first divided the 60 papers into 3 groups: level-4s (college-ready); level-3s / 2s / 1s (ESL-levels); and indeterminate splits (level-4 with 3 or 2 or 1 OR level-3 with 1). After split resolution, the 3 groups collapsed into 2 holistic levels: 11 papers earned holistic 4s; 49 resulted in holistic 3s / 2s / 1s.¹ Thereafter, analytic scoring provided information about grader satisfaction levels for 4 global skill areas—focus, development, organization, and clarity—as given by the contents of all 60 essays.

Scoring results, essay contents, and views from the 6 graders and the chief about the papers and about ESL writing in general are informative and have been organized as follows. In accord with the primary interests of this study, attention is first turned to an in-depth examination of the best 11 papers. Next is a detailed analysis of trends of overall and global qualities in all 60 papers, as suggested by the resulting holistic and analytic scores and grader commentary. Thereafter is a discussion of factors influencing judgments of quality. This is followed by a review of factors and trends observed in the papers ending in unacceptable holistic splits. A brief discussion of essays at ESL levels

¹ Appendix B tabulates the scores and other information for the 49 lower-scoring essays, some of which are discussed in depth later in this chapter.
is presented next, and this is followed by some afterthoughts about accord. A summary of all the results concludes the chapter.

The Level-4 Essays

Scores and other data for the 11 level-4 papers are given in Table 4-1. Eight essays received holistic 4s from 2 graders. Three, the chief resolved from splits to 4. All these papers are described, and some are presented, in the next section. The rest are given in

Table 4-1. Eleven holistic 4s (essays scored college-ready)

<table>
<thead>
<tr>
<th>Essay code</th>
<th>Language</th>
<th>Gender</th>
<th>Word count</th>
<th>1st Gdr scor</th>
<th>2nd Gdr scor</th>
<th>Chief's refereed scores for splits</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUKS</td>
<td>Greek</td>
<td>M</td>
<td>267</td>
<td>R2</td>
<td>R5</td>
<td>4 4 3 4 4 4 3 4</td>
</tr>
<tr>
<td>ZLDA</td>
<td>Polish</td>
<td>F</td>
<td>361</td>
<td>R2</td>
<td>R4</td>
<td>4 4 4 4 4 4 4 4</td>
</tr>
<tr>
<td>KAKE</td>
<td>Japan</td>
<td>F</td>
<td>178</td>
<td>R1</td>
<td>R4</td>
<td>4 4 4 4 4 4 3 4</td>
</tr>
<tr>
<td>KNGK</td>
<td>Russ</td>
<td>M</td>
<td>487</td>
<td>R2</td>
<td>R5</td>
<td>4 3 4 4 4 3 4 4</td>
</tr>
<tr>
<td>ROYY</td>
<td>Span</td>
<td>F</td>
<td>159</td>
<td>R1</td>
<td>R6</td>
<td>4 4 3 3 4 4 3 4</td>
</tr>
<tr>
<td>SALO</td>
<td>Portu</td>
<td>M</td>
<td>198</td>
<td>R1</td>
<td>R4</td>
<td>4 4 4 4 4 3 3 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Best-Scoring Essays</td>
</tr>
<tr>
<td>HOSE</td>
<td>Japan</td>
<td>F</td>
<td>207</td>
<td>R1</td>
<td>R6</td>
<td>3 4 / 4 4 4 3 3 3 3 4</td>
</tr>
<tr>
<td>PTGE</td>
<td>Frenc</td>
<td>F</td>
<td>219</td>
<td>R3</td>
<td>R6</td>
<td>4 4 / 4 3 3 3 3 3 3 2 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Second Best-Scoring Essays</td>
</tr>
<tr>
<td>KAOA</td>
<td>Russ</td>
<td>F</td>
<td>212</td>
<td>R2</td>
<td>R5</td>
<td>4 4 2 3 3 4 2 3 2</td>
</tr>
<tr>
<td>KKMA</td>
<td>Russ</td>
<td>F</td>
<td>176</td>
<td>R1</td>
<td>R6</td>
<td>3 2 3 2 2 2 2 2 2 2 2 1</td>
</tr>
<tr>
<td>HIAR</td>
<td>Hindi</td>
<td>M</td>
<td>212</td>
<td>R3</td>
<td>R6</td>
<td>3 4 / 4 3 2 2 2 2 2 2 1 1</td>
</tr>
</tbody>
</table>

Holistic: 4=college-ready; 3=advanced ESL; 2=intermediate ESL; 1=beginning ESL
Analytic: 4=all of the time; 3=most of the time; 2=some of the time; 1=not at all
Appendix A. The 11 essays are divided into 3 groups, according to their analytic scores: 6 Best-Scoring, 2 Second Best-Scoring, and 3 Third Best-Scoring essays. Each essay in the first group earned 5 or more (of a total of 8) analytic 4s, and both earned analytic 4s for development or organization, or for both. Each essay in the second group earned from 2 to 4 analytic 4s and no more than 1 lower-half analytic score—a 2. Each essay in the third group got at least 3 analytic 2s.

Two important findings have emerged, regarding the level-4 essays. First, graders generally gave a holistic 4 whenever they perceived that rhetorical or other strengths dominated in a paper. And upon the more deliberate analytic reading, when 2 graders recognized a global strength in an area, both gave that area an analytic 4—marking notable strength.

Two Noteworthy 4s

Two holistic 4 essays merit special attention for their preponderance of analytic 4s.

Figure 4-1 gives the beginning and end of essay ZLDA (at 361 words), which contains

As each and every one of us wishes to fulfill one’s dreams and expectations, my desire to study and obtain proper certification isn’t very much different from any other. I would actually like to become a property lawyer and be able to serve local community with my knowledge and professional service. . . . I have not formed myself becoming a criminal lawyer due to the fact that legal procedures in such matters would take a pricy toll on my private life. I am in general very sensitive person and would not be able to forget drastic or harmful situations which criminal justice creates. This doesn’t though mean that I wouldn’t know the difference between what would be right or wrong in particular situation. Maybe because I have been taught some simple life lessons and as anyone else, I am also a product of healthy relations in my childhood household. I have been taught something what my parents and the rest of family applied – being truthful in all one does and giving it the best of one’s tries. Obviously, I do not know what the future will bring for me or for our nation, but all aforementioned should be sustained as everyone’s creed.

Figure 4-1. The beginning and end of essay ZLDA (given holistic 4s by R2 and R4)
myriad features of strong writing: substantive length, a clear focus, strong arguments, full supports, fluid and easily readable language. Graders R2 and R4 gave 8 (of 8) analytic 4s to ZLDA on the global features—focus, development, organization, clarity.

Essay KAKE (at 178 words) is significantly shorter, but even so, strong rhetorically: focused, coherent, organized, unified, detailed, concluded. Graders R1 and R4 both gave it analytic 4s on focus, development, and organization. Figure 4-2 presents this essay, in its entirety.

I would like to work as a veterinarian in the United States in the future. I graduated from Veterinary School seventeen years ago and have worked as a veterinarian in Japan since then. While I studied at AMC (The Animal Medical Center) in New York, I met an American and married to him.

I and he came back to Japan and worked together at an animal hospital in Nagano, my home town.

Working the hospital in Japan was wonderful for us, but it wasn’t satisfied us completely. Because I think that veterinary medicine in Japan is still behind. In addition to it, animals are not protected by the law. I have seen many animals that were neglected badly, and everytime I treated them, I felt terrible and sad.

Our goal is that I and my husband work together again for the animals in America. It may be difficult to achieve the goal. It may take years to do it, but I try as much as I could. Studying English is the first step that I have to do.

Figure 4-2. Essay KAKE (given holistic 4s by R1 and R4)

Focus—In the 11 Holistic 4s

Figure 4-3 lists the first 2 sentences of all 11 holistic 4 papers, and their scores for focus. Essay HOSE, plus all 6 best-scoring essays save KNGK, got 4s from both graders on focus. Most of the 11 holistic 4 papers earned high scores for focus, which clearly shows that these graders considered focus important to good writing. The graders rewarded the skill of focus when the first sentence acknowledged the prompt, and the second clarified, or gave additional background information. On the other hand, they appeared less convinced by the focus when the first 2 sentences were simple or factual—
as essay KKMA’s—or noncommittal—as essays KAOA’s and HIAR’s. It is no surprise that effectively focused papers were rewarded analytically. Both second language and mainstream writing texts patently direct students to form definitive topic sentences and theses. Graders doubtless appreciate focus in timed writings.

**Essay CUKS**—4 Setting future goals is easy to do. The hardest is to actually go through all “adventures” that may come my way and achieve my goals.

**Essay ZLDA**—4 As each and every one of us wishes to fulfill one’s dreams and expectations, my desire to study and obtain proper certification isn’t very much different from any other. I would actually like to become a property lawyer and be able to serve local community with my knowledge and professional service.

**Essay KAKE**—4 I would like to work as a veterinarian in the United States in the future. I graduated from Veterinary School seventeen years ago and have worked as a veterinarian in Japan since then.

**Essay KNGK**—4 As they say only God knows what is going to be with one’s life. Every person may have only ambitions, dreams and goals.

**Essay ROYY**—4 One of my biggest dreams since I was in sixth grade has been to become a translator. I love languages.

**Essay SALO**—4 COMPUTER GRAPHIC ENGINEER. THAT’S GOING TO BE MY JOB TITLE IN A FEW YEARS.

**Essay HOSE**—4 My ambition is to get a job, which I can use my English for, in Japan. I used to work for a pharmaceutical company as an office worker and gave it up in Spring 1998.

**Essay PTGE**—3 I would like to work in a comunication company in the future, because I think it is a personal enrichment to have business relations with people since they learn you how the world works. As far as I am concerned, I am skilled in custumer relationships thanks to the jobs I have already got.

**Essay KAOA**—2 What do I want to do with my life? When I’m asking myself this question, I’m thinking about my passed and my future: what I did and what I’m planning to do.

**Essay KKMA**—2 I came from Russia last years. My husband is American.

**Essay HIAR**—2 Every person in this world wants to be dominant and sucessful and so I too, in lust of greed, walk in the same path. I would like to have my life in the pages of history.

Figure 4-3. Focus-analytic scores and the first two sentences (of the 11 holistic 4 papers)
Qualities in the Best-Scoring Essays—Development, Organization, Clarity

The 6 best-scoring essays—CUKS, ZLDA, KAKE, KNGK, ROYY, SALO—received holistic 4s from 2 graders. Each thus consistently met grader expectations, as a whole. Further, these 6 papers all got fewer analytic scores of 3 and more analytic scores of 4. Therefore, each also consistently met grader expectations well for the 4 global skills—focus, development, organization, and clarity.

The analytic scores of all 6 best-scoring essays revealed that development and organization were both of import to the English graders in my study. One or both of these global components received 2 analytic scores of 4. It is not at all surprising that English graders would reward papers having full development or effective organization.

Essay CUKS (at 267 words) was focused, gave clear main ideas and supports, and was clearly organized—all making it easy to read, holistically, for understanding. It got analytic 4s from graders R2 and R5 on focus and organization.

A critical finding, and one that should offer hope to ESL learners and their teachers, is that even lacking 2 scores of 4 on clarity, 3 of these 6 essays (CUKS, ROYY, SALO) were rewarded with many other analytic 4s. Of course, graders had been directed to concentrate primarily on the communication of meaning. On the analytic scoring sheet graders completed for each paper, this explicit directive appeared just above the box for their scores: “Notwithstanding the performance in syntax, grammar, and mechanics, I consider this student’s performance in the four Global Skills below as. . . .”

For instance, while analytically scoring essay CUKS, perhaps precisely because of this directive to focus attention primarily on the communication of meaning, neither grader R2 nor R5 made analytic marks to denote confusing parts, and R5 even noted being impressed by its organization. Not even the following verbal error was marked:
“It may take years to do it, but I try as much as I could.” And the few article errors were clearly considered incidental: “go through all ‘adventures’,” or “First and most important goal.” In fact, not even unidiomatic wording elicited a margin mark on the paper in the analytic scoring, from either grader. Neither marked any of the following portions, which seemed a bit awkward, especially the first sentence:

The world of Business is not something which describes me nor something which I aspire high to achieve. . . . I and he. . . . wasn’t satisfied us completely. . . . in addition to it. . . . I and my husband. . . .

Further, this comma splice elicited no margin mark: “Photography has great importance in my life, that is why. . . .” The 3 and 4 scores for clarity showed that grader R2 was hardly distracted by these potentially problematic elements, and grader R5 not at all.

Essay KNGK (at 487 words), which was prolifically detailed, was rewarded with analytic 4s on development and clarity from graders R2 and R5. Neither marked anything on this paper. Yet some might contend that it contained a couple parts not altogether clear in their meaning, namely these:

So what I am saying is when I complete the studies in the US, am planing to go back to Russian Federation as everyone knows the World’s Future is there. . . . Also would like to do my hobby – that is teaching youngers how to live. I think have everything for doing my best to reach or get all that.

This suggests that if an ESL learner’s response to a timed essay prompt includes plenty of information on the topic, some graders might overlook a few occasions of unclear expression, or even an occasional unsupported argument.

Essay ROYY (at 144 words) stood apart from other level-4 essays, for its comparatively short length and less in-depth detailing. Nonetheless, it demonstrated unity, coherence, and exemplification. It also gave a conclusion and showed fluency in English. Graders R1 and R6 gave it analytic 4s on focus and organization.
Essay SALO (at 198 words) was rewarded by graders R1 and R4 with analytic 4s for focus and development. Like essay ROYY in many aspects, it was unified, coherent, and fluent; it had arguments and supports communicating clear meanings; and it had a conclusion. But it had 6 indents (ROYY had none) and was considerably longer.

**Qualities in Lower-Scoring 4s**

Two essays that got 3 / 4 holistic split scores first—HOSE and PTGE—were both refereed to 4, and analytic scoring elicited more 3s and fewer 4s for both, as compared to what the 6 best-scoring essays got. This suggests that essays HOSE and PTGE were not as strong as those categorized as best-scoring in Table 4-1. Graders were thus divided on the quality of the whole of each essay, but then in independent global skill areas, they acknowledged strengths. Essay HOSE (at 207 words) only got 4 analytic 4s (of a total of 8, from 2 graders), while the best-scoring essays each got 5 or more. These student writers apparently were not as convincing to graders of their skills in focus, development, organization, and clarity. Figure 4-4 gives essay HOSE in full.

My ambition is to get a job, which I can use my English for, in Japan. I used to work for a pharmaceutical company as an office worker and gave it up in Spring 1998. As I didn’t have any remarkable certificates and licenses then, I couldn’t get the jobs that I really wanted to do. I decided to go abroad to study English and obtain ability of speaking and certificates of English.

I first went to England and studied there for a year and during that time, I succeeded in passing the Cambridge Examination (FCE). Then I realized that I would need some American English certificates as well because it is more recognized in my country. My stay in the U.S.A will be one year and I’m going to take TOEIC exam and obtain good score. If I can get high scores in TOEIC, I am sure I will be able to find my ideal job in Japan, I know it is not easy and I have to work very hard for it though.

I hope I can work for a trading company and after working there for several years, settle down and work at home using a computer looking after my children. That’s my ambition.

Figure 4-4. Essay HOSE, a split resolved to a holistic 4 (given a 4 by R1, a 3 by R6)
Holistically, grader R1 rewarded essay HOSE with a 4, so this essay looked like the level-4 anchors to her. Grader R6 disagreed, considering it to be more like a holistic 3. Yet in the subsequent analytic scoring, neither grader made any margin marks on this paper, and both gave it 2 analytic 3s and 2 analytic 4s, which suggests that both minimized problems in the writing, and recognized the writing’s global qualities. They seemingly discounted a somewhat awkward choice of coordinating conjunction (“and” vs. “but”), as well as a few article omissions—for instance, “obtain [the] ability” and “obtain [a] good score.” Barring the absence of paragraph indentations, essay HOSE has evident strengths—a clear focus, chronological organization, formal wording and sentencing, coherence and unity. It cannot be determined why, precisely, this timed response was not considered a holistic 4 by grader R6. In sum, essay HOSE has many strengths and only sporadic weaknesses.

Three other level-4 essays—KKMA, KAOA, HIAR—stand apart, in that graders gave them discordant score combinations. Two graders scored each of essays KKMA (at 176 words) and KAOA (at 212 words) 4s holistically. Thereafter, grader R5 and grader R6 gave KAOA and KKMA, respectively, 3 or more lower-half analytic scores (1s or 2s). In fact, grader R6 gave KKMA low analytic scores on all 4 global features, in contrast with the score of 4 he had given it holistically, but why? And why would grader R5 have given KAOA a 4 holistically, but then 3 scores of 2 analytically?

