NATE AND THE ONE-SIXTY-EIGHT

A DUAL ANALYSIS
OF TESTING CHARACTERISTICS AND ENROLLMENT TRENDS
AMONG STUDENTS WHO FAILED THE LEVEL 4 TEST
AT THE AMERICAN UNIVERSITY ALUMNI (AUA)
LANGUAGE CENTER, BANGKOK

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE MASTER OF ARTS IN TEACHING
DEGREE AT THE SCHOOL FOR INTERNATIONAL TRAINING
BRATTLEBORO, VERMONT

BY
MICHAEL S. MCMILLAN

MARCH 2008
IPP ADVISOR: MARTI ANDERSON
The author hereby does grant the School for International Training the permission to electronically reproduce and transmit this document to the students, alumni, staff, and faculty of the World Learning Community.

© Michael S. McMillan, 2008. All rights reserved.
This project by Michael S. McMillan is accepted in its present form.

Date: March 28, 2008

Project Advisor: Marti Anderson

Project Reader: Tana Ebaugh

Acknowledgements: First, I would like to thank my brother, Jon, to whom I am deeply indebted for introducing me to the concept of writing as a logical, disciplined, and manageable process. At a critical moment in my life, he sat me down and systematically showed me how to write my undergraduate thesis. Without his supportive intervention, I know I would not have earned my bachelors degree – perhaps to this day. I would also like to thank Debra Allee, my professor at Columbia University and employer for three years, who demonstrated the amazing power of the editor’s pen in her review of environmental impact statements that I helped to write.

There are more recent players in my life who have helped make my introduction into the world of education a path of adventurous personal discovery. The SMAT faculty at SIT, and all those who helped define the SIT experience in years past, have carved a niche in my heart that will be the source of deep personal inspiration in years to come. And for getting through and even enjoying the experiences of working at AUA, not to mention outstanding brainstorming sessions, Boojie Cowell and Tana Ebaugh, two friends and comrades for life. My advisor, Marti Anderson, has also framed the entire IPP experience with wise advice, thoughtful trust, and patient support. Her flexibility is much appreciated. Dee Parker, fellow supervisors, and teachers at AUA have all offered unending support, extending into my prolonged absences over two years to study in the US. Finally, a very special thanks to my partner, Alek, who showed extraordinary patience and support as I worked through this project.

To all, a heartfelt thanks.
ABSTRACT

This paper explores the fate of 168 students who failed the Level 4 test at the American University Alumni (AUA) Language Center in Bangkok, Thailand. It examines enrollment trends of all Level 4 students during a one-year study period spanning 2006 and 2007, and reveals what happened to the students – who dropped out (and when), and who continued to study. It also examines why students might have dropped out. It then takes a critical look at certain aspects of the Level 4 test, explores what makes a test a “good” test, and recommends a series of improvements to AUA’s test, as well as suggestions about broadening the perspectives of management. Together, these might result in better student performance, higher retention and enrollment rates, and a healthier, more productive work environment for teachers, administrators and staff.
Educational Resources Information Center (ERIC) Descriptors

Assessment
Alternatives to Standardized Testing
Testing Methods
Educational Administration
Grading
Listening Comprehension Tests
Scoring
Test Results
# TABLE OF CONTENTS

Chapter

1. BACKGROUND  
   - About AUA 2  
   - The Focus of this Project 12  
   - An Early Inquiry: How This Came to Be 12

2. THE ANALYSIS  
   - Introduction 16  
   - The Evolution of My Methodology 17  
   - Resources 19  
   - What the Study Reveals 21  
     - Part 1 22  
     - Part 2 37

3. LITERATURE REVIEW  
   - Introduction 41  
   - What I Found 42  
   - The Literature 44

4. RECOMMENDATIONS FOR AUA  
   - Introduction 53  
   - Recommendations 59

5. APPENDIX  

6. BIBLIOGRAPHY  

vi
CHAPTER 1

BACKGROUND

This project explores the fate of Nate, a student at the American University Alumni Language Center (AUA) in Bangkok, Thailand, as he began his journey into a program of learning to speak English. He was among a much larger group of 1,288 Level 4 students who studied in various terms spanning 2006 and 2007. Nate and 168 of his classmates had the misfortune of failing their first formal exam at AUA. They had recently embarked on a fifteen-level, 450-hour program and were being tested at the conclusion of Level 4. Despite a remedial plan designed to assist students who fail the Level 4 exam, only a small minority continued their studies to later pass the exam and proceed to higher levels. Indeed, very few will survive to complete Level 15 and receive their Certificate of Achievement.

Nate, who is 23, is from Nakhon Si Thammarat, a provincial capital in southern Thailand. He grew up in a ramshackle house behind one of the city’s hectic marketplaces, accessible only by crossing a low swamp on a series of heavy, unstable planks. He’s the youngest son in a large Muslim family, and he moved out of the house several years ago. For a few years, he lived in the back of a restaurant where he worked, driving his boss, Somkit, each morning to the market near his old home, where they bought vegetables for the restaurant.
Two years ago, Nate moved to Bangkok, where he could make more money. It wasn’t long before he realized the advantages of studying English, and AUA was a logical choice since its tuition is quite low. Nate is from a conservative background, and as a Muslim, he is a minority in Bangkok. He doesn’t speak a lot in class, keeping mostly to himself.

What experiences did Nate have at AUA? Is he still studying here after failing the Level 4 test? What about the other 168 classmates who didn’t pass the test? Did the system offer them support to get through this hurdle? These are the questions I will explore and attempt to answer.

ABOUT AUA

General Description and History.
AUA is no less than an “institution” in Thailand, serving for decades as the Kingdom’s primary source of English language instruction. About one-half million people have studied at AUA since it opened in 1952, and many more have attended cultural and entertainment events on its main campus in central Bangkok. There are now five other branches in Bangkok, as well as fourteen provincial branches. The main campus is called “Rajdamri.”
The AUA Language Center was established in 1952 to promote further mutual understanding between the people of Thailand and the United States through instruction in their respective languages, customs and traditions. AUA is a non-profit, non-political and non-religious organization. The affairs of the language center are governed by a board of Thai and American members of the American University Alumni Association (AUAA). AUAA originated in 1924 as a social organization for Thai graduates of universities in the US and their counterparts in Thailand.

Courses of study.

Three languages are taught at AUA: Thai, Japanese and English. The Thai and Japanese programs are quite small and administered separately from the English program; they are aimed at ex-pats living in Thailand. The English program, however, attracts several thousand students at any given time – mostly, but not exclusively, Thai. At Rajdamri, there are anywhere from approximately 2,000 to 4,000 students of English each term, depending on the time of year.

The English program operates on a cycle of seven terms per year, separated by short interims. Each term is six weeks in length. There are three (3) distinct curricula within the English program: the Regular Program, Special Courses, and the Academic English Program (AEP). Students in the Regular Program and Special Courses typically study an average of five hours per week, for a total of 30 hours each term. Students in AEP study an average of four hours per day, for a total of 120 hours each term.
The Regular Program. The Regular Program is where the vast majority of AUA students study. As stipulated by the Ministry of Education (MOE), the Regular Program consists of 15 levels; a placement test is used to position students in the most appropriate level. Thus, a student placed into Level 1 and continuing through the entire program would typically study five hours per week for six weeks, in fifteen terms over a period of just over two years. (His/her class and all other individual classes shall be referred to as “sections.”) However, if the student should fail any of the level exams that are held at Levels 4, 8 and 12, or the exit exam at Level 15, the story is likely to be quite different, as we shall see.

The Regular Program was expanded by two levels several years ago at Rajdamri and certain other branches in order to provide additional support for students placing at the lowest level. However, since AUA must adhere to the MOE stipulation of offering only fifteen levels, the two new levels were designated as “Pre-Levels,” thus keeping at least the appearance of the original 15-level plan intact. This addition expanded the program to 510 hours of study over seventeen terms. Those who complete the entire program are still considered “Level 15 Graduates,” whether or not they studied the two “Pre-Levels.”

Special Courses and the Academic English Program. Special Courses are designed for students who wish to focus their studies on a particular skill or group of skills, or who have completed Level 15 and want to continue their studies. Courses are grouped into four main categories: writing, speaking and listening, exam skills, and business English.
The Academic English Program is an intensive program designed for students preparing to study abroad in a university setting.

My role at AUA.

I am a supervisor in the academic office at AUA, where I oversee Special Courses and the busy evening shift. For the first five weeks of each term I have infrequent contact with students. In Week 6, however, many of those who have failed their level test come to the academic office for assistance. During the last few days of each term I spend the majority of my time counseling these students. I describe this in more detail below.

Assessment at AUA

In the Regular Program, grades are given for classroom performance and for level tests, as described below.

Classroom Performance. There are three passing grades and two failing grades, as follows:

<table>
<thead>
<tr>
<th>Passing Grades</th>
<th>Failing Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA = above average</td>
<td>RS = repeat level for poor scholastic performance</td>
</tr>
<tr>
<td>A = average</td>
<td>RA = repeat level for poor attendance</td>
</tr>
<tr>
<td>BA = below average</td>
<td></td>
</tr>
</tbody>
</table>

If a student receives a BA for two consecutive terms, he/she must repeat the level.

Level Tests.

Passing Grade = “P”  Failing Grade = “NP”
The Level Tests

In order to progress through the Regular Program, students must not only receive a passing grade at each level, but must also pass tests at the end of Levels 4, 8, 12 and 15, corresponding with the completion of each book used in the curriculum. Assessment is designed to become progressively more individualized as the student rises through the level system, culminating in a Level 15 test that includes an interview and at least two readers for each student’s writing sample.