It is not apparent why R5 and R6 gave these essays each a holistic 4. The qualities in these papers can be analyzed, however, in light of possible influencing factors. Essay KKMA, graders R1 and R6 each scored a 4 holistically, perhaps because they empathized
with this learner’s attitudes, personality, experiences, or circumstances. This portion might have elicited grader empathy, for instance:

… I have two beautiful daughters, I love my family very much. I’d like that my children could go to college. When I lived in Russia, I worked in hospital. I saw many sick children, unhappy parents. I’d like to practice here as a doctor. My dream is to improve my English skill than take medical test.

Figure 4-5 presents essays KKMA and KAOA.

**Essay KKMA**  My ambition: what I want to do with my life.
I came from Russia last years.
My husband is American. It is one reason why I am here. I have two beautiful daughters, I love my family very much. I’d like that my children could go to college. When I lived in Russia, I worked in hospital. I saw many sick children, unhappy parents. I’d like to practice here as a doctor. My dream is to improve my English skill than take medical test.

It nice to see a children in good health, with smile in their face. I’d like to help, to treat them. I have a lot of experience.
Although I’d like to get big house when my children’ll grow up and live me, they’ll come back home and have good rest.

I have another dream. I’d like to travel, see other countries, people.

Life is not long. Sometimes it seems not have time for everything. But it is not true. What I think about myself, I have to hurry, study hard, work. But life is wonderful. Unfortunately not didn’t take notice how!

**Essay KAOA**  My ambition: what I want to do with my life.

What do I want to do with my life? When I’m asking myself this question, I’m thinking about my passed and my future: what I did and what I’m planning to do. I changed my life a few years ago, so I’m thinking about my future again and again.

First, I think the things which I did is very important for me I have two children and I’m happy to be a mother and a wife.

I graduated from Medical Shcool and I worked in a hospital, so I know the medicine better, than many people and it helps me in my life.

But; I’m a communicative person, so I need to work, because it’s important to me. Next, I want my children to be happy and have all what they need and what they want. Because of that, I want to earn money. And, I need more respect from myself and the other people. It’s very pleasant for me to cure people, help them if I can, so it will be better if I can work in a medical field. Finally, I decide to choose the courses in a College and become a nurse.

Maybe, if I’ll do this, I’ll be happy, because now I have almost all what I need.

Figure 4-5. Essays KKMA and KAOA (given holistic 4s by R1, R6, R2, R5)

Analytically, grader R1 reported essay KKMA weak in organization and clarity.

And she highlighted 3 instances of inappropriate language use: “practical” (for “practice”), “It” (for “It’s”), and “in” (for “on”). Grader R6 reported it weak in all 4 global areas, a judgment which is strongly supported by the choppy sentencing, disjointed
ideas, and noticeable lack of unity. Indeed, this writer jumps from topic to topic—the Russian homeland, motherhood, hospital work, a dream house, traveling. She goes on to philosophize about life and to emphasize the importance of a formal education. Gaps in support and incomprehensibilities occur in a mix of subtopics—advantages and disadvantages, dreams and hardships—features often prevailing in lower ESL writings.

Essay KAOA was scored 4 holistically by graders R2 and R5. This learner’s writing reveals her self-assurance, educational prowess, high aspirations, and overall positive outlook. She states how her accomplishments fulfill her, and goes on to strongly assert her goals:

First, I think the things which I did is very important for me I have two children and I’m happy to be a mother and a wife. I graduated from Medical Shcool and I worked in a hospital, so I know the medicine better, than many people and it helps me in my life. . . .

Next, I want my children to be happy and have all what they need and what they want. Because of that, I want to earn money. And, I need more respect from myself and the other people. It’s very pleasant for me to cure people, help them if I can, so it will be better if I can work in a medical field. Finally, I decide to choose the courses in a College and become a nurse.

She offers evidence of her experience in completing a high degree and in meeting familial responsibilities. Her writing concludes with this positive statement: “Maybe, if I’ll do this, I’ll be happy, because now I have almost all what I need.” In sentence 2, several verb errors emerged, and yet these do not necessarily interfere with communicative meaning:

When I’m asking myself this question, I’m thinking about my passed and my future. . . . First, I think the things which I did is very important for me.

And 2 consecutive sentences display basic errors in punctuation and pronoun usage:

But; I’m a communicative person, so I need to work, because it’s important to my. Next, I want my children to be happy and have all what they need and what they want.
Other errors exist throughout and include a variety of article and capitalization problems. Nonetheless, upon a holistic reading, graders R2 and R5 were sufficiently convinced of this writer’s readiness for freshman composition, despite the language inaccuracies. The contents of the essay indicate considerable experience in life and in formal education, plus an overall positive outlook, all of which these graders probably gathered by reading holistically, and all of which may have led to the holistic 4s.

Analytically, however, grader R5 gave lower-half scores for focus, organization, and clarity—reflecting his dissatisfaction and calling attention to the discord in these analytic scores, in light of his earlier holistic score for this same essay. He made note of a pronoun misuse when he underlined “my.” His not marking other language inaccuracies, however, may show he deemed the language mostly communicative enough to understand. On the analytic score sheet he noted, “Wanders too much to be clear.” In contrast, grader R2 gave all high analytic scores, suggesting that her enthusiasm for this response was consistent, carrying across from the holistic to the analytic scoring.

The other best-scoring essay that earned discordant score combinations was HIAR (at 212 words). This paper got a holistic 3 / 4 split that was refereed to 4, and also raises questions. Why would grader R6 have scored HIAR 4 holistically when he and another rater (R3) considered it to have no clarity—scores of 1 from both—and to be developed and organized only some of the time—scores of 2 from both? Figure 4-6 gives essay HIAR in full.

This paper may have garnered holistic 4s from a grader and the chief because they shared expressed attitudes. This learner’s personality, positive outlook, and creative flair
strongly and immediately emerge. The 2 first sentences and the last are clear examples of an express dynamism and powerful yearning for success. The essay begins as follows:

Every person in this world wants to be dominant and successful and so I too, in lust of greed, walk in the same path. I would like to have my life in the pages of history.

Essay HIAR was closed with this proverbial expression: “Life shall stay forever this way, try for this I always may.” Many a mainstream composition teacher who would enjoy having such a personality in class would have scored this a 4 to help make this possible.

Perhaps such a fondness for this personality swayed grader R6 to give the holistic score of 4 to HIAR. Yet in the analytic scoring, R6 gave 2 indicators to show that this response failed to meet his expectations upon a more deliberate reading. He marked 2

My ambition: What I want to do with my life

Every person in this world wants to be dominant and successful and so I too, in lust of greed, walk in the same path. I would like to have my life in the pages of history. As a practical student I like to see my life as a successful Computer Science Engineer where I just have to give my mental stress. In a multinational company I wish to work, in a cabin of my own, with a salary fulfilling my needs and wants. Thus leading a satisfied life.

But, as a human I like to be much more than that. A person violinist who can play the tunes when ever he may. A poet who can write poetry what he wants and how he wants. A person who is even a part of the ‘big screen’. I would want to become a man being called ‘The One Man Show’, having tasted all of the life.

I would try to uphold my career as an Engineering, being my field of interest, but given an opportunity chance of enjoying the life my way I would not avail it. [“life shall stay forever this way, try for this I always may”]

Figure 4-6. Essay HIAR, a split resolved to a holistic 4 (given a 3 by R3, a 4 by R6) unclear portions, and he gave lower-half scores on all 4 global areas. As unclear, he marked:

. . . my life as a successful Computer Science Engineer where I just have to give my mental stress. . . . I would try to uphold my career as an Engineering, being my field of interest, but given an chance of enjoying the life my way.
Aside from these 2 unidiomatic sentences, it is striking that 4 of the remaining 9 sentences in this essay are fragments. After the opening 4 sentences is the first fragment: “Thus having leading a satisfied life.” Three more fragments soon follow:

A person violinist who can play the tunes when ever he may. A poet who can write poetry what he wants and how he wants. A person who is even a part of the ‘big screen’.

That nearly half of the content of this essay displays such weaknesses seems to show essay HIAR to be genuinely ESL-level and helps account for grader R6’s lower analytic scores. Grader R3 scored HIAR low for development, organization, and clarity, no doubt for the same problems.

**Overall and Global Trends**

Beyond the findings about qualities in the 11 best essays, overall and global trends in the 60 essays emerged and merit attention. What second language writers were able and unable to do well is suggestive for second language writing pedagogy. Mean scores (Table 4-2) provide some preliminary approximations of grader perspectives about

<table>
<thead>
<tr>
<th>Grader</th>
<th>Holistic</th>
<th>Focus</th>
<th>Development</th>
<th>Organization</th>
<th>Clarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Mean</td>
<td>2.60</td>
<td>3.20</td>
<td>3.00</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.10</td>
<td>.89</td>
<td>.73</td>
<td>.93</td>
</tr>
<tr>
<td>R2</td>
<td>Mean</td>
<td>2.80</td>
<td>3.10</td>
<td>2.50</td>
<td>3.10</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.89</td>
<td>1.02</td>
<td>1.00</td>
<td>1.17</td>
</tr>
<tr>
<td>R3</td>
<td>Mean</td>
<td>2.20</td>
<td>2.45</td>
<td>2.30</td>
<td>2.35</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.62</td>
<td>.76</td>
<td>.73</td>
<td>.67</td>
</tr>
<tr>
<td>R4</td>
<td>Mean</td>
<td>2.40</td>
<td>2.90</td>
<td>3.10</td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.00</td>
<td>.85</td>
<td>.72</td>
<td>.88</td>
</tr>
<tr>
<td>R5</td>
<td>Mean</td>
<td>3.10</td>
<td>3.35</td>
<td>3.25</td>
<td>2.85</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.79</td>
<td>1.04</td>
<td>.91</td>
<td>.81</td>
</tr>
<tr>
<td>R6</td>
<td>Mean</td>
<td>3.30</td>
<td>2.55</td>
<td>2.15</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.73</td>
<td>1.05</td>
<td>.75</td>
<td>.79</td>
</tr>
<tr>
<td>Total</td>
<td>Mean</td>
<td>2.73</td>
<td>2.93</td>
<td>2.72</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>.93</td>
<td>.98</td>
<td>.90</td>
<td>.92</td>
</tr>
</tbody>
</table>

R2 gave five holistic 4s—the mean of all graders’ holistic 4s
R1 and R4 gave upper-half holistic scores closest in number to ESL pre-scores
R3 gave no holistic 4s
R5 and R6 gave the most holistic 4s—seven and nine, respectively
overall quality (mean holistic scores) and the global skill areas (mean analytic scores) in all 60 papers in general, without making any special distinctions by holistic score level. Grader commentary will enrich these approximations. Table 4-2 depicts, grader by grader, the mean holistic and analytic scores for the 20 papers evaluated by each grader.\(^2\)

Recall that the 6 graders did not all evaluate the same 20 papers. Selection of the 20 essays for each grader packet was based on the distribution of relatively equal shares of papers at each ESL writing proficiency level, as detailed in Chapter 3. Considering this caveat, the total mean scores reflect relative approximations of the scoring tendencies for the group.\(^3\)

Trends in the Whole Paper

Recall that, after split resolution, 49 of the 60 papers (about 82\%) had determinate holistic scores below a 4, which indicated that to these graders and the chief, the overall quality of the large majority of the 60 essays fell generally below that of entry-level freshman writing. In support of this is the following commentary, offered in the post-scoring debriefing by the 2 graders who, coincidentally perhaps, gave the strictest scores analytically. Grader R3 reported that there seemed to be a need for encouraging ESL students to go through the writing process: brainstorming, outlining, organizing, and revising. This suggests that she considered the set of 20 papers she scored to be mostly comprised of drafts that reflected unplanned, unorganized, unedited writing. Likewise, grader R6 argued that these learners needed to develop the ability to break the act of

\(^2\) Each paper was scored twice, thus the 120 holistic scores and 120 sets of analytic scores.

\(^3\) Scoring particulars for the 60 papers are given elsewhere—scores for the 11 holistic 4s are in Table 4-1, and scores for the other 49 are in Appendix B.
writing into 3 time units—and to make sure they revised for one third of the time. He, too, perceived that few writers, of the 20 whose responses he had read, had done revising.

Individually, some variations in holistic scores resulted, to no one’s surprise, and are shown in Figure 4-7. Each cluster of bars represents the set of 20 papers a grader scored holistically; thus, the height of each bar gives the number of papers that the grader scored at that point value. Differences in the proportions of holistic 4s, 3s, 2s, and 1s the graders gave to their 20 essays are evident. Yet it must again be pointed out that not

![Figure 4-7](image)

**GRADER**
- R2 gave five holistic 4s— the mean of all graders’ holistic 4s
- R1 and R4 gave upper-half holistic scores closest in number to ESL pre-scores
- R3 gave no holistic 4s
- R5 and R6 gave the most holistic 4s—seven and nine, respectively

Figure 4-7. The six graders’ holistic-rank scores for their 20 essays (of the 60) every grader scored the exact same 20 papers, so these variations can only be taken as approximations of holistic scoring differences. Still, in every grader’s holistic scoring
folder (of 20 papers), either 11 or 12 essays had pre-scores from ESL placement graders that categorized them at advanced or college-ready ESL writing proficiency levels, and either 8 or 9 essays had pre-scores categorizing them at intermediate or beginning ESL writing proficiency levels.\(^4\)

Closer examination of agreements and disagreements on the 8 level-4 matches, and on the 14 level-4 splits, is revealing.\(^5\) Table 4-3 presents counts of grader accords and discords. Column 2 tabulates grader accords for the 8 matches between 2 graders.\(^6\) Column 3 gives total adjusted accords—column 2 totals combined with adjusted numbers of accords between the graders and the chief on his resolutions of the 14 level-4 splits.

Table 4-3. Accord on papers graders gave one or two holistic 4s (includes splits)

<table>
<thead>
<tr>
<th>Grader</th>
<th>Accords between Graders for the eight matching holistic 4 scores</th>
<th>Total # Adjusted Accords</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>+4 2 (R4) and 2 (R6)</td>
<td>+6</td>
</tr>
<tr>
<td>R2</td>
<td>+4 1 (R4) and 3 (R5)</td>
<td>+5</td>
</tr>
<tr>
<td>R3</td>
<td>0</td>
<td>+3</td>
</tr>
<tr>
<td>R4</td>
<td>+3 (see R1, R2)</td>
<td>+4</td>
</tr>
<tr>
<td>R5</td>
<td>+3 (see R2)</td>
<td>-1</td>
</tr>
<tr>
<td>R6</td>
<td>+2 (see R1)</td>
<td>-1</td>
</tr>
</tbody>
</table>

The total numbers of adjusted accords essentially support the pictorial representation of how graders’ holistic scores distributed in Figure 4-7. All in all, graders

---

\(^4\) See the exact distributions given under the section “The 20 Papers in Each Grader Packet” in Chapter 3.

\(^5\) See Appendix E for the holistic scores that graders and the chief gave for these splits.

\(^6\) Scores and other particulars for the 8 matches can be seen in Table 4-1 at the beginning of this chapter.
R1, R2, R3, and R4 saw slightly more eye-to-eye with each other and with the chief reader, as to what constituted a level-4 paper, than did graders R4 and R5.

A few variations illustrated by the information presented in Table 4-3, and that given in Figure 4-7, are worth pointing out. Probably the most noteworthy is grader R2’s giving 5 holistic 4s—precisely the mean score of all graders’ numbers of holistic 4s. This is noteworthy because it seems to bear out grader R2’s skill at meeting consensus views in holistic scorings, which in turn follows from her having the most years’ experience—of these 6 graders—at large-scale holistic scorings. Of course, this point and the next few, drawn from these data, are not necessarily generalizable, given the small number of graders in this study, but still, they enlighten the scoring results.

Also of note is that graders R1 and R4 gave numbers of upper-half holistic scores (advanced and college-ready levels) closest to the combined numbers of those same levels given by ESL graders. In their 20-paper folders, each of graders R1 and R4 had a total of 11 essays pre-scored at these 2 levels. Then in Figure 4-7, we see that graders R1 and R4 each assigned a combined total of 9 upper-half holistic scores (advanced and college-ready levels). It is probably no coincidence that these 2 graders have the most years’ experience (among these 6 scorers) teaching freshman composition, outside of grader R2. This means they have also had the most years’ exposure to ESL learner writing in freshman composition classes. This fact reflects in their having been able to distinguish advanced and college-ready level ESL writing (from intermediate and beginning levels) in distributions the most similar to those given by the trained ESL graders who assigned the pre-scores.

\[7\] Adding the bars for each grader’s number of level-4 scores (as shown in Figure 4-7) results in 30, which divided by six, equals five.
A final point is that graders R5 and R6 gave holistic 4s a bit more discordantly than the others, with slightly fewer adjusted accords, at -1 each (as shown in Table 4-3). And as the bars for the holistic score of 4 in Figure 4-7 indicate, graders R5 and R6 gave 7 and 9 holistic 4s, respectively—slightly higher proportions than the other 4 graders gave. In fact, a male scorer gave a holistic 4 to an essay in 11 of 13 instances in which a female scorer assigned the same paper a lower holistic score—to essays PTGE, HIAR, GACS, LUHA, OAPA, ROPT, SAAA, SDHA, SEEA, PYMA, and TNGA. Only 2 of these were resolved to holistic 4s, while the rest were rectified to scores lower than a 4, which shows that in the chief’s view, 9 of these given holistic 4s were unduly high.

Later in this chapter, it will be suggested that these male graders seemed to have rewarded positive thinking, creative style, or self-confidence with higher holistic scores, at times, than did 3 female graders—R1, R2, and R3. At the same time, however, the occasional lower-half analytic scores that these male graders gave to these same papers revealed that they scored less generously, analytically, when they evaluated these same writers’ global essay-writing skills.

It was somewhat puzzling that positive commentary from 3 of the 4 females—graders R1, R2, R4—and the lack of commentary from the 2 males—graders R5 and R6—seemed almost contradictory to the higher numbers of holistic 4s that the male graders gave. That is, graders R5 and R6 reported nothing positive, in general, about the papers they had scored. Perhaps on past occasions, when grading timed test writings by ESL learners for entry to freshman composition, grader R5 tended to score about 35% college-ready, and grader R6 tended to score about 45% college-ready—as on this
occasion. This might help explain why they found nothing out of the ordinary to report in this regard.

Three of the 4 females, who had scored a bit more strictly holistically than the 2 males, offered many opinions and specifics about what was good in the papers, in general terms. Grader R1 praised the best responses among those she saw: “Some of those [level] 4 papers were just as nice as a lot of things. . . . Only one or two little bitty things would let you know they weren’t native-born speakers.” Grader R2 complimented all the papers she scored, and explained that her positive reaction was based partly on her view of writing as a process:

I don’t teach ESL, but I’ve learned from the papers that I looked at. . . . I don’t see that they have more major problems. It’s simply . . . a writing process. . . . You know there’s very little that they ought to do, for instance, you know, to pass that CLAST test that they have to take. The other writing that they have to do, that would be cleared up in the process of revising and correcting.

To this she added a touch of irony: “They weren’t long enough to be irritating. You get irritated when you read so much of so much of so much of nothing [italics mine, to show verbal emphasis]. And suddenly a little bit of nothing. . . .” Apparently, she did not feel that these papers were padded or vacuous, but rather contained substance of worth.

Grader R4 praised all the writings: “They were, in many ways, charming.”

The female grader R3 was the strictest holistic evaluator in this group, giving no paper in her set of 20 a holistic score of 4. Of note, 7 of the papers in her set were given holistic 4s by other graders, and yet 5 of these resulted in splits that the chief refereed to a holistic score lower than a 4. In the post-scoring discussion, R3 reported that she respected the potential she saw, and the same potential she has seen in the ESL student papers in her own classes. This can be interpreted to mean that despite the below
college-ready caliber of papers in her set, in her view, their contents suggested to her that
in due time, many of these learners would be capable of passing freshman composition.

Trends in the Global Skill Areas

Approximations of trends in the global skill areas are suggested by the mean scores
in Table 4-2 for focus, development, organization, and clarity. Interestingly, 1 female
and 1 male—graders R3 and R6, respectively—were consistently stricter than the other
graders in their analytic grading, with mean scores of 2.55, or below, for all 4 global skill
areas. Each skill area will now be reviewed in turn.

Trends in focus

The total mean score of 2.93, for focus, is the highest scoring of the skill areas. This
shows, on a descriptive level, that for all graders as a group, expectations were more
frequently met on this global skill area than on any other. In post-scoring commentary,
grader R5 stated that students needed to learn to get to their point before completing 2
paragraphs. In Figure 4-8 are the first 2 sentences for 2 of the 3 papers (of his 20) that
this grader gave analytic scores of 2 or 1 for focus. These obscure statements typify the
sort of beginning that an ESL learner at a lower writing proficiency level tends to
produce. For this scorer, such a lack of focus called for a lower-half score.

Essay YAPA—1 What do I want to do with my life? I am not sure I know. I know, but I
am not sure I can.

LUHY—2 My ambition: what I want to do with my life. Every life of person has
ambition. But I have a great ambition.

Figure 4-8. Focus-analytic scores from grader R5 (and first two sentences)

Trends in development

The total mean score of 2.72 for development shows that for all graders as a group,
expectations were only sometimes met for this skill. Some of the post-scoring
commentary from 3 graders enlightens scoring results. Graders R1, R4, and R6 reported that development was strong in the papers. Counts of the numbers of analytic 3s or 4s graders R1 and R4 gave for development support their claims—they gave 15 and 16 upper-half scores, respectively. Grader R6, on the other hand, gave 16 lower-half scores for development—2 on papers he had even scored 4 holistically—which failed to support his claim about strong development.

Of note as well, examination of word count reinforces the widely held view of writing assessment evaluators and composition teachers that the number of words in a writing has no necessary bearing on a grader’s higher approval rating for the quality of its development. Some papers with fewer words earned high scores for development, and vice versa. Essays KAKE and SALO—both under 200 words—got 4s from 2 graders; essays IIKI and ROYY—around 160 words each—got upper-half scores; yet essay HIAR—over 200 words—got 2s on development from 2 graders.

**Trends in organization**

The total mean score of 2.67 for organization shows that, for all graders as a group, expectations were only sometimes met for this skill, similarly to results for development. In post-scoring commentary, 2 females—graders R2 and R4—stated that organization was strong in the papers. Counts of the numbers of upper-half scores they gave for this skill—at 14 and 12, respectively—supported their claims.

**Trends in clarity**

The total mean score of 2.49 for clarity shows that, for all graders as a group, expectations were only sometimes met for this skill, similarly to results for development, and for organization. This was no real surprise, as timed responses written by ESL learners tend to be weak in clarity. Early in the training, the chief had specified that
clarity was a critical factor to overall acceptability. He claimed that papers weak in this global area were intrinsically poor:

. . . so full of those kind of surface errors that you find yourself stopping and going back and rereading again and again, which is the worst thing that can happen to you when you're holistic scoring, because it really means there are some serious problems.

Looking at the mean analytic scores for clarity by individual graders shows 2 divisions. One male and 2 females were included in each. The female graders R1 and R3, and the male grader R6, tended to give lower-half scores, showing more strictness. Yet the female graders R2 and R4, and the male grader R5, had mean scores approaching an upper-half 3 on clarity, showing that these 3 appeared able to relax about language use, and even discount some inaccuracies in grammar and diction. Grader R2 gave 13 upper-half scores for clarity; grader R4 gave 15; grader R5 gave 14. It must be acknowledged, however, that these were not actual placements, but hypothetical rankings. Had these English graders been placing students into freshman composition in actuality, they may have delivered a higher proportion of low scores on clarity.

The lowest mean scores for the skill of clarity were given by 2 females—graders R1 and R3. Supporting this strictness, they both reported that clarity was the weakest skill in the papers, and gave the following pertinent commentary in the post-scoring discussion.

Grader R1 noted that issues with prepositions and articles were most notable, and went on to explain how she distinguished different levels of seriousness of ESL errors:

It’s the depth of the ESL errors, because ESL spelling doesn’t bother me and an occasional article goof doesn't bother me—I just skim way past it, and an occasional ESL misuse of preposition . . . It's when there are so many of them that I find myself rereading every line that I think, ‘How's this person going to survive?’ I suppose in the back of my head, ‘How's this person going to survive in any course
where they've got either essay exams or any job where they have to show they have facility in English?"

She also reported feeling that this scoring experiment reaffirmed that she was not grading too leniently: “It’s nice to have what I did know reaffirmed—that it’s okay to let some things slide, to focus on main idea and content . . . rather than be too extreme.” And judging from her having given upper-half scores (all 3s, no 4s) for clarity to 6 of her 20 papers, she de-emphasized errors in 30% of the essays she scored.

Grader R3 claimed that issues of clarity had more to do with sentencing, especially run-ons and fragments. She added, however, that her grading method did not penalize severely for such errors, as doing so would not seem fair and balanced:

I think all parts are important. I know I create a rubric that I use for myself, just so that I cannot . . . fail a paper because it has too many run-ons. Sentence structure isn’t 80% of the paper. . . . And . . . you’re not going to fail because you can’t differentiate between who and whom, or the object and the subject. . . . The ESL student in my class, if the student can organize and use transitions and all of that, but the grammar has some weaknesses. . . . Those are things that we can work through. It’s when the ESL student can’t organize and can’t use transitions, and can’t focus. . . .

Grader R3 only gave 3 of her 20 papers upper-half scores on clarity (all 3s, no 4s). She emphasized that ESL learners with serious errors needed to be advised into ESL classes:

I think we should also encourage our students more, if at all possible, to really take the ESL classes. . . . I know as a writing teacher I don’t have any skills in ESL. I never was trained, really, in ESL, and so I find it difficult for myself.

It was intriguing that neither grader R1 nor grader R3 gave a single analytic 4 on clarity, different from the others, who all gave between 3 and 9 of their 20 papers analytic 4s for clarity.

Grader R4 had been more generous on clarity, giving 3 of her 20 papers analytic 4s, and 12 other papers analytic 3s. She reporting regarding clarity globally, judging overall how ably the learner could communicate a message to the reader:
I ask myself, “On a scale from 1 to 4, how effective was this?” so that the different areas may change . . . but . . . how effectively did it communicate its main idea, its intent, its purpose, with a reader?

Grader R5 had also been more generous in scoring clarity, giving 9 of his 20 analytic 4s, and 6 other papers analytic 3s. He reportedly considered ideas more globally, mostly attending to the learner’s ability to express thought—creative thought and clear thought. He added that ESL students are not by default distinct from non-ESL students, in how well they express thought:

I also have a kind of a struggle between wanting to teach them how to think . . . in addition to how to be good grammar robots. I’ve had some students right out of high school who write perfect essays. There's nothing wrong with them; they're just vacant and brain dead . . . those people are going to be puppets of whatever company they work for, and I don't want to do that to them. . . . That’s something that, ESL and first language students have in common . . . just because they can’t do something . . . just because they can’t master the comma splice problem . . . the semester they have you, doesn’t mean they won’t, by the time they get their degree. And since lots of them don’t read and don’t write very much, then, I think those things take longer than a lot of people want to admit. . . . Sometimes when I read . . . one of the things I’m . . . trying to deduce how well those students think in English. When I’m reading their writing, are they thinking clearly enough in English?

The fact that half of the graders in this study gave relatively high scores to their papers for the skill of clarity—despite errors in the writings—is most encouraging. It suggests that some college-entry test graders would probably agree with the conclusion Haswell made as a result of his 1988 empirical research into errors in college student writing (detailed in Chapter 2). Haswell’s research showed that errors were inevitable in student writings at all college class levels, and that these included even basic errors that surfaced in the more sophisticated writing of students at the higher levels. This suggests that writing test scorers and writing practitioners ought to give more weight to composing skills than to grammatical and mechanical correctness, especially when evaluating students’ timed written responses.
Factors Influencing Judgments of Quality

Holistic scorings are successful due in large part to carefully planned training sessions and controlled scoring conditions, which encourage and reward consensus. It is common knowledge, however, that a grader scoring papers under uncontrolled circumstances can be unknowingly swayed by various factors. A grader might score 2 responses, seemingly parallel in quality to some other grader, incongruously, for instance, if 1 response strikes a chord—whether negative, or positive. Training is thus a crucial element in establishing like-minded scoring.

Yet notwithstanding this study’s training, influencing factors—grader empathy, idiomatic prose, a stated insecurity in English, and a shorter length—appeared to contribute to a handful of cases in which papers were given seemingly too high or too low holistic scores. These cases will now be discussed and illustrated.

Grader Empathy

Undeniably, human elements that writers leave in their responses—dry wit, vigor, enthusiasm—can touch a chord in a grader, which in turn encourages the reward of a high holistic score. Or some expressed attitude or character trait is judged important, and then the grader’s heart goes out, which may aid in overlooking weaknesses. This appears to have occurred with the 3 third best-scoring essays reviewed earlier in this chapter—HIAR, KAOA, and KKMA.

Idiomatic Prose

The English graders in this study generally rewarded the ESL writer able to produce a complete and on-topic response in moderately fluent English, notwithstanding weaknesses in certain rhetorical or grammatical features. While my study’s Holistic scoring guide (Table 3-1) encouraged this result, these graders might also defend these
scores, arguing that rewarding such an ESL response is not by default gratuitous, but justifiable, as it recognizes the communicative abilities the learner ably demonstrates.

In contrast, the ESL holistic scoring guide (Table D-1) formerly used by ESL placement teachers at this junior college discouraged their releasing entrants into mainstream English when their writings demonstrated an unconventional essay form, sporadic grammar errors, or various unsupported arguments, notwithstanding a general fluency in English. Many an entering ESL student at this college who failed to demonstrate some knowledge of rhetorical conventions, who inadvertently sprinkled errors throughout the writing, or who strung together generalizations with no supporting facts, examples, or other details, risked being placed into advanced ESL class, in accordance with the guide for placement that ESL graders used here. These graders hold that skills in formulating logical arguments, and in supporting them well can be learned in advanced ESL, which by design is much more than a grammar course. And they argue that it is better for the student, in the end, to gain those skills prior to entering college classes. Some students put off taking their composition courses until later in their college careers. They will quickly realize their disadvantaged position: Lacking abilities in formulating arguments, and formal composing, how can they expect to do well in mainstream courses outside of freshman composition, many of which require some writing?

The fact that the English graders in this study rewarded an ESL learner’s fluency in English may owe to the sort of papers they generally encounter in their own classes. Most of their students are native speaking; therefore, fluent, communicative English is probably commonplace among the writings they see. Weak development and
organization, and a lack of formal features, are probably equally common. This may incline them to pass an ESL writing that is idiomatic, for its facility with the idiom.

One example of this sort of acceptable ESL writing for English teachers is essay SALO (given in Appendix A). Graders R1 and R4 both gave this essay a holistic 4, and both gave it high analytic scores for all global areas. Considering its length—at 198 words—it is rather remarkable that this writing has only 1 article error—“lose the interest”—and 1 preposition error—“on the computer and fine arts fields”—signaling notable fluency. Another sign of fluency is the relative frequency of contractions: “that’s,” I’ve,” I’d,” “don’t,” “there’s.” One sign of a lack of fluency is a second language learners’ inability to recognize or produce English contractions.

English as a second language placement graders at this college had put essay SALO into advanced ESL. Aside from being dissatisfied, generally, with this paper’s relatively shallow development, these particular graders tend to disapprove when test takers break too many formal composing conventions. The student writer of essay SALO had broken several. Using all capital letters, this learner seemed to be masking an ignorance of the conventions for capitalization. Indenting approximately every 3 sentences, she displayed a complete lack of awareness of paragraphing rules. And not until the end did she include a transitional expression (“FINALLY”). In sum, the ideas in SALO come across as a bit disjointed and muddled, which most probably contributed to the ESL grader’s placing her in advanced ESL, despite her fluency.

A Stated Insecurity in English

In the case of essay IIKI (given in full in Figure 4-9), 2 English graders may have scored this writing low holistically partly because of the learner’s expressed insecurities about his English:
At the moment, 我 speak pretty good English but I don’t know difficult words and business talk. So, probably if I study English for another year, I will be able to do the job. I hope I will improve my English soon.

Belittling one’s own language skills on a writing test is self-defeating and seemingly submits to a lower holistic score. In borderline cases, this could be especially risky. In contrast to the pre-score at college-ready level that IIKI was given by an ESL grader at this college, essay IIKI got 3s from both English graders in the present study.

After I finish studying English, I’d like to work in my father’s company which is in Japan. His company makes molds for plastic products like cameras, walkmans, cd cases etc… He has some business between other countries. So he needs someone who can deal in English. That’s the main reason I am going to study English language. I also have studied design in England which will help me to do the job. 我 and my brother are already working in his company. There is still a big economy depression in Japan but I’d like to make a good family business and help my father. At the moment, 我 speak pretty good English but I don’t know difficult words and business talk. So, probably if I study English for another year, I will be able to do the job. I hope I will improve my English soon.