The Level 4 Test. At Level 4, students are given a multiple-choice test of 30 questions in the following format:

1. Listening 9 questions (with 2 examples)
2. Grammar 8 questions
3. Correct Response 9 questions
4. Reading 4 questions

Writing is not tested at Level 4. Neither is speaking. However, in lieu of testing speaking, the teacher assesses each student’s oral fluency on a scale of 1 to 4, based on the student’s performance over the course of the term. This “Oral Fluency” (OF) score is basically a backup score that is used only if a student’s score on the Level 4 test is at the borderline of pass/fail. In such cases, the scorer refers to the OF, and adds or subtracts points from the total score, as follows:

<table>
<thead>
<tr>
<th>Oral Fluency Score</th>
<th>Adjustment to Final Score of Paper Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>+2 points</td>
</tr>
<tr>
<td>3</td>
<td>+ 1 point</td>
</tr>
<tr>
<td>2</td>
<td>no change</td>
</tr>
<tr>
<td>1</td>
<td>- 1 point</td>
</tr>
</tbody>
</table>
For example, the pass/fail threshold for Form A is 16, and if a student with an OF of 4 scores 14, he or she will have gained two extra points on the test and pass it with a final score of 16. On the other hand, a student with an OF score of 1 who scores 16 on the test will lose one point and fail the test with a final score of 15.

The remedial plan for students who fail the Level 4 test.

Students who fail the Level 4 test must return to Level 2. Upon their completion of Level 2 they may take the Level 4 retest. If they pass the test at that time, they proceed directly to Level 5. If they do not pass the test, then their immediate fate is determined by their performance in Level 2; a passing grade for Level 2 (for their classroom performance, as opposed to the retest score) advances them to Level 3, whereas a failing grade would require that they repeat Level 2 again. The next term, upon completing their assigned Level, students may once again sit for the Level 4 test. This process continues, theoretically, until the student passes the test and advances to Level 5.

When a student fails the Level 4 test, registers for Level 2 and passes the retest according to plan, the system tends to work. However, as we shall later see, many students don’t pass the retest on the first attempt and are vulnerable to getting discouraged by having to repeat lower levels. This is when the policy becomes less clear, since procedures are somewhat cumbersome and not clearly documented or displayed. Further, the policy is often misinterpreted by teachers and support staff. It is now seen as “one of those grey areas,” which tend to emerge when a clear consensus has not been reached between academic supervisors, or between supervisors and the registration office. As a result,
what probably started out as “case by case” situations soon evolved into “it can’t be
done” situations.

For example, repeating students who pass Level 2 but do not pass the Level 4 retest
progress to Level 3. They are often incorrectly advised by members of the academic
office that they are not eligible to sit again for the test in Level 3, but must wait until they
have completed Level 4 for the second time. This is not true. A student who is repeating
Level 2, 3 or 4 may sit for the Level 4 test at any of those levels. This was true the last
time I asked, but may change again in the near future.

Distribution of Grades.

Students are informed of their grade on the last day of term. They are also expected to
register for the next term during their last class by obtaining their student card from their
teacher and going immediately to the registration office. This includes students who have
failed. I shall describe below the process for distributing grades at the testing level.

Prior to the final class. Students in testing levels can determine whether or not they
passed the test prior to the last class. On the last day of the term, test scores are posted on
a public bulletin board and students can review this list prior to going to class.

Confidentiality is maintained by using student i.d. numbers, rather than names. No
information is given about actual test scores; students who pass the test see only a “P”
and students who fail the test see “NP,” plus two codes that correspond to skill areas of
the test in which they performed poorly.
The codes and skills are as follows:

\[
\begin{align*}
L & = \text{Listening} \\
G & = \text{Grammar} \\
CR & = \text{Correct Response} \\
R & = \text{Reading}
\end{align*}
\]

The two skills are determined by the person scoring the test, who reviews the student’s answer sheet and selects two skills. Sometimes there are more than two areas of poor performance, and at other times errors are distributed evenly between the four skills. Still, only two skills are chosen.

**In the final class.** Each section teacher possesses a small card for every student in his/her section; the teacher uses these cards throughout the term to record attendance. Final grades and attendance are noted on the card, and the cards are given to the students in the last class. For students who fail the Level 4 test, the teacher is instructed to use a marker to marks the card with a large red “RS,” and on a line reading:

“Next Level ____________”

the teacher writes “TWO,” rather than “FIVE,” in red. Then, across the face of the card on the left side, the teacher uses a marker to write “RS” in red, block caps approximately 2 inches in height, indicating “Repeat Scholastic.” The required red marker is to facilitate processing in the registration office. Finally, students wishing to register for the upcoming term are directed, in class, to the registration office. This includes failing students who are to register for Level 2.
Support and guidance for students who fail the Level 4 test. Students who obtain their test results at the public bulletin board may or may not see a note at the bottom of the board informing them that they can speak to a supervisor in the academic office. Still, teachers believe that some of these students go home after seeing their test results and never return. Do they? That is one question this project will explore.

For students who fail the Level 4 test, guidance and counseling from the teacher is at the teacher’s discretion. With this discretion comes an unofficial flexibility in the grading system that may or may not be exploited by the teacher. For example, students whom the teacher believes could successfully complete the next level of study (but failed the test by a narrow margin) may be directed to me or another supervisor for an interview and possible adjustment of the test score to a passing mark, thus allowing for their progression to Level 5. Students who failed the test by a wider margin are generally not offered this option. Again, supportive intervention for these students, such as counseling them on areas of weakness, is at the discretion of the teacher. It should be noted that teachers have only five minutes between classes, so often there is virtually no time for a meaningful discussion between the teacher and students.

Whether or not a student who failed the test receives any support ultimately depends both on their teacher and the supervisor on duty in the academic office. Almost all teachers try to review test scores and determine which student(s) who failed might have a chance at Level 5. Those who don’t probably aren’t aware that they have this option, or don’t have the time, since test scores are given to teachers just one or two days before their final classes, and a supervisor is not always available when needed. However, certain
supervisors will look at a student’s failing score and, rather than conduct a brief interview, routinely advise the student that their only recourse is to register for Level 2.

In many respects, there is already a certain amount of flexibility in the assessment process, which is nice. For teachers who are both informed and fully motivated to push for their students and “work the system,” students are only at the mercy of the supervisor on duty, who may or may not decide to interview selected students.

**My role in providing support and guidance.** As I mentioned above, I am one of the supervisors who meets with students. If a student has failed the Level 4 test and I have not already heard the concerns of his/her teacher that the student is capable of studying successfully in Level 5, I am supposed to direct the student back to Level 2 as a matter of routine. However, if the teacher approaches me in advance and believes that the student is capable of studying in Level 5, I may conduct a quick interview to determine if the student should pass, and I will do so if at all possible.

This is when the “human” element enters, and thus subjectivity. For example, what I sometimes find is that I spend five minutes speaking with a student who failed by three or more points, but who shows good listening and speaking skills. If that student scored poorly in grammar, correct response, or reading, we can discuss the importance of working on these skills, basically by returning to Level 2. But if the student failed because of poor performance in the listening portion of the test, yet demonstrates good
listening skills with me, I begin to question other factors and wonder if perhaps the listening test is not accurately assessing our students’ listening skills.

The other very real “human” element is the process of advising some students that they must return to Level 2, which is a hard sell. I shall explore other options to this in a later chapter. My gut feeling has always been that such a measure fails outright in serving our students’ needs. In an era of increased competition and declining enrollment, it may need to be reconsidered.

THE FOCUS OF THIS PROJECT

This project attempts to explore some of the mysteries hidden in AUA’s testing statistics in order to reveal the “human face” of our students who fail. It looks at certain aspects of the test, the procedures in place to handle student failures, and most importantly, how those procedures actually impact these students. Finally, it asks and explores what AUA might do to better serve these students. It does not look at the other level tests, though experience shows that similar procedures and circumstances at higher levels appear to have similar results.

AN EARLIER INQUIRY: HOW THIS PROJECT CAME TO BE

I returned to work at AUA from my second summer at SIT in early September 2007. This coincided with the last week of Term 5, when level test results were being posted and students were notified of their grades. The failure rate among Level 15 students in Term 5 was interpreted by AUA’s testing supervisor as abnormally high, and he was
perplexed. Throughout that week my duties were to assist students who had failed the Level 15 test, in addition to those who had failed at other levels. The more I spoke to students, the more I learned about the individual circumstances affecting each student. My conclusion was that we don’t really know who our students are — even the comparatively few who make it all the way to Level 15 — and this has a direct impact on how we serve them. Nor do we have any idea about how our students interpret the Level 15 test.

Though my original intent for this project was to focus on both Level 15 and Level 4 tests and students, it became necessary to narrow my focus, as I describe in Chapter 3. It is worth noting, however, that my initial research that week in September 2007 included not only looking at histories of students who failed the Level 15 test, but also speaking to students about their experiences with the test. I learned, for example, that most students who failed the Level 15 test had given us no prior recorded indication of academic weakness. In fact, contrary to our assumptions, many of them had excellent academic and attendance records. I also discovered how several students felt about the reading portion of the Level 15 test. To paraphrase more than one student, they explained that they read in English every day at work and at home, but they had yet to encounter (and predicted that they never again would) a reading such as the one they found in the Level 15 test. If it is true that the students had not been exposed to similar types of readings either inside or outside the classroom, then both the content validity and construct validity of the Level 15 test is called into question. Alas, such as exploration must be held for another opportunity.
Finally, by examining every Level 4 answer sheet for Term 5, 2007, I learned that the majority of all Level 4 students were performing quite poorly during the first few moments of the test – the listening portion. In two cases, 100% of the students in a section answered question #1 incorrectly.