Figure 4-9. Essay IIKI (given holistic 3s by R3 and R5)

Let’s compare a similar paper—essay ROYY (given in full in Figure 4-10)—and its scores. Graders R1 and R6 gave essay ROYY a holistic 4. Note how this learner touts his facility with the English language, and with other languages as well: “I picture myself in the future as a person who is able to perform and excellent work as a translator mainly because I have a good ability to learn languages and I love learning them.”

One of my biggest dreams since I was in sixth grade has been to become a translator. I love languages. Therefore I think that would be my ideal job. I picture myself in the future as a person who is able to perform and excellent work as a translator mainly because I have a good ability to learn languages and I love learning them. I also think communication is one of the things we have to practice the most these days and working as a translator would contribute to make that happen. It would help us understand each other a little more.

Besides, being a translator would allow me to meet all kinds of people from many different places. I find it very interesting and exciting to know a little more about the different cultures we have in the world.

I have worked very hard to be a translator and I think I will be able to make it.

Figure 4-10. Essay ROYY (given holistic 4s by R1 and R6)
The data in Table 4-4 show just how similar essays IIKI and ROYY are, linguistically. And language use was a chief determinant in the pre-scores the ESL graders assigned. Essays IIKI and ROYY had nearly the same number of words, sentences and sentence types. As for errors, IIKI had 9 and ROYY had 5, which is the primary notable difference between them, linguistically. IIKI used 2 wrong verbs—“has” (for “does”) and “deal” (for “do”)—while ROYY used 2 wrong verbs—“be” (for “become”) and “make” (for “do”)—and 1 wrong verb form. ROYY had no preposition or article errors. IIKI contained 1 of each, but demonstrated a notable skill with these features, given that she is Japanese (learners with this L1 generally acquire articles only after extensive study and years of practice).

<table>
<thead>
<tr>
<th>Essay code</th>
<th># Words</th>
<th># sentences</th>
<th>Sentence Types</th>
<th># Errs</th>
<th>Verb or word ending</th>
<th>Mechanical Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROYY</td>
<td>159 words</td>
<td>9 sentences</td>
<td>complex—cx</td>
<td>5</td>
<td>be [become]</td>
<td>and [an]</td>
</tr>
<tr>
<td></td>
<td>159 words</td>
<td>9 sentences</td>
<td>compound—cd</td>
<td></td>
<td>make [making]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>159 words</td>
<td>9 sentences</td>
<td>simple—s</td>
<td></td>
<td>make [do]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>159 words</td>
<td>9 sentences</td>
<td></td>
<td></td>
<td>al [all]</td>
<td></td>
</tr>
<tr>
<td>IIKI</td>
<td>157 words</td>
<td>11 sentences</td>
<td>complex—cx</td>
<td>9</td>
<td>has [does]</td>
<td>(missing [the])</td>
</tr>
<tr>
<td></td>
<td>157 words</td>
<td>11 sentences</td>
<td>compound—cd</td>
<td></td>
<td>deal [do]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>157 words</td>
<td>11 sentences</td>
<td>simple—s</td>
<td></td>
<td>sister [sisters]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>157 words</td>
<td>11 sentences</td>
<td></td>
<td></td>
<td>brother [brothers]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>157 words</td>
<td>11 sentences</td>
<td></td>
<td></td>
<td>economy [economic]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>157 words</td>
<td>11 sentences</td>
<td></td>
<td></td>
<td>studding [studying]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>157 words</td>
<td>11 sentences</td>
<td></td>
<td></td>
<td>language [language]</td>
<td></td>
</tr>
</tbody>
</table>

Knowing that the learner who wrote ROYY was a speaker of Spanish natively, an ESL grader probably did not consider her ability to correctly use articles and prepositions as much a coup as the English graders may have. Further, the ESL grader probably did not see essay IIKI as markedly weaker—to her it did not warrant a lower holistic score—since its errors were reparable sentence-level problems that did not obscure meaning.
Moreover, the learner who wrote IIKI demonstrated a clear capacity to use effective English—the essay had unity, sentence variety, accurate word choice, verb usage, and good mechanics. Yet English graders R3 and R5 scored essay IIKI below 4 holistically, despite thereafter giving it all upper-half analytic scores (suggesting that holistic 4s might have been more appropriate).

Perhaps the number and type of errors in essay IIKI also persuaded graders R3 and R5 to score it below 4, holistically. The extra errors in IIKI were of types putatively irritating to English teachers: article, preposition, nominal plural, and word ending errors.

Shorter Writings

Additionally, these English graders tended to score inconclusive or comparatively short responses below 4 holistically. To illustrate, we have essay ENNT, which graders R2 and R4 both scored below 4. It is of note that during the holistic training, the chief had expressly cautioned the graders to avoid judging papers by their length:

By the way, just because a paper is short, doesn't mean it's undeveloped; some short papers are developed very nicely. However, it's liable to get you off, as soon as you see a short paper. . . . “This is going to be a bummer.” Read it before you make that decision.

In the post-scoring question-and-answer session, the chief reported being struck by the fact that the graders had scored this below a 4 holistically, when he had considered it to have the best command of the English language of the lot:

I thought it was the best paper of all the ones I had read . . . ENNT. It was very short. It was a very short one, but boy did that person have command of the language! . . . I told you the problem with that is it was very short, well, very short on the page, but it was dense, very densely written and clearly the person had command of the idiom, and I would venture to say that that person would take your Comp I class and make an A.

And knowing how difficult it is for a non-native speaker to compose an impromptu response under a strict time limit, ESL teachers at this junior college, and some elsewhere
no doubt, would consider it indefensible to place a second language learner with such a command of English into ESL. A learner whose principal weakness is the speed of composing does not need ESL coursework, they reason. Besides, writing timed responses under test conditions does not comprise the bulk of the requirements in freshman comp. To an ESL grader, the abrupt halt of essay ENNT merely underscores the student’s compliance with the test proctor’s directive to stop writing. Beyond its grammar control and precise diction, this 159-word essay (Figure 4-11) offers a playful tone, nice detailing for main points given, and a visible penchant for reflective thinking.

It is probably common for many people to be asked what they want to do or to be in the future. But how many of them would be able to come up with the answer right away? I am one of those whose goals are yet to determine. I used to dream of being a medical doctor, but later found out that Physics, Chemistry, and Biology are on top of my hate list. I used to dream of being a flight attendant, but my height and weight are not in the required proportion. I also used to dream and still dream of being a journalist or a columnist, but my writing skill doesn’t allow me a chance to the career. Well, what else can I be? And I keep asking myself that question every now and then. The right answer hasn’t come to my mind yet. I’m pretty tired when it comes to that thought. For one sure thing

Figure 4-11. Essay ENNT (given a holistic 3 by R2, a holistic 2 by R4)

Splits—Factors and Trends

Since this study set out to analyze best ESL papers, 33 papers with pre-scores corresponding to the advanced ESL writing proficiency level (my scoring guide’s holistic 3) were included. Many of these 4 training and 29 scoring papers resulted in split-pair rankings. Two interesting cases—essay KAKA (256 words) and essay PAII (267 words)—are given in Figure 4-12 and emerged as splits in the holistic training. What strengths and weaknesses could have prompted some graders to score these 4 holistically, but other graders to score them lower? Though a lack of time precluded their being scored analytically, analysis of variant features allows for conjecture about why some graders rewarded these essays with holistic 4s, while the chief put them at holistic 3s.
Essay KAKA

My ambitions

I have three things that I really interested in and want to do with my life which is voluntary activities, to be a bilingual person and to work at a company with using English.

What I’m really interested in is taking care of elderly people. Recent years, the number of old people are increasing and they have a lot of problems to live alone which are big issues in my country. What I can do for them is a little thing but I think I can join voluntary activities in a sort of nursing home and talk to them and encourage them sometimes.

The second one is to be a person who speak two languages, Japanese and English. English is a common language in the world. If I speak and be good at grammar of English, I can get a job in my country easier.

The last one that I want to do is working at a company which offers connections to go abroad. People who are planning to go abroad have a lot of worries about the way of living, custom and also language that you are going to go. I’d like to support them and help them through my experience of going abroad in the United States.

These are my three ambitions in my life. All I need to do now is to participate in a volunteer and how to take care of old people and also keep studying English and improve my English skills here in the United States of America.

Essay PAII

WELL, BASICLLY I CHOSE TO COME AND LIVE IN THE UNITED STATES BECAUSE OF THE INFINITE OPPORTUNITIES THAT LACK IN MY COUNTRY. I CHOOSE TO LIVE IN FLORIDA BECAUSE I THINK THAT THE COLLEGE EDUCATION AND THE FINANCIAL AID PROGRAMS ARE ONE OF THE BEST IN THE STATES. I’D LIKE TO PURSUIT MY DREAM TO BECOME A BUSINESS MANAGER IN THE FUTURE. TO COMPLETE MY DREAM I’LL NEED HARD WORK, GOOD WILL AND STUDY. BUT I KNOW I’M GONNA GET THROUGH. I ACTUALLY WORK AS A DELI CLERK. IT’S NOT THE GREATEST JOB, BUT STILL I HAVE A CHOICE TO APPLY FOR MANAGEMENT POSITIONS. DURING THE JOB PROCESS I CAN SEE THAT MY ORGANIZATIVE SKILLS PREVAIL ON THE OTHERS WORK. AND IT’S NICE TO BE APPRECIATED FOR WHAT YOU ARE DOING IN ORDER TO HELP YOUR FELLOW ASSOCIATES. BUT I’M NOT PLANNING ON A DELI CAREER. WHAT I LIKE MOST IS BEING THE MANAGER OF A BANK OFFICE. SO I CAN WORK AND DIRECT IN DIFFERENT ASPECTS OF THE ECONOMY. I LOVE TO BE ABLE TO UNDERSTAND ALL THE FINANCIAL PROCESSES. FROM BANK ACCOUNTS TO STOCK HOLDERS. A WORK AT THE BANK WILL LET ME DO WHAT I REALLY LIKE, FOR EXAMPLE WORKING WITH PEOPLE AND CUSTOMERS, BEING ABLE TO GIVE CONSULTS AND ADVICES TO EVERYBODY. AND WHEN YOU UNDERSTAND ECONOMY YOU ARE ABLE TO UNDERSTAND POLITICS TOO. I’VE ALWAYS BEEN ATTRACTION TO POLITICS ISSUES AND MOVES. IT’S INTERESTING. I KNOW THAT WHEN I’M GOING TO FINISH MY STUDIES I’LL BE ABLE TO FULLFILL MY DESIRES AND DREAMS, BUT IT’S STILL A LONG WAY AHEAD OF ME.

Figure 4-12. Training essays KAKA and PAII (split-scored, resolved to holistic 3s)
In essay KAKA, several strengths and weaknesses are evident. One strong point is that this writer conforms to the rhetorical conventions of formal writing. The paper starts with a clear purpose, and a plan for how this thesis will be developed and organized into 3 parts. Transitions and effective paragraphing follow, and some details help develop each subtopic (the second somewhat less in depth), mostly fulfilling the promise made at the beginning. Several strikethroughs indicate that the learner took time to read and modify code as he drafted ideas. The learner further reserved enough time to write a concluding paragraph with 2 sentences, the second of which reiterated the 3 main ideas presented, showing his ability to reaffirm main points and provide closure.

Despite all the strengths noted in essay KAKA, however, some graders would be reluctant to score this writing a 4, for its inconsistent use of correct “be” forms and its weaknesses in prepositions, evident immediately. The first 3 sentences reveal the omission of “am” where warranted, the use of “is” where “to do” is needed, a superfluous inclusion of the preposition “with,” and the omission of a needed preposition, “in.” Some misspellings also occur, but for the most part the vocabulary, verb usage, and syntax throughout the response are grammatical and appropriate for formal writing. In sum, the chief’s directive for graders to consider this a holistic 3-paper seemed well-justified.

Regarding weaknesses evident in essay PAII, this paper indicates an ignorance of the expected features of an academic written response. Arguably, the writer of essay PAII masked her inability to apply capitalization rules, by using all upper case—certainly her habitual penmanship. She further displayed occasional misspellings and wrong words (BASICLLY, OPPTUNITIES, TO PURSUIT MY DREAM, GONNA,
ORGANIZATIVE, FINANCIARY), no notion of how to paragraph, and a virtual lack of transitional expressions—all leaving her ideas disconnected and seemingly unplanned.

As for the strengths in essay PAII, some evidence reveals that this individual is well assimilated culturally—employed and hard working, striving for a better job, rather fluent in English—and therefore ready enough for mainstream college coursework. Nonetheless, the chief’s assessment of this paper as a holistic 3 seems accurate, as this close analysis suggests that in 1 semester of ESL coursework, this learner could improve the noted weaknesses, and become better prepared for writing at the college level.

Additionally, essay OAPA (at 233 words) and essay GACS (at 242 words), given in Figure 4-13, emerged as splits, as defined in this study. These each yielded a holistic score of 3 from 1 grader, a 4 from another, and were refereed to a 3 by the chief.

Strengths in both OAPA and GACS include: an adequate length (over 200 words), a stated focus, several main ideas, some supporting details, somewhat fluid style, and some form of conclusion. Weaknesses contained by both are mostly characterized by ESL-type language errors: omitted word endings, subject and verb errors, sentence punctuation errors, and improper paragraphing.

While 2 female English graders scored these writings below a 4 holistically, 2 males disagreed. Grader R6 rewarded essay GACS with a holistic 4, as grader R5 rewarded essay OAPA. Later, the low analytic scores she gave on organization and clarity showed weaknesses grader R1 perceived in essay GACS. She marked the paper in 3 places: “(body and brain) working together,” “Fannattics,” and “How is the real life?, the hard life?” And grader R6’s analytic score of 2 on the global skill of clarity revealed that he considered essay GACS unsatisfactory in this skill area.
Essay OAPA

I would like to be a marine biology because I always like animals but specially sea animals one of my dreams is to work in “SEA WORLD” or in any big aquarium, because I love whales, seals, etc.

All began when I came here for vacation when I was a kid, and I had the opportunity to see the whales very closer, I saw them in a show, they were jumping up and down, and they were very friendly. It was really funny. In that moment I change my decision, because before that, I wanted to be something related with science like cloning plants or vegetables, or maybe a doctor but I wasn’t sure. So, when I saw how the persons work with sea animals, I like it very much.

I hope some day I’ll have my own aquarium with all kind of fish, and all animals that lives on water.

In my country, Ecuador are the Galapagos island, I like to go there, is really amazing to see all the creatures, and the way they live, and how they care about their baby’s. I always see the TV programs about this topic and I like to know all the problems that Galapagos island has with the people that doesn’t care about them, and I want to help them, and to protect every little creature.

Essay GACS

Describe a job you would like in the future. Explain why you would be good at that job.

I would like to work in the future like a Baseball Pro Player, because this is my dream since I was a child.

I like it because in this work, you need to have Athletic talent and you need to be smart, very intelligent.

So, I think that I can be (body and brain) working together at the same time. I know that the way to be a Baseball Pro Player is very hard, so you should have three thoughts in your mind: Discipline, Dedication and Love for the game.

In fact, when you are a Baseball Pro Player you must have good manners because you’ll have communication with the Fannatics, News-Paper, T.V., Radio and specially with children.

After that, when I finish my Baseball Pro Player Career (like an active player), I would work with children and boys who play Baseball like a Coach and I want to teach them all I know and How is the real life?, the hard life?

I think that I could be good at this job, because I was working for my dream since I was 3 years old and now, I’m working harder than other years. How? I’m studying English and I will enrollment at this college in January; I’m going to The Gym and I practice Baseball every day, and I love working with children.

Figure 4-13. Essays OAPA and GACS, splits resolved to holistic 3s (OAPA given a 4 by R5, a 3 by R2; GACS given a 4 by R6, a 3 by R1)

Two other interesting splits (Figure 4-14) were essay LUHA (at 439 words) and essay PYMA (at 130 words), were given 2s holistically by female graders R1 and R3, but 4s holistically by the male grader R6. Finally, both were rectified to scores of holistic 2s by the chief. The composing in these shows some level of communicative competence. Yet the extent of the breaks to basic conventions in grammar and mechanics underscores the need for work on such issues prior to entry to mainstream college classes.
Essay LUHA

In 1996 I was came from Vietnam with my family. Since the first fews days we don’t know even a word of English, and after I went to middle school 8 grad. I try very hard for everythings. In the first few days in school. Some of students in class they’re teas me, I feel very upsad an ambarass, but I’m not give up, I told my self I have to learned somethings for my self, I’m not give that easily, because I want to be a nurse when I grow up to help the people who going to need my help.

I did learned a lot from school. Time pass very fast when the first day I’m started went to school was 8 graded, and now I’m graduate, I’m very prove my self. I can’t believe that I can finish the high school. In high school everything is hard because my English not very well. So everything for is hard ever had in my life.

When I’m senior years in high school the school law says all the seniors have to do the senior protected. I was suprise and afraid of, because I even know what the senior protect was, and also I never done of my life. Once against I’m not give up, I try very to done my senior protect. I do ask my teacher for help and also in the ESOL teacher they help to correct the, after that I found out only to do the protect, I do have to presentation. I’m so scare. I don’t think that I can presentation, but my teacher you going to be okay. You best take a deep break and you will be fine. I listen to her.

Day by day the date presentation come up. I presentation in front of 3 teacher. I’m scare to death but I listen to my teacher say take a deep break slow present. Final I’m done my senior protect. I got a 93% I think I did a good job. I try my best to learned and done the big senior protect. I graduate and I came back to the country I use born there. I’m so excited that I can take rest, don’t have to worry about senior protect any more. I live there 2 months. Everything there has change I almost forgot. like my brothers and sister they get older. and now I want to apply for college, hopefully that I can done a good job, but I still afraid that I’m going to fair, but I best my best doesn’t matter what the result is.

I'm sure

Figure 4-14. Essays LUHA and PYMA, splits resolved to holistic 2s (LUHA given a 4 by R6, a 2 by R1; PYMA given a 4 by R6, a 2 by R3)

As for essay LUHA, grader R1—who gave it a holistic 2—gave it analytic scores of 1 for clarity and 3 for focus, development, and organization. She substantiated her 1 score with 12 notations written over the first half of this paper. She marked 5 verb errors, including “was came,” and “I’m not give up.” Four word ending errors she
marked included “afraided” and “fews.” Two wrong words that caught her attention were “prove” (for “proud”) and “upsad” (for “upset”). And a phrase that she marked was this one: “fast when the first day.” Noticeably, grader R1’s 3 upper-half analytic scores were discordant with her holistic 2 score. Perhaps her efforts to stay focused on meaning (not on language) influenced her to give analytic scores that were unduly high, given the volume of weaknesses.

Essay PYMA, grader R3 gave a holistic 2—and analytic 2s on clarity, development, and organization, but a 3 on focus. She only marked the margin in 1 spot, beside an idea toward the end of the essay: “I have to many things in my mind, one of them is first study hurt and try to learn every day more and more.” This grader could have found problematic both the generality and the errors in this idea, which, by the way, the writer of essay PYMA had set apart as a paragraph. This learner showed no sense of how to paragraph, nor of how to support arguments with details. Grader R3’s 3 lower half-scores reflected her opinion that this learner needed additional ESL work.

In contrast, grader R6 gave holistic 4s to essay LUHA and essay PYMA. But later, upon the more deliberate analytic scoring, he gave all lower-half scores on the 4 global features for both essays. He scored both 2 for clarity—though he put no marks on either. He gave LUHA a 1 for focus, development, and organization. He gave PYMA a 2 for these. In essence, his analytic scores for LUHA and PYMA corresponded better to the chief’s rectified holistic scores of 2 than they did to his own earlier holistic score of 4s.

An additional interesting split case is essay TNGA (at 193 words)—refereed to a score lower than a holistic 4 by the chief (given in Figure 4-15). Grader R5 gave it a holistic 4 and 4 analytic 4s—on all global skill areas. And grader R2 gave it an
uncharacteristically high proportion of upper-half analytic scores—4s on all global skills except development—which she gave a score of 3. If her analytic scores validated her claim, then some weakness in this essay’s development, alone, help account for her scoring TNGA below a holistic 4. A closer look at this paper shows that some exemplification and unity exist. Yet the relatively shallow development could have influenced grader R2 to give it the holistic score of 3 and the analytic 3 on development. This college’s ESL grader probably also saw the development as too shallow to justify a 4, holistically. Grader R5’s holistic 4 might have been influenced by empathy or similar thinking to that which this examinee expressed.

A final interesting case—while not a split per se—was a paper that earned holistic 2s from both graders, but then got all 8 analytic scores of 3 or 4. This was essay HZJE, at 290 words, given in Figure 4-16. Upper-half analytic scores on development and focus from graders R2 and R4 seemed accurate enough, given that this writing made a promise
at the beginning and mostly fulfilled it by the end, and given that the volume of relevant detail included was sufficiently supportive.

Last year when I finished my high school all the people ask me about my future plans and all of them ask me the same question “What you want to do with your life?” I was afraid to answer that question because I didn’t know what I want to do with my life, but now almost a year after, my mind is more clear. Now I have more ambitions. I hope that all my dream come true. First I want to learn english second I want to study Medicine and my finally ambition is difficult but not impossible; I hope, in a nearly future that my family would come to live here, in America. When I were in high school, the english was so important, and I seems the opportunity to study here, and now after all the troubles, I’m waiting for enter in the college. When I arrived to florida, I supposed to enter at the college but, I didn’t have the paper so I was waiting for this paper almost 7 months. but now my dream come true. After I would finished this program I want to study Medicine, because I like to help all the people, and it is the way to help the persons. I know that woma be a doctor is hard but I have the capacity and the necessary intelligent for in the future would study this carreer. My parents told me that will be so hard for my, but I like it, and it is important for me wanna be a doctor.

And my finally ambition is hard, but I hope that I will come true. The situation in my country are difficult so I want to help my family . . .

Figure 4-16. Essay HZJE (given holistic 2s, but analytic 4s and 3s by graders R2 and R4)

However, it seemed surprising that these graders both gave essay HZJE the satisfactory upper-half score of 3 for clarity, despite the abundant language errors that left many ideas, especially in the second half, vague and undeveloped. Another apparent weakness, the writing jumped from past to present, to future, back to past, to an older past, back to future, and so on. And paragraphing was virtually non-existent; attempts to form paragraphs were clearly unprincipled. In essence, no apparent form of organization existed, and yet, each grader had also given HZJE an upper-half score of 4 for organization. All in all, these graders’ holistic 2 scores for HZJE appeared more accurate than their upper-half scores for clarity and organization seemed to be. That is, holistically, both graders indicated that weaknesses predominated, despite other strengths
in this response, and that this learner needed more language practice before freshman composition.

One interesting finding arises from the chief’s resolving 80% of the splits to ESL levels. The contents in the split papers more often diverged from the consensus expectations for a holistic 4 (college-ready writing). To describe prevalent weaknesses evident in the splits would contribute to the delineation of features of weaker ESL timed writing, so 4 splits—essays GACS, OAPA, LUHA, PYMA, which yielded from essays pre-scored at intermediate or beginning ESL levels—have been reviewed in this section. It is also important to carefully consider the qualities in the 3 split papers refereed to holistic 4, however—essays HIAR, HOSE, PTGE. Their strengths obviously suggested college readiness to the chief and to other graders, so these were also fully detailed earlier in this chapter, with the discussion of the level-4 essays.

**Essays at ESL Levels**

One additional important result—though not a primary purpose of this study—is the division of ESL writing into graded levels, allowing the advanced, intermediate, and beginning ESL writings (49 writings in all) to be distinguished from the highest group—11 ESL writings that were ultimately ranked college-ready by these English graders. Numerous essays emerged at the 2 lowest levels—intermediate and basic ESL. It seems useful to present 4 additional beginning-level ESL essays (in Figure 4-17) at this point. Presented and discussed earlier were 2 essays ranked at the intermediate ESL level by these English graders—OAPA, GACS, and HZJE—and 2 essays split-scored at first, but then resolved to also be at the intermediate ESL level—LUHA and PYMA.

What sorts of issues prompted these English graders to give a timed essay written by an ESL learner a holistic rank score of 1 (beginning ESL)? The following exchange
**Essay RZAA**

THE MOST IMPORTANT THING FOR ME IS BE THE BEST IN ALL THAT I DO. MY LIFE CAN BE EXPLAINED IN 3 THINGS --MY MAGISTER, MY PARENTS AND MY OWN FAMILY.

I ALWAYS HAD A DREAM, GO TO THE AIRFORCE BUT FIRST I’M GOING TO STUDY TO LOWYER, I WANT TO BE GOOD IN MY WORK TO GET MONEY AND HELP MY PARENTS, I WANT TO BUY HOUSE AND GET MARRIED WITH SOMEONE WHO LIKES THE FAMILY LIKE ME, AND HAVE TWO BOYS AND A GIRL. I’LL HAVE GOOD MONEY BECAUSE I’LL STAY WITH MY BABY HIS FIRST LIFE YEAR ALL TIME AND THEN WORK PART TIME. I WANT THEY GO TO KNOW MY COUNTRY ANYTIME OF THEIR LIFES. I WANT TO BE MARRIED WITH THE SAME MAN ALL MY LIFE, AND GO TO DANCING WITH MY SONS. THAT IS WHAT I WANT TO DO.

**Essay COIO**

My ambition is to go to College for two years to study for a Veteneran. and if I do good in College move on to a Unv in Texas. Work there and help my dad by sending him money to mexico to fixs his ranch. After three more years of school and know a Vet. I would like to look for a place to live. Some where by the beach close to Mexico. When all settold down I would like to get meirred. but I leve it all to god; and I pray for this to come true. Because only he can make my Ambition come true.

**Essay FAZA**

Everybody have different ambition

When I was young, I was thinking what I want to do with my life. On the one hand I would like to have happy personal life. On the other hand have own career. Look for success. I want to complete degree. That prepares me to enter the workforce immediately. I want to improve job skills. I plan to have happy life.

**Essay OOPO**

What I want to do with my life.

I want will make a major in United States. When it happen I hope can living in Florida and working without any problems. If this is not possible, I will come back to Colombia my country and will work in my major.

I will want to study “sports training”. It is who preparats in physical condition the sportsmen of a team. I like soccer in special. But, right now I want to learn English and after I will think in which University to study. In conclusion, I will live in Florida and will marriage.

**Figure 4-17. Four additional beginning-ESL level essays**

among some of the graders and the chief seems to encapsulate well their thoughts about the comprehensibility of portions of essay RZAA, one such beginning ESL-level writing:
R1: “My MAGISTER” … what the heck is that?

CR: Yeah.

R4: What is “magister”?

CR: I put a line out there too.

R1: “His first life year all time," I could figure it out, but I had to reread it.

R4: Yeah, I marked that too.

CR: Down here, "and have two boys and a girl . . . I'll have good money because I'll stay with my baby.” That's really kind of clumsy writing.

R1: "I want they go to know. . . ." I'm sure that what the person meant was, "I want them to grow to know my country," but that's not what it says. And just because if I reread it, I intuit what it means, somebody else would intuit. . . .

R4: Well, I think that's where I got thrown off the mark. I assumed that if I could intuit it . . . then that's not good writing, but it has gotten its point across, which is ultimately what I'm looking for.

R1: I guess what I have in the back of my mind is that I'm used to interpreting the goofy syntax that comes through my class, so I'm trying to think of it . . . would somebody who's NOT a comp teacher be able to figure it out, in a professional situation.

CR: But what if, say, a student who comes to you and says, but you know what I mean.

R4: If I don't know what he means, then ok, and generally, I don't put words into their mouths. I do not make that jump, but I didn't have to make a jump in any of these papers.

Some Afterthoughts about Accord

Since the holistic scores—for dividing the papers into non-ESL and ESL categories—and the analytic scores—for informing about global qualities—were all hypothetical, accord was not of primary import in the environment established for this study. Nonetheless, considering accord in the 2 scorings would show how well the resulting scores represented views among graders. In the holistic training, all graders
scored all 14 samples. Accord existed whenever 4 or more of the 6 graders gave the same score to a training sample, constituting majority agreement. Accord was achieved on most level-2, level-1, and level-4 samples. Three level-4 anchors were established: essays GKUA, OIVM (in Appendix A), and essay PAMA (at 197 words, in Figure 4-18).

Strikingly, all

I already have quite a good idea about my future job. Finding a job that suits me well is very important for me as I would like to become a real professional and enjoy my work as much as possible.

I would like to work in a public relations agency or in a public relations department of some company.

I already have some experience in PR. I have worked as an Account Executive in a PR agency back at home in Czech republic. I have learned a lot of important things that are necessary for PR – my duties included writing essays for different types of newspapers, media relations, client meetings, developing PR strategies for clients etc.

What I enjoyed the most about my job was the possibility to meet many interesting people, to be able to improve my writing skills and to be an important part of our team.

I would love to become an Account Manager in the future, which would include more responsibilities but it would also mean a more independent position.

I hope my studies at this college will help me to achieve my goal and open new opportunities for myself.

Figure 4-18. Training 197-word essay PAMA (pre-scored holistic 4)
graders gave essay PAMA a 4, and the chief strongly commended it:

This is the kind of guy or gal you want to get in your Comp I class. . . . One of the things that strikes me . . . awfully good content. . . . This person knows what he’s talking about, or what she’s talking about. . . . I’ve read enough writing by professional people who can’t do the parallel structure as well as this. . . . Ok, that’s definitely a 4-paper. In other words, we want this person in Comp I.

Two female raters wrote comments on the analytic scoring sheet: “generally clear sentence patterns,” “focus and specifics,” “well thought out,” and “moves from past to present well.”
Discord predominated, however, in the holistic training on all samples pre-scored 3 (advanced ESL); some even got holistic 4s. Recall that all holistic score pairs of a holistic 4 with any lower score were defined as discordant (see Table 3-2 in Chapter 3)—to ensure that the best papers (4s) could be first identified definitively, so that their global strengths—which were of great import—could be most closely analyzed. All splits containing only 1 score of 4, thus, had to be resolved. Table 4-5 charts the discord resulting on all 4 training samples with the holistic pre-score of 3. Despite the clear lack of accord resulting, however, the chief designated all 4 of these to be level-3 anchors (2 of these—essays KAKA and PAII—were discussed in detail earlier in this chapter).

Table 4-5. Discord among six grader scores—for all pre-scored level-3 samples

<table>
<thead>
<tr>
<th>Pre-Scored level-3 anchor</th>
<th>Level-4 (college-ready)</th>
<th>Level-3 (advanced ESL)</th>
<th>Level-2 (intermediate ESL)</th>
<th>Level-1 (beginning ESL)</th>
<th>Final level designation for this anchor</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAXA</td>
<td>3 graders</td>
<td>3 graders</td>
<td>--</td>
<td>--</td>
<td>3</td>
</tr>
<tr>
<td>ZVOA</td>
<td>--</td>
<td>3 graders</td>
<td>3 graders</td>
<td>--</td>
<td>3</td>
</tr>
<tr>
<td>KAKA</td>
<td>3 graders</td>
<td>3 graders</td>
<td>--</td>
<td>--</td>
<td>3</td>
</tr>
<tr>
<td>PAII</td>
<td>2 graders</td>
<td>3 graders</td>
<td>1 grader</td>
<td>--</td>
<td>3</td>
</tr>
</tbody>
</table>

Then in the holistic scoring, a satisfactory 75% accord resulted (graders agreed on 45 of the 60 papers). Level-3 papers yielded high rates (40%) of splits, as they had in the training. This indicated that the advanced ESL writings were the most difficult for these graders to score holistically (details about the discordant 15 papers are shown in Table E-1).

In the analytic training, the issue of score variations immediately surfaced, and grader R2 reported her own assumption: “It seems like they [the analytic scores] are validating the holistic score, but they’re a little bit more accurate.” The chief carefully explained that papers with low holistic scores might get low and high analytic scores, but
added that papers with high holistic scores should only get high analytic scores. Time allowed for scoring only 6 samples. On the 3 training papers with holistic 2 and 1 scores, most graders gave an overabundance (3 or 4, of 4) of high analytic scores. This suggested that some papers with low holistic scores would probably get discordant analytic scores. Yet grader accord was outstanding on the only holistic 4 training paper scored analytically—PAMA. Four graders gave it analytic 4s for all 4 skills; the others gave 3s and 4s. Recall that all graders had given essay PAMA a holistic 4. The scores and reactions to essay PAMA were consistently favorable, further raising confidence levels in the accuracy of upper-half analytic scores for papers with 2 holistic 4s—they should be trustworthy.

Then in the analytic scoring, discordant scores resulted for 3 of the 11 holistic 4s (discussed at the beginning of this chapter). Yet on the 49 scored below 4 holistically (Appendix B), only 1 essay got discordant values—essay HZJE (discussed with splits, earlier in this chapter).