Where might these initial discoveries take me as I attempted to define my own IPP? One very real risk was that the path would be ever widening – that I would indeed find that the more I looked, the more I found. What might I discover? How could I define and clarify my intentions? Of course, this happened, and it took some time to develop a clear sense of purpose. There seemed to be opportunities for exploration everywhere I turned.

Ultimately, the students defined my purpose for me. Since I interview students at all testing levels who failed their test, I concluded that it was the students at Level 4 who were being let down the most. After all, there are far more Level 4 students lined up on the sofa outside my office than any other level! In addition, Level 15 students can see their goal within reach and have already demonstrated their motivation and perseverance. Level 4 students, on the other hand, are just getting started. They have just established friendships with their classmates. They are struggling with enormous amounts of grammar. Most likely, this is the first test they have ever failed (since it is accepted policy throughout the educational system in Thailand to pass all students to the next level). At AUA, students who fail the Level 4 test are being sent back two levels, an unprecedented loss of face in a culture where “face” defines social standing.
Having determined to focus only on Level 4, questions surfaced. Is there something about the test, or how it is administered, that prevents students from performing well in the listening portion? If so, how many students fail the test as a result? What can we do about this?

And of course, my original question: What really happened to Nate and those 168 students who failed the Level 4 test.
CHAPTER 2
THE ANALYSIS

INTRODUCTION

This chapter focuses on the actual results of my explorations into the Level 4 test. Specifically, it looks at three related things. First, after the Introduction it describes the evolution of my methodology. Next, it describes the most useful resources available to me at AUA, and identifies the reasonable parameters of the project. Finally and most importantly, it presents my actual discoveries. They reveal almost all that we can determine about what happened (or is happening) to Nate and his classmates who sat for but failed the Level 4 test during the period that was analyzed (hereinafter the “Research Period”), which was seven consecutive terms, Term 5, 2006 and Term 4, 2007, spanning the period from August 2006 through July 2007. Table 3.1 in the Appendix defines the exact Research Period.

I was fortunate that the first book I read in preparation for this project was James Dean Brown’s Understanding Research in Second Language Learning (1988, Cambridge: Cambridge University Press). This is an excellent resource for the teacher who wants to undertake a well-planned research project. After initially preparing to lay out a more formal process, I concluded that the goals of my project are more modest and that it is not an experimental statistical study to be designed, but is a look at statistics as they already
exist within AUA’s system. Still, I remained conscious of Brown’s basic characteristics of statistical research, as summarized below (Brown, 1988).

1. Systematic
2. Logical
3. Tangible
4. Replicable
5. Reductive

It has been my goal to keep these characteristics at the forefront of my planning, so as to present the most reliable and useful conclusions possible.

THE EVOLUTION OF MY METHODOLOGY

In Chapter 1, I described my original inquiry in September 2007, which looked at both Level 15 and Level 4 test results for the recently completed Term 5, 2007. My Level 15 inquiry focused on the personal histories of the students who failed the test, while the Level 4 inquiry looked at a broad trend in how students answered an individual question on the test. What evolved into this project is actually a combination of these different perspectives — an attempt to look at both the stories of our students and broader enrollment and dropout trends, using these statistics to build the stories. This sounds somewhat counterintuitive, since statistics by their nature appear impersonal. Surely, they reveal far less about our students than face-to-face interviews could. Yet, when properly examined, they speak, telling a surprising story that itself merits attention. My goal is to put a personal face to the statistics.
Finally, practical considerations prevented me from interviewing students for this project. Thus, the results of the project revolve around the data available to me as defined by a reasonable project scope.

**Parameters.** As the project took shape, more questions arose. For example, should I look to see whether or not students of a particular age were more likely to fail the test and/or subsequently drop out? Likewise, what about class times? Do students fail more in the evening, after a day of work or school? In mid-afternoon, when the urge for a nap strikes? And what of the teachers’ role? And absentee rates? Finally, is class size a factor? Are students in larger classes at a disadvantage?

I looked at these factors in my research and saw no immediately discernable trends. Thus, they regrettably fell by the wayside, again for reasons of scope and practicality.

**Key Factors to consider.** The first thing I had to do was define was a research period. I needed to choose a period that was both manageable and long enough to offer insight into trends. I chose the period as defined described above, which represents when the bulk of my research takes place. (See the note accompanying Table 3.1 in the Appendix for exceptions.) I also looked at a “follow up” period (Term 5, 2007 to the present), during which I continued tracking the progress (or lack thereof) of students who failed the exam, in order to more accurately determine their long-term fate at AUA.
However, a different, somewhat shorter period of time applied to my analysis of the Level 4 test listening portion. Since this relied solely on student answer sheets, which are routinely destroyed (see below), I had no choice but to study only what was available. Therefore, this portion of my studies is limited to Terms 5, 6 and 7 of 2007, a period corresponding roughly to August through December 2007.

RESOURCES

The most useful resources accessible to me were the AUA computerized registration system, and archived answer sheets, grade sheets and retest results.

The Registration System. The computerized registration system is surprisingly limited in terms of the information it provides and how a user interacts with and manipulates this information. Nonetheless, it is searchable by both student name and student i.d. number. In a cumbersome process by which the six-digit student i.d. number is entered into the system, it provides an “enrollment history” for each student. This history shows enrollment by term, section level (the name and level for the individual class), final grade, and the number of hours the student was absent. For testing levels, it also shows the test results as “P” (pass), “NP” (no pass), or “NS” (no show). For failures, two weak areas are noted out of the four areas tested (“G” = grammar, “R” = reading, “CR” = correct response, “L” = listening). Table 3.2 in the Appendix summarizes relevant information available from the registration system.
Caveats: The computer system is capable of showing the names of section teachers in a separate search by class lists, but this is unreliable. This is because teachers are assigned their own i.d. number in the computer system from a rather limited pool of numbers, so they are far from unique. As a result, almost all current teachers share their i.d. number with at least one former teacher. Thus, an examination of teachers for a student (by requesting individual class lists, term by term) often reveals the names of teachers who no longer teach at AUA. Further, student i.d. numbers are also recycled, though not as frequently as teacher i.d. numbers, leading to more mismatches in the system. Any analysis of students must be undertaken with care to maintain accuracy and to weed out mismatches. (An examination of the limitations and usage of the registration system are clearly beyond the scope of this project, but will be considered in Chapter 4.)

Archived Answer Sheets, Grade Sheets and Retest Results. Individual student answer sheets are kept on file for approximately four weeks after the tests are administered, when they are destroyed to make room for the upcoming term. During their brief life, the answer sheets are filed in a cabinet in my office (which I share with the testing supervisor), with each section clipped and collated by score, with low scores on top. This facilitates a quick review of answer sheets when students come to the academic office in the final week of each term with inquiries, since most inquiries are from students who just failed the test. The brief shelf life of the answer sheets is not an accurate reflection of their potential value, as we shall later explore. Final scores for each student are recorded on separate grade sheets, as described below.
Grade Sheets are modified class lists that show only the grade given to each student and the number of hours absent. They are archived indefinitely in large binders. Once archived, they serve no known function in the monitoring and administering of the testing program.

Retest Student Test Results are forms that are manually generated by the testing secretary each term and archived indefinitely. They summarize the results of students who have retaken any level test due to previous failure(s). They show the grade (NP or P), and two weak areas in cases of a failing grade (as described above and in more detail in Chapter 1). They also show the current section level for each student. So, for example, if a student who earlier failed the Level 4 test is retaking it from Level 2, the Level 2 section is indicated. The information obtained from each of the above sources is summarized in Table 3.3 in the Appendix.

**WHAT THE STUDY REVEALS**

This section has two parts. The bulk of the first part explores the actual testing experience that our friend Nate most likely had — events that typically occur for students and proctors as a Level 4 test is administered. Through this it reveals relevant discoveries about the 168 students who failed the Level 4 test, and eventually about Nate himself. It also analyzes certain aspects of all Level 4 students during a single term, in order to provide a basis for comparison of student retention rates between the overall Level 4 population and those who failed the test. Late in my studies I realized that the results of this study would be more meaningful if we knew not only about students who failed the
Level 4 test, but also about students who passed. How many of them have we retained? Of those we lost, when did we lose them? Determining the fate of both groups allows AUA to better understand the trends at work and how it serves its students.

The second part of the study looks only at a single aspect of the listening section of the Level 4 test, as a brief exploration of how this section may or may not be hindering our students’ performance on the test.

PART 1 – TRACKING STUDENTS WHO FAILED THE LEVEL 4 TEST

The testing experience.

Some aspects of the testing experience I describe below cast doubt on the overall validity of the Level 4 test. They are later addressed in both Chapter 3, Literature Review, and Chapter 4, Recommendations.

From August 2006 through July 2007, 1,288 students registered to study Level 4 at AUA Rajdamri. As any AUA teacher will tell you, the vast majority were likely to be motivated and enrolled out of their own self interest. This would certainly describe Nate, who enrolled in a 7:00 to 9:00 PM class entirely on his own free will. Others were undoubtedly being pushed or required by employers or parents, or felt compelled to study in order to support future endeavors. Nevertheless, it’s easy to imagine that on the first day of each term, Nate and his classmates were brimming with enthusiasm, if not a touch of anxiety about the impending test that awaited them at the end of the term in six weeks.
Students had been on a break for a week since completing Level 3, as there is a one-week interim between terms. During that week, Nate’s routine didn’t change that much. He still spent his days working in the restaurant, but put in extra evening hours since classes were not being held. On that first day, he recognized all of his classmates, since new students are not allowed to start at Level 4. That is because it is a testing level and students have almost completed their course books.

After about one week of class, Nate’s teacher would most likely have read a detailed announcement to the class about the Level 4 test, informing them of the test date, the format, and the repercussions of failing — returning to Level 2. Sometimes, however, the testing secretary overlooks this form and it doesn’t get distributed until later in the term. Other times, the form gets lost and just doesn’t get read. I know this because I’ve lost it more than once myself while teaching Level 4, and know others who have as well.

As the seven terms of the Research Period progressed during 2006 and 2007, a total of 104 students out of the original 1,288 who enrolled would drop out, representing 8.1 percent of all those who registered. Interestingly, the dropout rate would be consistent between large classes with eighteen or more students, and small classes with only seven to seventeen students. If students were receiving less individualized attention in those larger classes, it didn’t discourage them to the point of dropping out any more than it did in small classes. Nate’s class was large, with 24 students.
Those 104 students who did drop out didn’t return to sit for the Level 4 exam in Week 6. That left 1,184 students remaining who actually sat for the test during the Research Period. Table 3.4 summarizes enrollment, absentee dropouts, and student numbers for the Level 4 test.

<table>
<thead>
<tr>
<th>Total enrollment</th>
<th>1,288</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of dropouts from absenteeism</td>
<td>104</td>
</tr>
<tr>
<td>Remaining students who sat for L4 test</td>
<td>1,184</td>
</tr>
</tbody>
</table>

**Week Six: The Day of the Level 4 Test**

The Level 4 test is almost always administered early in Week 6, the only exception being when a holiday falls in Week 6 and tests start late in Week 5. Except for classes that began at 7 AM, which meet daily for one hour, the tests are administered during the second hour of two-hour classes. The first hour on the day of the test would have been a normal class, except for the nervousness of the students. Then, during the five-minute break, their teacher would have erased anything on the whiteboard, packed up his/her materials, wished the students good luck on the test, and departed. Moments later, a proctor would arrive. The proctor would also have just finished teaching for an hour in another classroom, packed up his/her materials, and returned to the academic office to retrieve the testing materials, which are packed in plastic folders and stacked in the testing supervisor’s office. Before leaving with the pack, the proctor would first have to sign his/her initials on a sign-out sheet.
If the proctor were a new teacher, he or she might have looked at the testing materials in advance, perhaps even with a mentor. (It is assumed that more experienced teachers know the routine.) In the package is an instruction page. For me, it takes about four minutes to read, and another few minutes to review the detailed procedures until they become clear and properly ordered in my head.

As mentioned, proctors have five minutes to end their previous class, pack up their materials, return to the academic office, retrieve their test pack, sign out for it, review the instructions, take care of any personal needs, and report to their new testing room.

Once in the classroom, the proctor would check to ensure that the students’ regular teacher had already set up the chairs into well-spaced rows. He would unpack the materials, set up the listening CD, distribute answer sheets and provide instructions for students to complete the required personal information at the top of the sheet. He would then pass out the test booklets, and instruct the students to write the number of the student booklet in the space provided on their answer sheet, perhaps paraphrasing the instructions on the proctor’s sheet for the students, which were intended only for the proctor. However, the test itself has no introductory instructions for the students, so the proctor may (or may not) choose to share the basic framework of the test with the students. Once the test is completed, the class is dismissed. Both the students and the teacher are free to go – an incentive for the teacher to move along.
The listening CD was recorded in-house approximately ten years ago, using teachers and staff members. The CD begins with spoken instructions for the listening section. For the most part, these instructions on the CD are clear. However, once the actual listening questions are presented, the volume drops off sharply and the sound quality deteriorates significantly. The teacher may or may not adjust the volume, but he would have to do so while the students are listening to the test, which would be distracting. The proctor’s instructions state explicitly that the CD is not to be stopped once it is started. Also, the proctor may feel that stopping the CD would create even more problems, since many of the CD players are unreliable, and attempts to adjust them are met with unpredictable results.

Since the conversations on the CD were recorded using early digital technology, the timbre of the voices is very tinny. If the proctor adjusts the volume upward, and he most likely will since the drop in volume off is so significant, the voices become grossly distorted. It’s as if the speakers’ mouths were too close to the microphones. Further, the speakers on the CD player vibrate noisily due to the distortion. It is, in fact, impossible for me, a native speaker of English, to understand entire phrases at this volume. Throughout all of this, the students are just setting out on their first formal test at AUA.

One helpful technique is for the proctor to turn off the air conditioners, thus allowing the volume to be reduced somewhat. This suggestion isn’t in the instructions, so it may or may not occur to the teacher. Finally, the listening section is played only once. Afterward, students continue with the remaining sections.
Two days after the test, the students would have returned for their final class and to receive their test results. They should have been instructed by their teacher to bring tuition for the upcoming term, as registration takes place during the final class. Most students would have stopped by the bulletin board outside the academic office to check their test score, a process that is described in more detail in Chapter 1.

As we know, 168, or 14.2% of students who took the test, didn’t pass the test. From the board they would have learned about the two “problem areas” they had in the test, but they wouldn’t have seen an actual score. A note on the board would have instructed them to visit the academic office if they had a question. Often, students can be seen looking sheepishly at the glass doors of the academic office, perhaps gathering the courage to enter and speak to a farong (foreign) supervisor about their problems.

Once they saw their score, some of the students who failed the test left the campus without attending the last class. Others, perhaps anticipating that they failed, didn’t show up for class but instructed a friend to call them with the news. They wouldn’t have received the benefit of any guidance from a supervisor, or their teacher — had the teacher chosen, or had the time, to offer it. It’s impossible to know how many students left, because most teachers mark all students “present” for the final class. Therefore, no reliable record exists. This is understandable, because teachers are required to finish their record keeping — filling out individual student cards, grade sheets, and perfect attendance certificates — prior to class time. Thus, there is no manageable way to keep
accurate attendance records on the final day without possessing superhuman skills, especially for teachers in larger classes.

Students who failed the test and reported to class would have seen their student cards, with “Next Level: TWO” printed on them in red ink, as well as a large “RS,” also in red. Their teacher might have spoken to them about the implications of their failure and the next steps for them to take. Then again, the teacher may not have explained this. Either way, if students want to continue their studies, they are right away hustled down to the registration office to pay for the next term.

**What happened to students who failed the test?**

Table 3.5, below, reveals that of the 168 students who failed the Level 4 test during the Research Period, 117 of them (69.6 percent) immediately dropped out (line 3). Another 51 students (30.4 percent) took the Level 4 retest in the *next term*, indicating that they registered for and studied Level 2 again (line 4). Of those 51 students, 31 students (60.8 percent) passed the retest.

In *later terms*, another ten students from the 117 who dropped out returned to AUA, registered for Level 2, and took the Level 4 retest (line 6). Of those ten students, five (50.0 percent) passed (line 7).

The table indicates that students who did not take one or more terms off outperformed those who did on the retest, passing it at a rate of 60.8 percent, compared to 50.0%
percent (lines 5 and 7, respectively). Finally, the 36 students who eventually took and passed the retest (lines 5 + 7) represent about one in five of the 117 students who failed the Level 4 test (21.4 percent).

### Table 3.5
The Level 4 Test
Summary of Retest Pass/Fail Rates During the Research Period

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number of students who sat for the Level 4 test</td>
<td>1,184</td>
</tr>
<tr>
<td>2</td>
<td>Number of students who failed the Level 4 test</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>(14.2% of 1,184)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Failing students who dropped out</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>(69.6% of 168)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Failing students who took the retest the next term</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>(30.4% of 168)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Students who passed the retest</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>(60.8% of 51)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Dropouts who later returned and took the retest</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>(6.0% of 168)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Dropouts who later passed the retest</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(50% of 10)</td>
<td></td>
</tr>
</tbody>
</table>

**Of students who failed the test, who dropped out?** For students who failed the Level 4 test, one important thing to look at is who registered for Level 2. Is there any way for us to determine what factors contribute toward a student’s decision whether or not to register? Actual rates of registration for Level 2 by students who failed the Level 4 test vary significantly from class to class and term to term. Table 3.6 in the Appendix summarizes these rates by term. The table reveals that in Term 4, 2007, 57.1 percent of students registered for Level 2, as opposed to Term 2, 2007, where only 10.3 percent registered. Were students in Term 4 more motivated than those in Term 2?

A closer look at these rates, class by class, raises the question of whether or not teachers or other factors play a role in influencing students who fail the test to register for Level 2. Registration rates varied widely, and rates for one term, Term 5, 2006, are shown in
Table 3.7 in the Appendix. The table shows that of the 35 students who failed the test in Term 5, 2006, only ten registered for Level 2 and took the retest in Term 6, 2006. That’s 28.6 percent of the 35 who failed the test that term. In some sections, none of the students registered for Level 2 (see especially Sections A303 and TD406). The variation in registration rates from class to class indicates a need to look carefully at the factors involved in guiding students who fail the test. There is insufficient data to further explore what these factors might be, but methods to generate useful data in the future are discussed in Chapter 4.