Patterns of analytic scores emerged and are noticeable when one views the essays grouped by their holistic scores (Appendix B). Essays scored the highest holistically tended to earn greater numbers of upper-half analytic scores, and vice versa. Essays that earned 1 or 2 holistic 3s got more analytic 3s and 4s than analytic 1s and 2s. And for essays that earned holistic 2s, the trend reversed. The essays that earned 1 or 2 holistic scores of 1 show the clearest patterns—an abundance of analytic 2s and 1s.

**Summary**

A variety of strengths in 8 of the 11 holistic-4 ESL essays set them apart. Two essays earning the greatest number of analytic scores of 4 were similar globally, but distinct in another way. Essays ZLDA and KAKE adhered well to formal essay style—
an explicit focus, full development, given the time limit, notable organization, and clear language. Both also demonstrated control of cognitive / academic language proficiency (CALP), discussed in detail in Chapter 2. Recall that CALP is characterized by grammatically and syntactically rich sentences, and is often tested for by institutional placement tests. These 2 essays showed learner facility with CALP, both in rich sentencing, and in academic word usage with correct word endings. Yet in essay ZLDA, the learner’s clear finesse with the idiom also shone forth. In fact, 6 other level-4 essays received 5 or more analytic scores of 4 and the remaining 3 analytic scores of 3, showing that graders saw them first worthy of the top holistic score, and then again worthy of high scores in the 4 global skill areas, upon a more deliberate analytic reading. And these 6 essays all gave evidence of control of CALP features.

Three of the 11 level-4 essays were strikingly distinct, however, in receiving a number of incongruent low analytic scores. These low scores indicated noticeable global weaknesses in language or rhetorical features. It was speculated that a self-assuredness and a creative flair for expressing ideas (in essay HIAR, which also showed some control of CALP) or experience in meeting educational or familial expectations (in essays KAOA and KKMA) convinced some graders and the chief of these learners’ college readiness. Capabilities communicated clearly to a grader thus appeared to tip the balance in the writer’s favor in the holistic scoring.

This study has illustrated that judgments of quality are very much influenced by the scoring instrument used. It seems at first counterintuitive that non-ESL graders would give 11 (18%) of the 60 papers the holistic score of 4 (using the Holistic scoring guide in Table 3-1), when ESL graders gave a holistic 4 to a mere 5 (8%) of the 60 (using the ESL
Holistic Scoring Guide in Appendix D). The scoring instrument the ESL graders used, however, required that a level-4 paper (given as a level-6) demonstrate correct usage of complex sentence structures, few grammar errors, and correct advanced (academic) vocabulary, in contrast to the more lenient demands in these areas specified by the instrument the non-ESL graders in my study had used for scoring holistically. It is thus fully understandable that different numbers of papers emerged as college-ready, given the divergent specifications of the 2 instruments.

Surprisingly, 2 essays that received holistic scores below 4 from English graders in this study looked like others that they rewarded with holistic 4s. For essay ENNT, 2 English graders attested to giving a holistic score below 4 primarily because of its inconclusiveness. The chief strongly argued that this was insufficient justification for withholding a holistic 4, given this examinee’s finesse with the idiom. Another essay, IIKI—which resembled essay ROYY, which received a holistic 4—was speculated to have been given a holistic score below 4 in part because the writer denigrated her English. Otherwise, these 2 essays had many similar features, excepting that IIKI had a couple more types and occurrences of errors. Of course, one would hope that if the writing were far from the target, it would not be given a high holistic score—college preparatory ESL teachers do not, as a rule, want inadequately prepared ESL learners to be let into college classes. They generally agree that it is unwise to forego additional preparatory work if it is needed.

The analytic scoring method used in this study—despite its setbacks—was useful in that it offered information that explained or repaired a few questionable holistic scores. Results showed that this method could be both stricter and more generous than holistic
scoring. Its slower reading gave graders more time to mark weaknesses in papers scored
4 holistically, with lower analytic scores. Yet it was also more generous in that it allowed
graders to give upper-half scores—3s and 4s—to the global features of focus,
development, and organization when those aspects were strong enough to override
weaknesses in clarity.
CHAPTER 5
DISCUSSION AND CONCLUSIONS
IMPLICATIONS AND RECOMMENDATIONS

Discussion and Conclusions

Perspectives of non-ESL English faculty graders on qualities in ESL timed writings have seldom been studied. But since non-ESL English teachers commonly score college-entry tests, it behooves teachers of college-bound ESL learners to attend to the expectations and standards that mainstream English teachers hold. Much pertinent research exists by now, whereby ESL teachers can gain valuable information about the unfavorable effects of the errors and mechanical weaknesses in ESL writings, and how to deal with such issues. And yet, L1 writing assessment experts have argued convincingly that more global feedback is a better means for more improved learner revisions (Elbow, 1995; White, 1995). And as second language writing scholars further, many ESL-type errors are actually “minor sentence-level errors which are a natural and inevitable aspect of learning a new language” (Sweedler-Brown, 1993, p. 4), and they should be seen as such, particularly in ESL writings that contain noticeable rhetorical control.

For all of these reasons, the chief purpose of the present research was to offer a new perspective on ESL writing, by first having college English professors holistically determine which of 60 essays were college-ready, and by then having them analytically determine the strong global qualities in all the essays. Previous research suggests that college tests reward demonstrations of control of CALP (Bachman and Cohen, 1998; Bardovi-Harlig and Bofman, 1989; Chafe and Danielewicz, 1987; Cummins, 1985;
Hamp-Lyons, 1991a). The notably high analytic scores that 8 of the 11 level-4 essays in my study received, plus the demonstrated control of CALP shown by the writers of these 8 essays, strongly support these claims.

Encouraging from a practical standpoint for teachers of ESL learners, and for ESL learners themselves, these non-ESL graders were able to discount reparable errors (those not associated with communication breakdown) and give 7 of the 11 essays that they scored 4 holistically analytic scores as high as 3 and 4 for clarity. These high scores reflect these graders’ views that the written communication in these 7 papers had few or no confusing parts, notwithstanding the existence of some ESL-type errors.

Of course, this finding must be interpreted within the context of the scoring instrument by which these English graders scored analytically. This device instructed them to be lenient on errors not associated with communication breakdown. And this result must be considered cautiously for other reasons. This group of English graders may have been exposed to a sufficient number of native speaker problems in timed writings so that their expectations for clarity did not presuppose error-free prose in a timed writing. Alternately, they may have scored clarity more leniently because these scores were merely hypothetical. Still, the English graders in this study were able to discount errors not causing communication breakdown. This finding lends support to the convincing argument forwarded by one empirical research study: As they are inevitable in the writing of students at various college levels, and even in the writings of post-college employees, errors should not command our attention (Haswell, 1988).

Significantly, then, the global focus of my scoring methods (the directive to discount errors not impeding comprehension) revealed that some graders will de-
emphasize less serious errors and reward a timed writing with a high holistic score, when other elements indicate college readiness. While development and clarity are generally considered by non-ESL graders to be weaknesses in ESL timed writings (Silva, 1993), it was encouraging to find that when an ESL writer could produce a timed writing resembling an entry-level freshman composition draft, it was scored college-ready holistically, notwithstanding weaknesses, by a grader who was de-emphasizing reparable errors.

Furthermore, this dissertation adds to the body of empirical studies demonstrating the usefulness of holistic scoring for categorization purposes (Enginarlar, 1991). Holistic scoring results (along with analytic) have yielded useful categorizations of ESL composing at different writing proficiency levels. Holistic scores were trustworthy—as per the acceptable .75 interrater reliability that resulted—and formed the basis for dividing the papers into levels.

In support of the assertion that borderline, or midlevel, writings can be challenging to score holistically (Elbow, 1996b), some such essays in the present study resulted in holistic splits in which only 1 grader assigned a holistic 4. As examples, take essays GACS, LUHA, and PYMA. The chief rectified all 3 splits to below 4—one holistic 3 and 2 holistic 2s. And in the subsequent analytic scoring, 1 or both graders recognized rhetorical weaknesses, besides problems with clarity of expression, in each essay (scoring them low analytically on 2 or more of the 4 global skill areas). The split holistic scores, the low analytic scores, and the chief’s lowering of the discordant holistic 4 assignments, all give further evidence for previous research attesting to graders’ expectation for prose that conforms to features of correct academic written style (Halpern, 1984). That is, for
these 3 ESL essays, which contained a mix of strengths and weaknesses, at least 1 grader (of 2) recognized that the weaknesses were too serious to be discounted.

In fact, the emergence of 10 holistic splits in which only 1 grader assigned a holistic 4 serves to illustrate the inherent variability in ESL learner writing, which is probably partly responsible for grading differences. It also shows just how distinct different learner abilities in a range of elements of the skill of writing can be (Hamp-Lyons, 1996). Hamp-Lyons calls attention to the uneven development of fluency and accuracy. Varied proficiencies in vocabulary, syntactic control, and rhetorical structure and style are quite common. As she argues, such ESL variability contrasts with the relatively stable fluency and accuracy in L1 performances, and thus may result in lower holistic scores from graders in large-scale testing. That is, a rater may lose sight of the quality of ideas and arguments present in an ESL paper full of errors, especially if that rater is not trained to favor ideas and coherence over correctness.

In essence then, holistic and analytic scoring results for the midlevel ESL responses in this study demonstrated that analytic scoring can help inform us about the strengths and weaknesses in the more problematic ESL papers. Many SL learners are unable to display an above average language mastery under the constraints of a timed essay test—errors of nonnative-like English are commonly sprinkled throughout their responses. Thus, even for ESL teachers with training and experience in evaluating ESL writing, judging borderline pass / fail writings can be problematic (Ruetten, 1994). Yet the more deliberate reading of the analytic scoring can help graders to clearly identify global strengths and weaknesses, despite their being intertwined.
Implications and Recommendations

The implications and recommendations of this study can be fairly applied only to similar testing and teaching environments. The scoring guides, timed writings, and graders used by this research all pertain to an undergraduate institution of higher learning where teachers of college-bound ESL students work to equip these learners to produce passing-level, or better, writing for non-ESL college credit courses.

Implications for Pedagogy

It is impossible for those of us who teach ESL writers to mark every error our students make, even in second drafts, so what should we emphasize most? Teachers of college-bound ESL students need better methods for guiding learners to write responses under test conditions and within time constraints. Timed writings are the equivalent of a first draft. There is no chance for real, meaningful revision. And while we cannot change the faculty grading standards, we can do much to help students become better prepared to do their best. Revelations from this study in 3 areas—overall quality in the whole paper; focus, development, organization, and clarity of expression in college-ready essays; holistic and analytic scoring methods used in combination—have been informative. Insights suggest more effective pedagogical approaches to ESL writing than the traditional emphasis on errors. These areas will now be reviewed.

The whole paper and the global skill areas

As a group, the graders in this study consistently scored essays with only minor language problems, which were later shown to be convincingly strong in the 4 examined global features, a 4 holistically. That is, papers that earned holistic 4s from 2 graders generally also earned (in a subsequent scoring) several 4s analytically, and the rest 3s, demonstrating that they possessed noticeable global strengths—a descriptive and clear
These results suggest that teachers preparing ESL students for college-credit classes make efforts to help them with prewriting—jotting down a focus, a few main points, and a conclusion. In this way, they can ensure student essays will include a stated purpose, adequate development, and some closure. Also, teacher feedback about the weaknesses in clarity of expression that affect comprehensibility of parts or of the whole will be of much help. One useful practice is to request that a couple of non-ESL English teachers holistically score—even notate lightly—a class of advanced ESL practice timed-writing tests. This provides students invaluable “real” opinions about their readiness for mainstream college English.

To the graders in this scoring experiment, focus had to be clear, specific, on-topic, and presented early rather than later in the writing. Regarding development, these English scorers seemed inclined to prefer definitive statements—which signaled personal confidence—over ambivalent ones—which showed indecision or insecurity. The ESL graders at this college, in contrast, tend to be more willing to accept indecisiveness and diffidence in a writing, and writing that ends inconclusively, whenever the learner can demonstrate the ability to shine linguistically.

These English graders considered all of the following to give evidence of the global skill of organization: transitional expressions, logical paragraph indents, and ideas that followed reasonably from those that preceded. As for clarity, these English graders
demonstrated that they could relax somewhat about language problems, so long as the ideas expressed were comprehensible.

All of these results have implications for teachers preparing ESL students for timed college-entry tests and college-credit classes. An effective strategy for helping students gain competency in CALP—the correct use of academic diction, grammar, and sentencing, which these graders rewarded with top holistic and analytic scores—would combine classroom interactions about form-meaning relationships in shared readings. For instance, students could be asked to read aloud and then to explain key ideas by paraphrasing them, or by giving a different word arrangement (forcing word ending changes or tensing of new verbs). Using synonyms, rewordings, and rearrangements, the meanings of verb forms, tenses and polysyllabic words could be made clear, with students actively participating. Vocabulary work focused on word parts—prefixes, roots, and suffixes—and diligent work with verb forms and tenses are both essential in helping students make form-meaning connections, which improve their CALP.

Students must acquire the ability to draft under time pressure, through repeated practice. They need to learn how to give definitive and clear statements of focus near the start of the essay. And they need to learn to apply principles of organization and development, to write using comprehensible and appropriate vocabulary, grammar, and syntax, and to give a conclusion to close. Finally, they need to be able to independently revise and edit their writing, for compliance with grader expectations in these features.

As for revising, ESL learners need to be able to repair as many “basic” errors—prepositions, articles, nominal plurals, and adjective endings—as possible. When some English graders encounter what seems to be an excessive number of sentence-level
errors—inadequate fluency—in nonnative-student writing, they discount rhetorical strengths (Hamp-Lyons, 1996; Sweedler-Brown, 1993). Teachers of college-bound ESL learners who give ample practice in timed writings can help students work through issues to achieve improved future drafts, through guided revising.

Using the holistic and analytic scoring methods together

A teacher’s providing students principled feedback about the global strengths and weaknesses in their drafts would facilitate work on trouble spots, and at the same time reward their successes. Results of this study indicate that teachers of college-bound learners might use the holistic and analytic scoring methods in combination to better advise their students about which global areas most need improvement in second drafts. First, they could use the holistic method to provide an evaluation of the overall passing quality. Then they could use the more deliberate analytic method to give discrete global evaluations for the student’s performance on 4 or more global aspects.

Limitations

The 74 essays constituting the dataset—14 for training and 60 for scoring—are all timed responses written several years ago for entry and placement into the junior college where this scoring experiment took place. Because of this limited sample size and population, and because the holistic and analytic rankings resulting from this study were not actually used for placing students into freshman composition or ESL courses, resulting scores must be interpreted with caution. Other ESL timed writings may be different, and graders might have delivered lower analytic and holistic scores had these been tied to actual course placements.

The chief scorer and graders participated voluntarily for 4 hours on a Saturday. This appeared to induce some participants to limit their oral commentary both during the
2 training sessions and in the post-scoring discussion. Most probably, aside from their scores, 2 grader perspectives are represented primarily by researcher speculation (R2 and R6 contributed minimally to verbal discussions during the trainings and the post-scoring question-and-answer session).

The chief scorer and graders did not, and had not been asked to, give written and oral commentary on every high-scoring or low-scoring paper, although they scored all 60 by both methods. Because grader rationale was not always available, researcher speculation could not be avoided altogether for some noteworthy cases of high and low scores.

The non-ESL college English graders—except for the chief—are all connected to the same junior college. This shared teaching environment probably represents a limited grading mindset that resulted in scores, commentary, and interpretations not generalizable to other populations of such graders. From another standpoint, however, the homogeneity afforded by these same circumstances could well have served as an asset, creating a degree of uniformity that might have contributed in positive ways to reliable scoring results.

The chief scorer and graders were less comfortable with the analytic scoring method used than with the holistic, in part because of its novelty to them, in part due to unforeseen limitations arising during the analytic training (discussed in Chapter 3). Therefore, analytic results are sometimes speculative, and cannot necessarily be relied upon as conclusive.

A reported weakness in the analytic scoring method of this study was its failure to define terms. This left graders guessing at how to interpret them, and probably relying on
their own personal ideas, which might have varied. In particular, graders reported that
the global skill of “organization,” as well as the 2 expressions of frequency for qualifying
the lowest points on the 4-point scale—“some of the time,” for 2, and “none of the time,”
for 1—should have been defined up front, before commencing this scoring. This mention
indicates that analytic scoring variations probably occurred. While writing assessment
research has explored the benefits and drawbacks to various scoring methods, analytic
designs included, I am not familiar with particular studies that have explored such issues
as the ones raised here about the definitions of terms and qualifiers.