Similar stories can be derived from other terms. For example, in Term 2, 2007, 29 students failed the Level 4 test, but only 4 registered for Level 2. That’s only 10.3 percent of those who failed the test (shown in Table 3.6 in the Appendix). Clearly, AUA would benefit by exploring the reasons why so few students who fail the Level 4 test are registering for Level 2. The rationale for obtaining student feedback on matters such as this is explored in Chapter 3, while suggested methods are offered in Chapter 4.

**Do we know what happened to Nate?** Yes, we do. Nate responded to his failure the way most AUA students do — he dropped out of the program. He had never failed a school test before (all students were routinely passed at the public school he attended in Nahkon Si Thammarat), and we can assume that he felt sufficiently discouraged by the test results and the limited options available to him to drop out of the program. Nate’s Level 4 teacher has no recollection of whether or not he attended class on the last day, though his attendance card shows him as present. No one at AUA has attempted to
contact Nate since he dropped out. His telephone number, Bangkok and Nakhon Si Thammarat addresses, and his email address are on his original application, but that information was never entered into the appropriate fields in the computerized registration system. So, in truth, we don’t really know what happened to Nate after he dropped out, so we have the power to find out. All we need is the interest, which is something I address in Chapter 4.

For visual learners, Nate’s story, and the story of his 168 classmates who failed the Level 4 test at some point during the Research Period, is represented in Figures 3.1 and 3.2 on the next two pages.

First, we meet Nate again.
1. Meet Nate’s classmates who also failed the test – all 168 of them.

2. 117 of the group said “No thanks” to repeating Level 2 and just went home. Unfortunately, this included Nate. 51 took Level 2 again and retested.

3. Of the 51 who retested, 31 passed the retest (about 61%) and were eligible to study in Level 5.
4. And what of our 31 torch-bearing heroes who passed the Level 4 retest? Where are they today? 21 students dropped out at some stage of the program.

21 students:

6 QUIT in LEVEL 5:
1. Mr. Amnart Molispelee
2. Miss Tanita Kituerkasamwong
3. Miss Sasinee Boonseang
4. Mr. Assawapong Assawalerpaiboon
5. Mr. Khongchat Chaisrivibul
6. Miss Sudarat Aremwit

6 QUIT in LEVEL 2, right after the retest:
1. Miss Nichanat Tareuluk
2. Miss Walai Phattarakijkul
3. Miss Waranya Thaweewitthananaskul
4. Mrs. Kanyarat Thanabat
5. Miss Kornthip Kalapak
6. Mr. Sanchai Veerannas

2 QUIT in LEVEL 6:
1. Mr. Tanapope Pratheerjarusrun
2. Miss Sukanya Pakakpong

2 QUIT in LEVEL 7:
1. Miss Areerat Chuesailucang
2. Miss Anchalee Jusee

4 QUIT in LEVEL 8:
1. Miss Netnapha Buaban
2. Miss Saranrom Sitpol
3. Miss Changfong Nimitpontorn
4. Miss Pakpanichchar Pharwalsirihihgool

1 QUIT in LEVEL 9:
1. Miss Pornnaphat Nonkrathok

10 students:

1. Mrs. Budsayamart Fungfuang
2. Miss Arpapon Konprajong
3. Mr. Sakarin Petchpornsikikul
4. Miss Passarin Intachaom
5. Mr. Anusat Ketwong
6. Miss Luksika Ruangsung
7. Miss Sukanya Yeanghup
8. Mr. Satit Lapcharoen
9. Miss Pichmear Soonsap
10. Miss Jeeranun Kumpeng

10 ARE STILL STUDYING TODAY!!!
**Who remains?** As Figure 3.2 shows, ten students out of 168 students who failed the test are still studying. That’s only 6.0 percent, with 94.0 percent dropping out at some point before completing the program. However, the ten students who are still studying actually represent 32.3 percent of the 31 students who took the retest and passed it. So, about one-in-three students who passed the retest is still studying at AUA. We will later compare the dropout rates among students who failed to *all* students within a single term.

**When did the students who passed the retest actually drop out?** Did our 31 students who passed the Level 4 retest continue to study at AUA for several terms, or did they drop out quickly? A look at Figure 3.2 shows that of the 21 students who dropped out after passing the retest, six students did so immediately (quitting in Level 2 after taking the retest), and another six students did so in Level 5. That’s already more than half (57.1 percent) of the post-retest dropouts. Another two students dropped out in both Levels 6 and 7, four students in Level 8, and one student in Level 9, bringing the total to 21 students. After Level 9, the remaining 10 students stayed with the program until completing it.

**Comparing Our 168 Students to the Student Body Overall**

How might our 168 students who failed the test compare to AUA’s overall Level 4 student body? How many of the 1,288 people who originally registered for Level 4 during the Research Period are actually still studying at AUA? I examined one term — Term 6, 2006 — in order to take a sample. With registered enrollment of 191 students in that term, I determined that the sample size would be adequate in order to provide reliable
data enough for the comparison. For reasons of practicality, I did not look at all 1,288 students, since this portion of the study necessitated looking at enrollment histories one-by-one.

In Term 6, 2006, I noted (a) if the student was still studying today (Term 2, 2008) and (b) if not, in what Level the student dropped out. A student on a “normal” progression would be in Level 14 during Term 2, 2008. Table 3.8 summarizes my findings.

Table 3.8
Summary of Enrollment Trends for All Students
Enrolled in Level 4 During Term 6, 2006

<table>
<thead>
<tr>
<th>Level of Dropout</th>
<th>Dropouts</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quit after Level 2*</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td>Quit after Level 3*</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Quit after Level 4</td>
<td>47</td>
<td>24.6</td>
</tr>
<tr>
<td>Quit after Level 5</td>
<td>40</td>
<td>20.9</td>
</tr>
<tr>
<td>Quit after Level 6</td>
<td>33</td>
<td>17.3</td>
</tr>
<tr>
<td>Quit after Level 7</td>
<td>20</td>
<td>10.5</td>
</tr>
<tr>
<td>Quit after Level 8</td>
<td>16</td>
<td>8.4</td>
</tr>
<tr>
<td>Quit after Level 9</td>
<td>13</td>
<td>6.8</td>
</tr>
<tr>
<td>Quit after Level 10</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td>Quit after Level 11</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>Quit after Level 12</td>
<td>2</td>
<td>1.0</td>
</tr>
<tr>
<td>Quit after Level 13</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>Quit after Level 14</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>191</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

| Currently enrolled    | 12       | 6.3        |

* Students who dropped out after initially failing the Level 4 test and subsequently registering for and studying in a lower level.
As Table 3.8 indicates, almost half (45.5 percent) of all students in Term 6, 2006 dropped out in Levels 4 and 5. Another 36.2 percent dropped out shortly thereafter, in Levels 6, 7 and 8. Thus, from the point where students were ready to enter *Interchange 2* (Level 5) to the point they would finish that book (Level 8), AUA lost 81.7 percent of its students who were studying at Level 4.

We can also see that very few students drop out after Level 9. However, of the 20 students who did drop out in higher levels (9 and above), fully 65 percent do so in Level 9 – at the Introduction of the *Sky High* series.

How do overall dropout rates for Term 6, 2006 compare to those for students who passed the Level 4 retest? The rates are nearly identical. A full 93.7 percent of the total student body in Term 6, 2006 dropped out, compared to 94 percent that dropped out among students who passed the retest. Statistically, the rates are the same.

**Conclusion.** For teachers and supervisors alike who have worked at AUA for several years, none of the above statistics would be very surprising. Yet, as AUA embarks on forming a strategic plan, initiated as a result of declining enrollment, the above may prove useful in helping to begin to understand enrollment trends and how AUA serves, or fails to serve, its students. There isn’t much good news in this chapter. However, one statistic that I feel is encouraging is that 32.3 percent (10 out of 31, as shown in Figure 3.2) of those students who passed the Level 4 retest are still enrolled at AUA. That compares to only 6 percent overall. Perhaps in this irony those students who originally
failed the Level 4 test have a story to tell that would benefit students, teachers and administrators. In Chapter 4, I explore opportunities for students to share more of their experiences.

PART 2 – QUESTION #1 ON THE LISTENING PORTION OF THE TEST.

As I described in Chapter 1 and above, part of my original inquiry in September 2007 was to look at the answer sheets of students who took the Level 4 test to calculate if many were answering the first question incorrectly. This was because I had interviewed a number of students who failed the test, often showing particular weakness in the listening section, but whom I assessed during interviews as having good listening skills. Could there be something about the test that was hindering our students’ performance? If so, was this hindrance significant enough to cause students who might otherwise be able to study successfully in Level 5 to fail the test?

Methodology. My examination of student answer sheets was limited to Terms 5, 6 and 7, 2007, since answer sheets from previous terms had already been destroyed, as I describe in my introduction. For each section, I also looked at other factors:

1. test version (there are eight forms);
2. proctor’s name;
3. section teacher’s name;
4. classroom size (large or small); and
5. number of students in each class.
I noted which test version was used because I wanted to see if higher rates of incorrect answers were associated with particular versions of the test. I noted the name of the proctor because I suspected variations in the way the tests were administered. I noted the section teacher simply because the information was available (I do not suspect a correlation between students’ scores on the listening section of the test and teacher performance over the course of a single, 30-hour term). Next, I noted classroom size (S/L) as an example of an environmental variant. Finally, I noted the number of students in each class. This was necessary in order to calculate percentages of incorrect answers, but it might also be an indicator of student performance.

Of the above five factors, none showed a notable correlation with student performance in regard to the first question on the test. Table 3.9 in the Appendix summarizes all of the information I collected over the three terms available to me.

In Term 5, 2007, 61.0 percent of all students answered the first question in the listening section incorrectly. Terms 6 and 7 were 56.9 percent and 54.5 percent, respectively. The average for all three terms was 57.4 percent.

Surprisingly, the number of students in the classroom did not seem to have a negative affect on student performance. Classes with 18 or more students are summarized in Table 3.10 in the Appendix. Some of the best student performances can be found in those classes, and overall performance among this group was actually better than that of the entire study group: incorrect answers were chosen by 50.4 percent of students in
large classes, vs. 57.4 percent overall. Two sections shown in Table 3.10 had 100 percent of students answering the first question incorrectly. These classes had only nine and ten students, respectively.

The wide range of performance between sections and between class sizes would indicate that other variables are at play. Overall, is a 57.4 percent rate of incorrect answers on the first question statistically significant? Just what is an acceptable rate? The answer to this question can be calculated and needs to be known by AUA’s testing administrators.

What is particularly striking about the rates of incorrect answers with the first question is that, despite the overall average that falls in the mid-range, there are indeed certain class sections where the rates are entirely unacceptable, by any standard. What happened in the sections where the vast majority of students answered incorrectly? Table 3.11 in the Appendix shows that there are nine class sections where at least three out of four students answered incorrectly. That represents nearly one in four sections.

**Conclusion.** Despite the fact that for all class sections the overall average falls below sixty percent, what of those classes with much higher rates? How can these rates be lowered? Is there a valid argument for “easing” students into a test? These and other questions will be explored in Chapters 3 and 4.

Chapter 3, Literature Review, examines the views of others in the profession in regard to various aspects of assessment that are relevant to this analysis. Chapter 4 then applies the
principles outlined in the Literature Review to a series of suggestions for how AUA might address these issues. Therefore, I refer the reader to Chapter 4 for a more complete conclusion.
CHAPTER 3
LITERATURE REVIEW

INTRODUCTION

This review is a modest sampling of the literature that addresses specific aspects of assessment that are particularly relevant to this project. The review was revealing to me in a number of ways. First, it highlighted the need to clearly define one’s project early on, in order to give the search a useful focus. Second, it helped to define the beliefs that respectable leaders in our profession hold toward assessment, and it allowed me to compare those beliefs to my own. Finally, it provided support to me as I thought of how I might offer suggestions to AUA to improve upon its assessment and support of its students.

I am, in all honesty, still not certain of the precise focus my literature review should have taken. Initially, one would head toward literature on assessment, as I did, but the story is be incomplete. I suspect there is a rich, relevant body of literature to be discovered and exploited. Yet, if there is, I didn’t find it.

Part of this was outright resistance. I wanted to depart into uncharted territory on a unique exploration; to be free of constraints imposed by academicians. I wanted to read
literature that soared, as Parker Palmer’s does to me, rather than analyses that anesthetize.

Yet, the cold reality was that I was analyzing numbers. Nothing fancy.

Despite the numbers, I wanted to avoid margins of error, scales, and the like; cast aside the hard science and look at the social science, focusing on the human faces just under the numbers. Surely they would have stories to tell.

And what literature might guide me through this process? Most of all, I wanted literature that would validate the outrageous awkwardness and callousness I feel each time I tell a student who has failed a test, and who needs support, that his only recourse is to head back to Level 2 — again. Or a third time. Or a fourth. If there were a useful body of literature out there, it would tell me what on earth we are doing at AUA to repel 94 percent of the Level 4 students into leaving the program whether they pass or fail the test. It would tell me how to instill the love of learning within a system that has neither the time nor the interest for such. It would teach me how to persuade others to construct a supportive remedial program for those in need. As I would eventually discover when writing Chapter 4, Recommendations for AUA, it would guide me through a process of cultural healing, and teach me how to motivate and lead.

**WHAT I FOUND**

In my search of literature on assessment, I mostly found a selection of books and articles that explain how to design tests and how to conduct research. I drew from a limited number of resources: the personal libraries of three friends, and the library at AUA. I
was also fortunate to receive an electronic version of a complete class module on assessment from a friend who is a candidate for a master’s degree at Wollongong University. This module contained 71 separate articles on assessment, including a few from authors I cite in my bibliography. However, only one article on washback, by Luke Prodromou, was relevant to my particular focus. The remaining articles focused solely on test design from a perspective that seemed to exclude my actual goals.

As I mentioned in Chapter 2, my first source of supporting information was James Dean Brown’s *Understanding Research in Second Language Learning*. I spent considerable time in this book fully expecting that I would soon be laying out a complex statistical analysis. Yet my goals seemed anything but complex; I merely wanted to trace the progress of a group of students, and see how many of them answered a single question incorrectly. The real questions I wanted to answer are *how to help those students*, and how to improve a test that is merely one component in a much larger system that is itself *based on beliefs often far different from my own*.

Still, I have learned from what literature I read. I learned more about constructing valid tests and how feedback from teachers supports motivation and learning. I learned that my suspicions about the important role of proctors are valid. I learned about important and beneficial alternatives to testing. Finally, I learned more about testing listening.

All of these discoveries are supported herein through the literature I explored. They shall be applied to this study in the final chapter, which provides recommendations to AUA.
I often think about the crowds of students who gather around the bulletin board during the final week of class each term to check their scores on the level tests. As I described in Chapter 2, the only “score” they receive is “P” (pass”) and “NP” (not pass), followed by a brief indication of two “problem areas” for students who fail.

**Washback.** There is much about *washback* in the literature on assessment. Bachman and Palmer (1996, 30) start off on a solid footing by stating that “it is widely assumed to exist.” Assuming it indeed does, washback is sometimes referred to as a potentially negative characteristic of testing that can result in teachers “teaching to the test,” to produce the highest scores among students. James Dean Brown defined it in 2005 as how the test can affect the curriculum. However, washback also describes the potentially positive feedback that an institution and teachers can give to students as part of their assessment. H. Douglas Brown set a worthy goal for teachers in 2004 by stating that good tests should lead the way toward positive washback. Yet, he also acknowledged the perspective that dominates test administration at AUA, which is that practicality usually wins over washback.

There is no denying the day-to-day importance of practicality at an institution such as AUA, where several hundred tests are administered every seven weeks. However, there are also considerations that should not be automatically overshadowed by practicality.
James Dean Brown summarized (2005) the widespread beliefs that not only administrators, but also teachers need to be involved in all aspects of test administration, and should understand the purpose and goals of the each test. Further, teachers need to be properly supported with instructions, and receive and give feedback about the test. Finally, it is very important that students get supportive feedback, including actual test scores.

Michael Rost (2002) addresses what some teachers at AUA describe as a Catch-22 scenario, where they feel the test is not an accurate reflection of what is taught in the course books. Many teachers feel that the textbook doesn’t adequately prepare students for the Level 15 exit test, so they devote the majority of classroom time in Level 15 working on the specific skills addressed in the test. At AUA, the result is a tacit acknowledgement between administrators and teachers to put the course book aside, assess the weaknesses of the students in relation to what is on the actual test, and to teach to the test.

Turning again to the positive attributes of washback, Rost (2002) believes that positive washback results in more motivated students and makes students feel as if they have more access to their teachers. H. Douglas Brown (2004) takes this further by laying out a framework for teachers and test designers to incorporate useful feedback into the assessment process. First, students need more than a simple letter grade. They need tests that promote feedback, since students actually learn about their language skills in this manner. Washback increases both learner autonomy and confidence.
Bachman and Palmer (1996) sum things up by reminding us that washback isn’t just a simple matter of the effect of tests on teaching, or of negatives or positives, but is actually quite complex. It is worth noting that in Chapter 8 of Bachman and Palmer, “Identifying, Allocating and Managing Resources,” the otherwise detailed planning process for developing tests fails to include provisions for providing positive washback to students.

**Instructions to Students.** It is interesting to note that the AUA Level 4 test contains no introductory instructions to be issued prior to the start of the test. However, each section in the test booklet does include brief procedural instructions. Bachman and Palmer (1996) make a strong case for more complete instructions, stating that instructions have the potential to clarify the range of students’ questions, doubts and expectations about tests regarding their relevancy and fairness. Clearing this up, or at least minimizing it, helps to motivate students and to improve their performance.

The basics of test instructions are well addressed by Bachman and Palmer (1996), and include a statement regarding the test purpose, the actual language abilities that will be tested, a description of the layout of the test, a summary of the procedures that will be followed, and information on the scoring method.

**Variance in the Testing Experience.** External/environmental variables are important considerations. While the test itself may remain consistent from one class to the next, the
methods with which the test is administered can vary greatly. James Dean Brown (2005) addresses how variables such as equipment (in AUA’s case, CD players), lighting, room temperature, noise from external sources and variation in timing can become sources of measurement error. Indeed, the proctor may also contribute to errors by giving the instructions poorly, failing to answer initial questions that students may have, rushing, showing nervousness, etc. Unclear instructions within the test itself may also contribute toward measurement variances that have nothing to do with the purpose of the test.

**Validity.** And what of the test itself? Of the many factors that join to make a good test, which might AUA examine with a renewed interest? Content and face validity are crucial. James Dean Brown (1988, 102) defines content validity with a question, managing to be concise, yet a bit circular at once: “Is the test a representative sample of the content of whatever the test is claiming to test?” Indeed, Michael Rost (2002) addresses both forms of validity, explaining that tests should be directly related to classroom material and tasks. Without this, tests will not have face validity and will not feel familiar to the students. Further, they will lack content validity by not having similar topics or pulling from the same source of information that was covered in the class. Rost also discusses the notion of procedural validity, or the relationship between the test-taking conditions and the actual teaching situations (1990).