Recommendations

It has been widely held for some time now that to make an important decision
based on a single written response is unreasonable. And for some students, whose L1 is
not English, a timed writing test presents a particularly problematic challenge. Some cite
the difficulty of responding to an obscure or uninteresting topic, others, the lack of
practice writing under time pressure for this audience. English as a second language
examinees tend not to perform to their potential on this test. Yet notwithstanding
objections, writing tests prevail as viable and respected measures for testing readiness for
entry to college, entry to junior level studies, and exit from some programs. The holistic
scoring method, moreover, is widely used and respected for the cost-efficient and
accurate evaluations it yields, and for its capacity to classify papers into distinct
proficiency levels. Judgments made in professionally managed scorings are solid and can
be trusted, because of the checks and balances.

Lacking new methods and approaches, some ESL teachers continue to rely upon a
timeless, tiresome, ineffective method: marking or correcting grammar errors and wrong
words. Their perfunctory focus on language conventions leads learners to believe that
grammatically and mechanically correct language equates to good writing, which is not the full picture. While the 1980s and 90s witnessed increased teacher understanding about errors that appear linked to failures, this expanded awareness has done little to reduce failure rates.

Clearly, ESL teachers need more help in apprising themselves of writing test practices and procedures, and of the holistic and other more global scoring methods, so that they can provide their students with ample preparation and practice, without having to invest great amounts of their time. In face of rising failure rates on college-entry barrier writing tests, and in college coursework, ESL teachers need better methods of helping learners generate prose that non-ESL scorers expect.

Perhaps in some college preparatory teaching environments, as in the one of this study, ESL teachers are doing a disservice to some of the more fluent advanced ESL students, by retaining them unnecessarily in ESL classes when English teachers would accept their composing. The case in point here was one fluent but somewhat rhetorically weak response—SALO (discussed in Chapter 4)—which highlighted a difference between the expectations of the ESL and English teachers at this college. Two English graders were more willing to reward his fluency with a holistic 4, while the original group of ESL placement graders (evaluating by the ESL Holistic Scoring Guide given in Table D-1) seemed to discount this same fluency, giving holistic scores below a 4, perhaps because the content was not adequately developed.

This case raises a basic question: Which determination would be best for the student, in the end—a fail, requiring further advanced ESL preparatory work, or a pass, allowing entry to freshman composition? That is, where do most such learners better
acquire composing skills and skills of argumentation? Further investigation of this issue is needed and could help us to determine whether students, whose papers were passed primarily on their strengths in fluency, more often passed or failed college writing classes thereafter. Obtaining an answer to that question could improve placement decision-making.

The results of this study reinforce the need to redirect our attention in the ESL writing classroom and in SL writing research, as writing and assessment experts have advised. As White (1994) succinctly puts it, teachers should center their commentary on the conception and organization of a writing, which will likely induce substantive improvements in the quality of the writing. Redirecting our attention away from student errors and onto the ideas would be a way of treating the learner in a more caring professional way, showing concern for the deeper issues of central ideas, main ideas, and meaningful support for these, which would probably tend to produce more meaningful student reflections and revisions (Elbow, 1973).

New evidence of what ESL learners can do well recognizes good overall writing and global strengths. Moreover, it suggests that teachers might respond to drafts more globally, instead of to errors (not shown useful in raising pass percentages). Particularly useful could be a new emphasis on helping students strengthen the ideas they wish to express. Findings of this research have confirmed that it is possible to acknowledge what constitutes a focused, organized, developed, comprehensible essay by staying attentive to global aspects.

The essays by ESL learners that non-ESL professors in this study rewarded with college-ready holistic rankings have underscored the power of 3 global strengths in timed
writings—focus, development, and organization—to help essays earn high holistic scores. Knowing more, we teachers who need to prepare ESL learners for timed writing tests of college-entry can emphasize these 3 global areas more in our own classrooms, as we continue to help students deal with the comprehensibility issues in their language.

Receiving such direction from us on global qualities in practice timed tests of writing, our learners might learn to compose and revise more broadly, and to better examine and improve the organization and content in their writing. Some might become more vested in expressing themselves fully and meaningfully. Writing and revising absolutely can provide a positive experience for both the learner and the teacher.

Reparable sentence-level errors, unless overly repetitive, can be de-emphasized. Pointedly complaining about errors is not useful, on the other hand, so teachers need to choose other methods. Nonetheless, serious errors—those confounding comprehension—in second drafts, and in all classroom exchanges must be cleared up. Learners need to be reminded, furthermore, that unless they repair serious errors, they will not likely pass college classes. Because some ESL teachers become adept at comprehending variably expressed SL writing and speaking containing non-nativelike inaccuracies, they must take care to ensure that their ability to comprehend obscure written expression does not transform into the assigning of passing scores to poor writings. The obvious risk is that students fossilize to the writing of incorrect forms or adopt the lackadaisical attitude that approximations are good enough, when they are not.

All said, most SLA experts agree that SLA processes, rate of SLA, and level of ultimate attainment all benefit from formal classroom instruction (Long, 1988). Moreover, formal instruction is now recognized as a means of facilitating SLA for the
learner by “speeding up the process of ‘natural’ acquisition” (Ellis, 1994, p. 654). It thus would help second language learners target language attainment in written English if their ESL teachers would observe some of the following principles and practices, among others:

- clear writing and revising directives
- checklists to delimit expectations and grades
- models of good writing (professional and student)
- teacher-student conferences about writing
- information about timed writing tests
- feedback from non-ESL teachers on practice timed-writings
- clear written feedback on strengths and improvements.

Many in the field of SLA agree: Under optimal instructional and environmental conditions, an adult can even approach native-like fluency in the target language. As writing, especially under test conditions, presents a great challenge to most learners, plenty of modeling, practice, guidance, and reinforcement are vital to real progress.
Discuss in detail three specific future goals that you now have. Why are they important to you and how will you achieve them?

Setting future goals is easy to do. The hardest is to actually go through all “adventures” that may come my way and achieve my goals.

First and most important goal for me is to finish at least a two year education at college in the field of Business. The world of Business is not something which describes me nor something which I aspire high to achieve. However, it is my father’s wish to see his son in the field of Business. That is why achieving this goal is so important for me. My father will see how greatly I respect him and how much I appreciate his financial support as well as his trust for me to let me travel thousands of miles away from home.

Secondly, I intend to keep photography as my hobby and develop my knowledge as far as I can. Photography has great importance in my life, that is why my goal is to not leave the camera out of my life. It helps me emotionally and it helps with the anxiety I have to face, the cruelty and the evilness of the world.

Achieving my first goal will help me achieve my third goal which I now have in mind. Finishing college, I could find a job with which I can support my self and I will be free to choose and set my own personal goals.

Financial independence is my key to my own success and to my freedom to set my own personal goals or dreams. With patience and persistanse I believe I will be able to achieve my goals.

Figure A-1. Essay CUKS, 267 words, a best-scoring essay, scored by R2 and R5
As each and every one of us wishes to fulfill one’s dreams and expectations, my desire to study and obtain proper certification isn’t very much different from any other. I would actually like to become a property lawyer and be able to serve local community with my knowledge and professional service. I happen to have a friendly and easy-going approach towards the others and think it should be considered an asset in today’s ‘law and legislation’ environment. To treat all people with equal dignity became the highest standard which is still, unfortunately, far from being practised. We all judge each other by the financial status, and that is what I find especially disturbing. The fact that the United States is a strong economical country makes us prone to such thinking and how inappropriate it seems. I know that through individual example much could be achieved. I am also pro preserving all good and moral values – not only through my private life, but the more in my professional practise. However, by realizing that today’s world brings constantly newer technologies and therefore modifies our social state, I remain open for learning all new ways in my occupational field. Still, a good memory and ability to connect facts are the very basics for every lawyer. I have not formed myself becoming a criminal lawyer due to the fact that legal procedures in such matters would take a pricy toll on my private life. I am in general very sensitive person and would not be able to forget drastic or harmful situations which criminal justice creates. This doesn’t though mean that I wouldn’t know the difference between what would be right or wrong in particular situation. Maybe because I have been taught some simple life lessons and as anyone else, I am also a product of healthy relations in my childhood household. I have been taught something what my parents and the rest of family applied – being truthful in all one does and giving it the best of one’s tries. Obviously, I do not know what the future will bring for me or for our nation, but all aforementioned should be sustained as everyone’s creed.

Figure A-2. Essay ZLDA, 361 words, a best-scoring essay, scored by R2 and R4
My ambition: what I want to do with my life.

As they say only God knows what is going to be with one’s life. Every person may have only ambitions, dreams and goals.

As for me personally, I have enough strong ambitions and goals, also do have opportunities for most of them to come true. And all what is needed now – is God’s help and time.

Right now when I am in the US, I do have to complete my MBA within the next year and a half or two. During this time I may also obtain one or maybe couple more specializations in your college. So what I am saying is when I complete the studies in the US, am planing to go back to Russian Federation as everyone knows the World’s Future is there. Within the next 10 years US is going to have economic stagnation and todays Russia has many opportunities to offer and the future will give even more.

Before coming to the US I was supposed to be an English Teacher in my Country. But I didn’t see my goals to be completed with only BA Degree and my profession as a Teacher. That is why I am here in the US – going to school now. / MBA Program and here full-time in order to obtain F-1 visa / .

When everything is completed and I am back to Russia I would probably start teaching in my University / Economics or Management in Business or it depends on what they would need. I would also complete the Doctors Degree. As for regular life: would probably start my own business after finish building my house. So as you can see everything is going to take lots of time.

All I want to be in this life is just a regular businessman with couple billion dollars. Also would like to do my hobby – that is teaching youngers how to live. I think have everything for doing my best to reach or get all that. I have good educational background now/ BA in Linguistics, Psychology, Diploma in Management, also Diploma in Journalism (BBC contest in Russia 2nd place winner 1996 / , Diploma in Technical Translation / Used to work in Grass Factory in Russia, interpreted also for SPARECS Concern representatives, this company unites 16 countries European and US / , And some other minor courses completed.

As for doing business have enough connections in Russian Federation, some other European and Asian Countries, such as Germany, Check Republic, G.B., China, Middle East Countries. US also by the way am doing my business here already. Run a Corporation / General One/. Just a small Business.

Do have also Political Connections starting from ex candidate for presidency in Russia, Presently the Governor of the major territory in Russia and finishing with City Majors, etc.

I do have also one big Dream: to build a Big University in Russia. So let’s see what is going to happen.

Figure A-3. Essay KNGK, 487 words, a best-scoring essay, scored by R2 and R5
One of my biggest dreams since I was in sixth grade has been to become a translator. I love languages. Therefore I think that would be my ideal job. I picture myself in the future as a person who is able to perform and excellent work as a translator mainly because I have a good ability to learn languages and I love learning them. I also think communication is one of the things we have to practice the most these days and working as a translator would contribute to make that happen. It would help us understand each other a little more. Besides, being a translator would allow me to meet all kinds of people from many different places. I find it very interesting and exciting to know a little more about the different cultures we have in the world.

I have worked very hard to be a translator and I think I will be able to make it.

Figure A-4. Essay ROYY, 159 words, a best-scoring essay, scored by R1 and R6

COMPUTER GRAPHIC ENGINEER.
THAT’S GOING TO BE MY JOB TITLE IN A FEW YEARS. THIS IS NOT THE HIGHEST PAYING JOB IN THE FINE ARTS FIELD, BUT IT’S THE ONE I’D HAVE THE MOST PLEASURE TO WORK WITH.

I’VE ALWAYS LIKED TO DRAW, SINCE MY CHILDHOOD. WHEREVER I GO, I’D TAKE MY SKETCH BOOK ALONG. I DRAW EVERY DAY OF MY LIFE.

AND I KNOW IT NOT ONLY TAKES DEDICATION TO BECOME A GOOD FINE ARTIST, BUT IT ALSO TAKES MONEY AND TIME. I DON’T HAVE EITHER ONE OF THEM NOW.

BUT I HAVE ANOTHER PASSION YET: COMPUTERS. FOR ABOUT 4 YEARS I’VE LEARNED FROM AND WORKED WITH COMPUTERS.

OF COURSE, THERE’S A GREAT VARIETY OF SPECIALIZATION COURSES ON THE COMPUTER AND FINE ARTS FIELDS, AND COMPUTER GRAPHICS IS THE ONE THAT SUITS ME BEST.

THE ONLY LIMIT IS THE IMAGINATION. THE IDEA OF CREATING SOMETHING NEW EVERY DAY AMAZES ME. MAINLY BECAUSE I CAN’T ADAPT TO AN EVERY DAY ROUTINE. I USUALLY GET BORED AND LOSE THE INTEREST.

FINALLY, I THINK THIS JOB WOULD BE INTERESTING BECAUSE I’D DO SOMETHING THAT I LIKE, WORK WITH COMPUTERS AND BE GIVEN THE CHANCE TO USE MY IMAGINATION TO CREATE.

Figure A-5. Essay SALO, 198 words, a best-scoring essay, scored by R1 and R4
My ambition: what I want to do with my life.

I would like to work in a communication company in the future, because I think it is a personal enrichment to have business relations with people since they learn you how the world works. As far as I am concerned, I am skilled in customer relationships thanks to the jobs I have already got. Indeed, I greeted people and gave them some pieces of advice in order to satisfy them with the best way possible. And this applies to the period when I was an insurance representative as well as the years I worked in a resort company. Therefore my experience made me discover my ability to communicate. In addition, it is all the more interesting for me to plan to work in a communication company that I could deal not only in a local area but also out of the United States. Actually I can speak several languages and I hope it will help me to work with international customers. But for now, it is just an aim, a challenge. I can still change my mind and anyway this is just an idea. The more important for me is to enjoy with what I do and it makes me happy to have a talk (even simple) with somebody.

Figure A-6. Essay PTGE, a 219-word second best-scoring essay, split by R3 and R6

My ambition: what I want to do with my life.

I have been always asking myself what I want to do with my life, whom I want to be and how I see my future. I have to admit that those questions are still very difficult for me to answer.

Sometimes I want my life to be a big adventure. I enjoy traveling, meeting new people, seeking new jobs and studying in new schools. This way, I feel like I learn a lot and have fun, too. There is always some interesting place I want to visit, but I definitely want to have my own house, somewhere in the world, to which I can always come back after my “journey”. I want my husband and son to be there for me. I love family life, so I will probably have another child in the future.

But, because the life is not only family and fun but also hard work I would love to finish college and find a good job. I think I would like to study Accounting. It is very interesting for me, especially because I know this field a little from my experience from my country. That would be exciting to learn how this “machine of business” works in the United States. After a couple of years in college I will try to find a job as a bookkeeper, learn as much as I can, get experience and maybe start my own business.

There are big plans. But for now, my goal is to learn English, start college and wait what future brings. That’s what I want to do with my life right now.

Figure A-7. Training 275-word essay GKUA, pre-scored holistic 4
In a world where IT (Information Technology) supersedes all sources of income, I feel uncertain as I jot down the job I would be doing in the near future. Private and Public relations management is a field where I believe I could use all my personal abilities to fulfill the requirements of the job. This job would heavily require patience, a lot of stress and a whole lot of communication – to name a few. But, what I feel is that if you like what you do then you are not only completing your duties successfully but you achieve self satisfaction. Self-satisfaction would go a long way when you choose to do a job for the rest of your life. And the most important thing is that you are developing as an individual. You learn not only to behave well with your customers but this sort of virtue is engraved in yourself. There are a lot of good things that come with this job, as in, a lot of people can be reached and they come to know you. And if you live up to the expectation of yourself first, any job would be as interesting.