**Feedback about the testing procedure itself.** Teachers, proctors and students experience tests first-hand, but test developers may not. Thus, Bachman and Palmer (1996) include feedback to these developers about the testing procedure as a useful and
necessary element of test planning. Their suggested methods for obtaining feedback include questionnaires, think-aloud protocols (accounts given by students of the processes they go through when taking a test), descriptions by an outside observer, and interviews of test-takers.

**Alternatives to Testing.** I have already stated my lack of success in identifying literature that directly addresses how AUA can better serve students who fail the Level 4 test, *after* the assessment process. However, at least some of those students could be better served by altering, or at least adding to, existing assessment. This will be explored in Chapter 4.

H. Douglas Brown brings tests off their lofty pedestals. “Tests are simply measurement tools…Assessment includes all occasions from informal impromptu observations and comments up to and including tests.” (2004, 251). He clearly places trust in teachers’ ability to assess, not just in formal tests themselves. In fact, he alludes to the superiority of alternative techniques, saying they offer more washback, can be potentially more accurate, and carry greater face validity through increased authenticity.

An inherent trust in teachers allows us to expand into further possibilities. H. Douglas Brown advocates **triangulation**, or the consideration of two or more performances before assessing, relying as much as possible on observable performance. His rationale is that all abilities of a student are not revealed in a single situation. He also stresses the reliability of teachers’ intuition, which may include highly accurate assessments that are
otherwise difficult to measure. Brown believes that these observations carry real validity and should be included in the determination of a student’s final grade.

Triangulating empowers teachers, but also places greater responsibility on administrators to be rigorous in determining criteria for evaluation. Finally, it asks for a greater involvement from the students and teachers alike. Since the process of triangulation is broader and much richer than that experienced in a simple multiple choice test, students and teachers are likely to be more motivated.

Like triangulation, conferences rely on an increased role of the teacher, with H. Douglas Brown (2004) believing that conferences put teachers in the role of facilitator and guide, rather than simply that of a test administrator. Students would therefore need to redefine their own views towards teachers and tests, and accept the role of teacher as ally.

The great challenge, of course, is how teachers at institutions such as AUA can manage these additional roles, when they have neither the time nor the support from school administration to do so. Genesee and Upshur acknowledge similar limitations, but offer a compelling case for integrating conferences into the curriculum, in a manner that would likely be achievable in almost any academic setting (1996). In their view, conferences can be as simple as conversations or discussions between teachers and students about school work. They can vary in the number of participants, and can relate to work completed or in-progress. Often, the topic of discussion is designed specifically for the purpose of the conference.
Genesee and Upshur (1996) also believe that conferences offer teachers a special insight into the lives, culture, and most importantly the learning styles of their students, which teachers may otherwise overlook. This is particularly noteworthy for western teachers in Bangkok. At AUA, managers would think that the above would be a prerequisite to holding conferences, when in fact the conferences themselves would provide learning experiences for the teachers.

**Interviews** are useful tools that are employed to a certain extent at AUA. As I mentioned in Chapters 1 and 2, I and other supervisors have the option of interviewing students who failed the Level 4 test. True, as presently structured they are not very practical because of time constraints on supervisors, and reliability varies from one interviewer to the next. This can easily be solved by having a trained, designated teacher on duty (“admin” time) during Week 6 to conduct interviews. H. Douglas Brown believes reliability can be maintained if interview questions are carefully constructed to produce focused responses (2004). He also believes that appropriately constructed interviews and conferences can have a high degree of content validity if they remain focused on course objectives.

**Assessing Listening**

Since listening is a receptive skill, it is difficult to isolate it from the other skills and to assess it without also assessing those skills. Rost (2002) and Brown (2004) believe it is very difficult to test listening only, and that any test will invariably be looking at other skills as well. Brown states that “…all assessment of listening and reading must be made
on the basis of observing the test-taker’s speaking or writing…so all assessment of receptive performance must be made by inference!” (118). Listening samples should be authentic, as opposed to the sterile studio recordings we often hear. Rost believes that testing listening should include the idiosyncrasies of authentic natural speech, such as background noises, negotiated meanings, pauses, intonations, etc.

Bailey (1998) addresses the importance of including appropriate schema in assessment, believing that mismatched schema can result in misinterpretations of messages. This is especially likely when the schema is culturally specific. Bachman and Palmer concur. They present a question that every test designer should ask when given a task (1996, 145): “Is this test task likely to evoke an affective response that would make it relatively easy or difficult for the test takers to perform at their best?”

Finally, Michael Rost makes an interesting case for designing listening assessment so that the receptive role of the test taker remains clearly defined and manageable (2002). He reminds us that one of the first tasks in a listening test is to make sure that the students understand his or her role – how to respond. Is the student an outside observer, a participant, a judge? Is the role consistent? If not, the act of switching roles should be considered a skill in itself.
Conclusion. The perspectives offered by the literature I have selected provide a solid basis for the following chapter, which focuses on how AUA might examine its own views toward assessment and assessment construction, with the goal of providing students with the best possible opportunities to advance their learning, and the best possible remedial support in those cases where it is needed.
CHAPTER 4
RECOMMENDATIONS FOR AUA

This chapter builds on the results of my analysis presented in Chapter 2 and the review of literature in Chapter 3, in order to provide suggestions for how AUA might improve both its Level 4 test and its overall support of Students who fail the Level 4 test. It is assumed that many, if not all, of the suggestions could be applied to other tests at AUA and to students in other levels as well. Unlike Chapter 2, which is divided into two sections, this chapter blends recommendations into a unified presentation, since the suggestions often relate to both the test and the students taking it. The Introduction, however, presents crucial suggestions that are necessary to consider before AUA can undertake the recommendations made in this chapter.

INTRODUCTION

Certainly no test is perfect. It is also evident that at AUA a great deal of effort has gone into making the level tests as fair, useful and practical as possible. It is to everyone’s credit that hundreds of students are successfully tested every few weeks, and there is widespread agreement that, for the most part, the tests are doing their job.
Yet, the stories revealed in this report are deeply discouraging. Teachers at AUA know that large numbers of students drop out of the program, and the data herein confirms this suspicion. Administrators may or may not know this. To date, dropout rates have never been analyzed. Further, administrators acknowledge that in many respects they know little else about AUA students.

What has lead to an environment where policies are built upon suspicions and tacit acknowledgements, rather than more positive attributes? No doubt the answers are complex, compelling, and worthy of deeper analysis. They are addressed below, however, as necessary prerequisites for the recommendations made in this chapter.

Three factors are explored: first, challenges of working in a bi-cultural setting; second, attitudes of management toward teachers; finally, the interweaving of these two factors into a all-inclusive, self-perpetuating paralysis.

**Challenges of Working in a Bi-Cultural Setting.**

Our Thai hosts are among the most gracious, generous and accepting people in the world. Nonetheless, some from the west find it challenging at times to understand and work successfully within Thai social structure and customs. As a result, certain perceived barriers exist at AUA that have grown over the years to effectively inhibit communication between the academic office, where management is entirely western and support is Thai, and all other departments, where both management and support are Thai.
From the perspective of some western staff, many existing policies and procedures feel locked into place, because changing them would require the cooperation of other departments. This perception has discouraged managers and teachers over the years from maintaining a “can do” work ethic. True, AUA operates in a very traditional style, where the factors at work among those who set policy and procedures, and hire personnel are quite complex and not always logical to westerners. Since this style does not always recognize merit as a primary criterion for leadership, it is often difficult for the western staff to understand and negotiate through the intricacies of effecting change.

However, the Thai staff comprises unique individuals that represent a wide range of age, experience, skills and competence, social backgrounds, interests and goals. They are not merely one body of workers united by a shared heritage. Those who are involved in day-to-day support operations know their jobs well, perform their jobs with great skill, courtesy and accuracy, earn a fraction of the salary that western staff earn, and are just as eager for change as the westerners.

And times are changing, albeit incrementally and not always noticeably. For example, while long-term managers in the academic office remember a time when the former registrar ruled with an iron fist and is said to have scoffed at its proposed changes, the new registrar shows a tremendous desire and ability to cooperate with the academic office. With this single improvement, it is now possible for the academic office to imagine any number of changes that would have been much more difficult only a few years ago. Thus, the stage is set for a dialogue between the registrar and the academic
office to create a progressive and comprehensive support network for students who fail the Level 4 test. This opportunity needs to be acknowledged by the academic office, and similar opportunities should be formally identified and nurtured through a joint effort between the offices.

It is all too easy for management to let the past define the present.

**Attitudes of Management Toward Teachers.**

The lack of outreach by managers toward teachers is likely the single most debilitating factor that prevents positive change at AUA. Specifically, managers exclude teachers from all administrative matters. The results are predictable: (potentially) overworked supervisors and demoralized teachers. However, the role of teachers is inextricably linked to student assessment, since teachers are the lifeline between students and learning. Further, they are the ones most able to give feedback on the testing experience, and to convey student feedback to test administrators.

Negative attitudes toward teachers are deeply rooted and habitually perpetuated among management. In fact, it is widely believed by management that teachers have little or nothing to offer in the form of professional input, irregardless of the teachers’ education, professional development, interest, talent and classroom experience. Thus, there are very few meaningful administrative functions given to teachers. A more typical function is secretarial work, such as re-filing supplemental materials. This type of function is sometimes given out in lieu of a class (five hours per week).
An example of the lack of support for teachers can be seen in the following conversation that was shared between this writer and another supervisor in 2006, who were discussing what classes to assign to a female teacher for an upcoming term.