Figure A-8. Training 195-word essay OIVM, pre-scored holistic 4
APPENDIX B
ESSAYS SCORING LOWER THAN A HOLISTIC 4

Table B-1. Forty-nine essays earning holistic scores of 3 / 2 / 1

Ten Essays Scored 3 / 3 Advanced ESL

<table>
<thead>
<tr>
<th>Essay code</th>
<th>1st gdr</th>
<th>scor</th>
<th>2nd gdr</th>
<th>scor</th>
<th>Retained holistic scor</th>
<th>C</th>
<th>R</th>
<th>Analytic Pairs</th>
<th>1st Grader Score focus</th>
<th>2nd Grader Score development</th>
<th>2nd Grader Score organization</th>
<th>clarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAMA</td>
<td>R1</td>
<td>3</td>
<td>R4</td>
<td>3</td>
<td>2 2 3</td>
<td>3</td>
<td>3</td>
<td>2 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSJR</td>
<td>R2</td>
<td>3</td>
<td>R4</td>
<td>3</td>
<td>4 3 3 4</td>
<td>4</td>
<td>3</td>
<td>3 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACS*</td>
<td>R1</td>
<td>3</td>
<td>R6</td>
<td>4</td>
<td>3 3 3 3</td>
<td>4</td>
<td>3</td>
<td>3 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IJKI</td>
<td>R3</td>
<td>3</td>
<td>R5</td>
<td>3</td>
<td>3 4 3 4</td>
<td>3</td>
<td>3</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JCDR*</td>
<td>R1</td>
<td>4</td>
<td>R6</td>
<td>3</td>
<td>3 3 4 4</td>
<td>4</td>
<td>1</td>
<td>4 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OAPA*</td>
<td>R2</td>
<td>3</td>
<td>R5</td>
<td>4</td>
<td>3 3 3 4</td>
<td>3</td>
<td>4</td>
<td>3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAAA*</td>
<td>R3</td>
<td>3</td>
<td>R5</td>
<td>4</td>
<td>3 4 4 4</td>
<td>4</td>
<td>3</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEEA*</td>
<td>R3</td>
<td>3</td>
<td>R6</td>
<td>4</td>
<td>3 3 3 3</td>
<td>3</td>
<td>3</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEJO</td>
<td>R2</td>
<td>3</td>
<td>R4</td>
<td>3</td>
<td>3 4 4 3</td>
<td>3</td>
<td>4</td>
<td>4 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIMN</td>
<td>R2</td>
<td>3</td>
<td>R4</td>
<td>3</td>
<td>3 4 3 4</td>
<td>2</td>
<td>4</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Twenty Essays Scored 3 / 2 Advanced / Intermediate ESL

<table>
<thead>
<tr>
<th>Essay code</th>
<th>1st gdr</th>
<th>scor</th>
<th>2nd gdr</th>
<th>scor</th>
<th>Retained holistic scor</th>
<th>C</th>
<th>R</th>
<th>Analytic Pairs</th>
<th>1st Grader Score focus</th>
<th>2nd Grader Score development</th>
<th>2nd Grader Score organization</th>
<th>clarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIAH</td>
<td>R1</td>
<td>3</td>
<td>R4</td>
<td>2</td>
<td>4 3 3 3</td>
<td>2</td>
<td>3</td>
<td>2 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANAA</td>
<td>R1</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>3 2 3 2</td>
<td>2</td>
<td>1</td>
<td>1 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANCA</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>2 3 3 2</td>
<td>3</td>
<td>2</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AODS</td>
<td>R1</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>3 3 3 2</td>
<td>3</td>
<td>2</td>
<td>2 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENNT</td>
<td>R2</td>
<td>3</td>
<td>R6</td>
<td>4</td>
<td>4 2 2 3</td>
<td>3</td>
<td>2</td>
<td>4 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIER</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>2 3 3 1</td>
<td>3</td>
<td>1</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GKEA</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>3 4 2 4</td>
<td>3</td>
<td>3</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOMO</td>
<td>R2</td>
<td>2</td>
<td>R4</td>
<td>3</td>
<td>2 4 1 4</td>
<td>2</td>
<td>1</td>
<td>1 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JNCA</td>
<td>R3</td>
<td>2</td>
<td>R5</td>
<td>3</td>
<td>2 3 3 3</td>
<td>3</td>
<td>3</td>
<td>5 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCAA</td>
<td>R2</td>
<td>2</td>
<td>R5</td>
<td>3</td>
<td>1 4 2 3</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIFA</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>2 4 2 3</td>
<td>1</td>
<td>4</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMMM*</td>
<td>R3</td>
<td>1</td>
<td>R5</td>
<td>3</td>
<td>3 2 3 3</td>
<td>1</td>
<td>4</td>
<td>1 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SADA</td>
<td>R2</td>
<td>3</td>
<td>R5</td>
<td>2</td>
<td>2 4 1 2</td>
<td>1</td>
<td>2</td>
<td>1 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDHA*</td>
<td>R3</td>
<td>3</td>
<td>R5</td>
<td>2</td>
<td>2 3 3 2</td>
<td>2</td>
<td>2</td>
<td>2 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SZNE*</td>
<td>R2</td>
<td>4</td>
<td>R4</td>
<td>3</td>
<td>3 2 4 3</td>
<td>2</td>
<td>2</td>
<td>2 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNGA*</td>
<td>R2</td>
<td>3</td>
<td>R6</td>
<td>4</td>
<td>3 3 3 4</td>
<td>4</td>
<td>3</td>
<td>3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSBO</td>
<td>R2</td>
<td>2</td>
<td>R5</td>
<td>3</td>
<td>2 4 1 4</td>
<td>1</td>
<td>3</td>
<td>2 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VCNA</td>
<td>R3</td>
<td>2</td>
<td>R5</td>
<td>3</td>
<td>3 4 2 3</td>
<td>3</td>
<td>3</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAJN</td>
<td>R2</td>
<td>2</td>
<td>R5</td>
<td>3</td>
<td>2 4 1 4</td>
<td>1</td>
<td>4</td>
<td>3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YESN</td>
<td>R1</td>
<td>2</td>
<td>R6</td>
<td>3</td>
<td>3 2 4 1</td>
<td>3</td>
<td>2</td>
<td>2 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Eight splits on college readiness—each constituted by a 4 with a 3
Table B-1. Continued

<table>
<thead>
<tr>
<th>Essay code</th>
<th>1st gdr</th>
<th>1st scor</th>
<th>2nd gdr</th>
<th>2nd scor</th>
<th>Retained holistic scor</th>
<th>C</th>
<th>R</th>
<th>Analytic Pairs</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASSA</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>BYJE</td>
<td>R1</td>
<td>2</td>
<td>R4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>GAAA</td>
<td>R1</td>
<td>2</td>
<td>R4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GZCN</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>HLMA</td>
<td>R1</td>
<td>2</td>
<td>R6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>HZJE</td>
<td>R2</td>
<td>2</td>
<td>R4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>LUHA*</td>
<td>R1</td>
<td>2</td>
<td>R6</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PYMA*</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ROPT*</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>SOAI</td>
<td>R3</td>
<td>2</td>
<td>R5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>YAPA</td>
<td>R3</td>
<td>2</td>
<td>R5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Seven Essays Scored 2 / 1 Intermediate / Beginning ESL

<table>
<thead>
<tr>
<th>Essay code</th>
<th>1st gdr</th>
<th>1st scor</th>
<th>2nd gdr</th>
<th>2nd scor</th>
<th>Retained holistic scor</th>
<th>C</th>
<th>R</th>
<th>Analytic Pairs</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>BACA</td>
<td>R1</td>
<td>2</td>
<td>R4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>BHCA</td>
<td>R3</td>
<td>1</td>
<td>R5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>EONA</td>
<td>R1</td>
<td>1</td>
<td>R4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>GZMA</td>
<td>R2</td>
<td>2</td>
<td>R4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>LUHY</td>
<td>R2</td>
<td>1</td>
<td>R5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>MATA</td>
<td>R1</td>
<td>1</td>
<td>R4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>PEMU</td>
<td>R2</td>
<td>2</td>
<td>R4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

One Essay Scored 1 / 1 Beginning ESL

<table>
<thead>
<tr>
<th>Essay code</th>
<th>1st gdr</th>
<th>1st scor</th>
<th>2nd gdr</th>
<th>2nd scor</th>
<th>Retained holistic scor</th>
<th>C</th>
<th>R</th>
<th>Analytic Pairs</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
<th>1st Grader Score</th>
<th>2nd Grader Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOLO</td>
<td>R1</td>
<td>1</td>
<td>R4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

* Three splits on college readiness—each constituted by a 4 with a 2
## Table C-1. Method for identifying and resolving discrepant holistic score-pairs

<table>
<thead>
<tr>
<th>Two Readers’ scores split&lt;sup&gt;a&lt;/sup&gt;</th>
<th>First Refereed score by the chief&lt;sup&gt;b&lt;/sup&gt;</th>
<th>First Refereed score and less discrepant score&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Second Refereed score by another reader&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Second Refereed score and less discrepant score&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 with 3 college-ready advanced ESL split</td>
<td>4</td>
<td>4 with 4</td>
<td>4</td>
<td>4 with 4</td>
</tr>
<tr>
<td>3</td>
<td>3 with 1 (STILL A SPLIT)</td>
<td>3 with 1</td>
<td>3</td>
<td>3 with 3</td>
</tr>
<tr>
<td>2</td>
<td>2 with 2</td>
<td>this refereed pair needs further refereeing</td>
<td>2</td>
<td>2 with 1</td>
</tr>
<tr>
<td>4 with 2 college-ready intermediate ESL split</td>
<td>1</td>
<td>2 with 1</td>
<td>1</td>
<td>1 with 1</td>
</tr>
<tr>
<td>4 with 1 college-ready beginning ESL split</td>
<td>3</td>
<td>3 with 1 (STILL A SPLIT)</td>
<td>3</td>
<td>3 with 3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>this refereed pair needs further refereeing</td>
<td>2</td>
<td>2 with 1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1 with 1</td>
<td>1</td>
<td>1 with 1</td>
</tr>
<tr>
<td>3 with 1 advanced ESL beginning ESL split</td>
<td>2</td>
<td>2 with 1</td>
<td>4</td>
<td>4 with 4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1 with 1</td>
<td>3</td>
<td>3 with 3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4 with 3 (STILL A SPLIT)</td>
<td>2</td>
<td>2 with 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>this refereed pair needs further refereeing</td>
<td>1</td>
<td>1 with 1</td>
</tr>
</tbody>
</table>

<sup>a</sup> defined splits; <sup>b</sup> possible 1<sup>st</sup> replacements; <sup>c</sup> 1<sup>st</sup> replacement scores; <sup>d</sup> possible 2<sup>nd</sup> replacements; <sup>e</sup> 2<sup>nd</sup> replacement scores
APPENDIX D
ESL HOLISTIC SCORING GUIDE

Table D-1. Descriptions of the placement and exit paragraph ratings

<table>
<thead>
<tr>
<th>Score of 6</th>
<th>The student displays the ability to satisfy every feature in the description for the Score of 5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score of 5—includes 5↑ 5↓</td>
<td>The paragraph has a topic sentence which contains a controlling idea. Major and minor supports are obvious and logical through clear transitional markers and are well developed with concrete details and specific examples. The writer successfully uses complex sentence structures and sentence combining. The paragraph contains few or no grammatical errors. Use of punctuation is varied and accurate. The writer uses advanced vocabulary with minimal word choice and spelling errors.</td>
</tr>
<tr>
<td>Score of 4—includes 4↑ 4↓</td>
<td>The paragraph has a topic sentence which contains a controlling idea. Main support is obvious and logical through clear transitional markers though minor support may need expansion. The writer is usually successful in using complex sentence structures and sentence combining. The paragraph may contain some grammatical errors which are not pervasive and do not interfere with readability. Spelling errors are restricted to advanced or technical vocabulary with occasional word choice errors.</td>
</tr>
<tr>
<td>Score of 3—includes 3↑ 3↓</td>
<td>The paragraph attempts to use a topic sentence, but may have no controlling idea. Support is evident, but may consist mainly of a series of generalizations or related personal experiences, often reflecting rhetorical patterns of the writer’s native language. The writer attempts complex sentence structures and sentence combining, though not always successfully. The paragraph may contain numerous grammatical errors, but these errors generally do not interfere with readability. The writer shows minimal competency in spelling and a knowledge of basic punctuation rules. Vocabulary is basic with frequent word choice errors.</td>
</tr>
<tr>
<td>Score of 2—includes 2↑ 2↓</td>
<td>The paragraph often lacks a topic sentence. It may, however, contain a restatement of the writing prompt. Support may be given, but in list form or rambling order. It contains serious grammatical errors which often impede the reader’s ability to understand the meaning. The sentence structure of a 2↑ / 2 / 2↓ paragraph is normally simple, with attempts at more complex structures often unsuccessful. The paragraph may reveal some basic familiarity with punctuation and spelling rules. Vocabulary is limited and basic.</td>
</tr>
</tbody>
</table>

1 Before 2000, ESL graders at this college placed ESL applicants by using this guide. A 6 corresponded with Comp I; a 5, Advanced II ESL; a 4, Advanced I; a 3, Intermediate; a 2, Basic II; and a 1, Basic I.
Table D-1. Continued

<table>
<thead>
<tr>
<th>Score of 1—includes 1↑ 1 1↓</th>
</tr>
</thead>
<tbody>
<tr>
<td>The paragraph usually lacks a topic sentence. It may, however, contain a restatement of the writing prompt. It will often contain strings of words or phrases, fused or run-together sentences and comma splices rather than complete sentences. When complete sentences are used, they are simple in nature. Word order, usage, and sentence structure are generally inconsistent. There is little or no support for the topic, resulting in a very brief response. Phonetic, rather than English, spelling may be exhibited, and even basic vocabulary may be misused. A 1↑ / 1 / 1↓ paragraph is also one which is written off the topic.</td>
</tr>
</tbody>
</table>

---

2 From archived files, by a committee of ESL teachers responsible for the evaluating of ESL writing tests for course placement. Copyright 1999 by St. Petersburg Junior College. Reprinted with permission.
## Table E-1. Fifteen splits—and their rectified score-pairs

<table>
<thead>
<tr>
<th>Chief's resolution code and gender</th>
<th>1st grader</th>
<th>Holistic score</th>
<th>2nd grader</th>
<th>Holistic score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIAR M</td>
<td>R3</td>
<td>3</td>
<td>R6</td>
<td>4</td>
</tr>
<tr>
<td>HOSE F</td>
<td>R1</td>
<td>4</td>
<td>R6</td>
<td>3</td>
</tr>
<tr>
<td>PTGE F</td>
<td>R3</td>
<td>3</td>
<td>R6</td>
<td>4</td>
</tr>
<tr>
<td>Resolved to 3 / 3 or 3 / 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JCDR M</td>
<td>R3</td>
<td>3</td>
<td>R5</td>
<td>4</td>
</tr>
<tr>
<td>SAAA F</td>
<td>R3</td>
<td>3</td>
<td>R5</td>
<td>4</td>
</tr>
<tr>
<td>SDHA F</td>
<td>R3</td>
<td>3</td>
<td>R5</td>
<td>4</td>
</tr>
<tr>
<td>SEEA F</td>
<td>R3</td>
<td>3</td>
<td>R6</td>
<td>4</td>
</tr>
<tr>
<td>SZNE F</td>
<td>R2</td>
<td>4</td>
<td>R4</td>
<td>3</td>
</tr>
<tr>
<td>TNGA F</td>
<td>R2</td>
<td>3</td>
<td>R5</td>
<td>4</td>
</tr>
<tr>
<td>MMMM M</td>
<td>R3</td>
<td>1</td>
<td>R5</td>
<td>3</td>
</tr>
<tr>
<td>Resolved to 2 / 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROPT M</td>
<td>R3</td>
<td>2</td>
<td>R6</td>
<td>4</td>
</tr>
<tr>
<td>Resolved to 4 / 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40% (11) of the 29 essays (with pre-scores corresponding to 3*) received split-pair scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolved to 3 / 3 or 3 / 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13% (2) of the 16 essays (with a pre-score corresponding to 2*) received split-pair scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolved to 2 / 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22% (2) of the 9 essays (with pre-scores corresponding to 1*) received split-pair scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolved to 2 / 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ESL raters determined placements on the basis of rhetorical features, grammaticality, and academic word choice and usage—3*=advanced ESL; 2*=intermediate ESL; 1*=beginning ESL.
REFERENCES


Spaulding, Cheryl. (1992). The motivation to read and write. In Judith W. Irwin & Mary Anne Doyle (Eds.), *Reading/writing connections: Learning from research* (pp. 177-201). Newark, Delaware: International Reading Association.


Swain, Merrill. (1998). Focus on form through conscious reflection. In Catherine Doughty & Jessica Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge UP.


BIOGRAPHICAL SKETCH

Cultivating her interest in linguistics, Jeanna Ojeda moved to France to study French and then to Costa Rica to study Spanish. She acquired her M.A. in Applied Linguistics from the University of South Florida in 1991. Thereafter, she taught all levels and skill areas of English as a second language at St. Petersburg Junior College in Clearwater, Florida. While in residence in Gainesville, she taught Scholarly Writing in the Linguistics Program at the University of Florida and various ESL courses at the nearby English Language Institute.

Since July of 1999, she has been living in Pinellas County, where she implemented the newest ESL Program at the Tarpon Springs Center of St. Petersburg College (formerly St. Petersburg Junior College), and has been teaching various ESL courses since.