A. Let’s give her the materials development slot. She’s good at that.
B. Correction: you mean materials maintenance. Teachers don’t develop materials.

This mentality has not gone unnoticed by AUA teachers, who mostly feel they are viewed with contempt by management. Do teachers want to be involved? For many, absolutely. For a few, no. Yet even those few have expressed concern for their students and naturally want to enjoy their jobs. Part of management is to lead, and part of leadership is to motivate; it is incumbent upon all supervisors to believe that most teachers care deeply about their jobs, their students, and their profession.

The contempt for teachers has spread from management to some members of the Thai staff. In the academic office, staff are mistrustful that teachers will follow basic procedures. This doubt extends all the way to the top. Certainly, the Thai director of AUA has shown little respect toward teachers as a whole, though he does make himself available to individual teachers who wish to discuss personal matters with him.
Self Perpetuating Paralysis

Add to the above scenarios other work-related factors, such as aging facilities and equipment, declining enrollment, and low wages, and one can quickly envision an environment that perpetuates stagnation, if not outright paralysis. Managers say “It can’t be done,” hindered by valid, though often outdated memories of resistance between the academic office and other departments. Teachers scoff when asked to give their free time toward a project, when their ideas and concerns are not validated by management. Thai staff in the academic office resist setting up procedures that depend on the cooperation of teachers. When a challenge arises, everyone searches briefly for an answer, remembers that they are typically not forthcoming, and gives up.

Yet, so many of the challenges faced by the academic office could be addressed through a supportive and inclusive approach toward teachers. AUA has some of the best teachers in the business in Thailand. All are certified and hold Bachelor’s degrees, and several have masters degrees. Many give freely of themselves and care deeply about the institution. Still more have demonstrated skillful and creative teaching. Collectively, they offer a vast resource that could be tapped for a wide variety of uses. Some of these are explored in this chapter. Many more remain to be discovered.

Putting it Together: The Role of Leadership in Effecting Change.

Instituting real change at AUA, such as changes in testing procedures, will take the cooperation of all involved, and that will not happen without the support of all. Thus, it is my firm belief that the only way AUA can make significant improvements is for
leaders — managers — to assume the responsibility inherent in leadership to reassess their views of teachers and Thai staff and to begin a more inclusive and respectful relationship with both. Collectively, clear tasks can then be defined, planned and undertaken by supervisors, teachers and Thai staff alike.

However, this route it will not be recognized by managers until they reach a consensus, perhaps guided by an individual or perhaps jointly determined, that cooperation, respect, and outreach are valid and essential components of effective management and leadership. In doing so, the opportunities – rather than the obstacles - will become visible.

In order for the suggestions that follow to be realistic at AUA, it would also be necessary for management to develop a working plan whereby clear, common goals are set, timeframes are established and followed, and adequate leadership is put into place to provide the vision, motivation, material support, time support, and administrative support to bring goals to fruition.

RECOMMENDATIONS

Overall Improvement: Making a Better Level 4 Test.

This section looks at recommendations for improving the Level 4 test, and also the overall assessment experience for students, teachers and management. These recommendations are organized around general themes and for the most part are
supported by the literature reviewed in Chapter 3. Occasionally, I offer a suggestion that may lack clear support in the literature, but as I have previously noted, I consider my success at identifying the full body of relevant literature to be limited indeed.

**Improve the test itself.** The Level 4 test has no instructions. In the literature review, Bachman and Palmer outline the essential components of good test instructions, and AUA should provide them. Not only will this help students perform to the best of their abilities, but it will also help them understand the basis for the test and what lies ahead for them.

The role of schema and certain aspects of test validity should also be re-examined. It can be argued that portions of the Level 4 test, particularly the listening section, present students with unfamiliar scenarios, bringing the procedural validity of the test into question. At AUA, students are taught listening in a multi-step process that includes pre-listening tasks such as introducing the topic, activating schema, introducing new vocabulary, and predicting. Only then is a listening task introduced, and it is a general task, followed on a second listening by a detailed task. In the test, students are asked to listen and respond to a number of different conversations, and are given only a one-sentence introduction, and one chance to hear the conversation. This is the first time in their studies at AUA that listening is presented in this format. I suggest that at the very least, clear examples be given and reviewed with care prior to the test. Finally, administrators should confirm that students have adequate time to read the questions (as they are instructed to do) prior to each listening question.
Rost (2002: 186) provides an example of instructions for a listening test that differ significantly from AUA’s. They are worth examining for their focus on students understanding tasks and having adequate opportunities to hear the listening samples. *(Italics added.)*

Begin playing Part 1 of the test. Stop the CD after conversation 1, question 1. *Be sure that all students understand the procedure.* (You may wish to use questions 1 and 2 as examples.) *Replay conversation 1* and continue through question 20 without stopping.

Begin playing Part 2 of the test. Stop the CD after conversation 1. *Be sure that all students understand the procedure.* (You may wish to use conversation 1 as an example.) *Replay each conversation so that students have an opportunity to hear each conversation in part 2 two times.*

*Return scored test sheets to the students and go over the test one more time for feedback.*

I believe that a more student-centered approach toward test administration should be considered.

Tests are also administered in an inconsistent manner. As described in Chapter 2, proctors simply don’t have enough time in their five-minute breaks to adequately prepare for the test. AUA should consider a standardized briefing for proctors, and better support for proctors who have questions, can’t locate their test packs, or show signs of not understanding the testing process. Proctor instructions should be edited and simplified, and included as an addendum in the (soon-to-be?) teachers’ handbook. They should highlight the need for proctors to consider environmental variables such as noisy air
conditioners on the test takers. Clear reminders about key points (test dates, procedures, etc.) should be posted in the teachers’ lounge in advance of the tests.

As currently implemented, testing classes are dismissed at the conclusion of the test. This is a strong incentive for proctors to move as quickly as possible. If the experience were to include new components, such as more detailed instructions, better examples, feedback sessions, and a closure activity, the testing hour of class could become a rich learning experience, and better paced to support the students. Early dismissal should be re-evaluated.

Positive Washback. Positive washback is the potentially positive feedback that an institution and teachers can give to students as part of their assessment. This usually takes the form of details about actual performance on a test or assessment project, with emphasis on strong areas and areas for improvement. Often, suggested methods for improving are included.

AUA should reassess its perception of balance between usefulness and practicality and determine how it can give students more positive washback on the Level 4 test. True, there are real constraints on practicality if the feedback were to be presented on the public bulletin board, because the additional information would have to be produced by the testing secretary, who must already process testing for all students within a very short period of time.
However, positive washback could be brought into the classroom, where I believe it would be most beneficial. Since the final class of each term is considered wasteful by many teachers and students alike, teachers and supervisors should design a questionnaire for final classes that includes opportunities for individualized feedback to students. Of course, people will say “What will the other students do while the teacher speaks to individual students?” This is not an insurmountable problem. For example, I have seen two classes join together on the final day to form very successful sessions. This could free up one teacher to meet with students individually. With proper mixing of abilities, it could also serve as a useful motivational and marketing tool for students. For example, if Pre-Level A students interact in English for 30 minutes with Level 4 students, they can see how they will progress in the near future. Level 4 students would see how they have progressed. This example is but one of several possibilities that could be considered.

Finally, students who fail the test need significant support at this point. They need to clearly understand their strengths and weaknesses, and also their options (see below). These options need to be provided to students in writing, so they may discuss them with their parents. AUA also needs to explore how to get more students who fail the test to attend the final session of class. If they are not present, they cannot benefit from feedback. For those who still do not show up, AUA should reach out in the form of supportive telephone calls or emails. First, however, this information would have to be entered into the computer system.
Formalize teacher involvement in assessment of students who fail the test. As I described in Chapter 2, when a student fails the Level 4 test, the teacher has the option to consult with a supervisor if the rate of failure is marginal and the teacher believes the student can successfully study at the next level. This is terrific but doesn’t go far enough. Teachers need to know that they have this power, and they need more of it. They need to receive not only the test results, but also clear instructions about available options and/or requirements made of them, plus the time to carry them out. There needs to be a collective understanding between all teachers and supervisors that supportive intervention on behalf of students is a priority. During the final week, a designated supervisor needs to be available for this purpose at designated times. A more informal structure may lead an otherwise occupied supervisor to see this process as an administrative inconvenience, or a supervisor may simply be unavailable.

Alternatives to testing. Conferences and interviews warrant serious consideration for students who perform poorly on the test. Presently, a few students make it into a supervisor’s office and may be interviewed. AUA should explore the potential benefits of alternative reassessment. If it is adapted as policy, it should be formalized and required, rather than as a whim of the supervisor on duty. Even better, include teachers in the process.

Provide more appealing options to students who fail the test. Rather than seeing test failures as the prerequisite for punitive action, it might be possible for AUA to see them as a marketing opportunity. Creative solutions should be considered. Currently, AUA’s
sole remedy for students who fail the Level 4 test is to send them back to Level 2. This causes students to lose face. Losing face in Thailand is simply unacceptable, and is avoided at all costs. This needs to be recognized at the institutional level and alternative programs need to be available. An obvious alternative is a remedial course. However, this is not practical in smaller branches, although it might be possible to offer a “special” course that is open to all lower-level students that would focus on the necessary skills to support Level 4 students. In larger branches, it might be possible to offer a short, intensive course during the interim week, followed by a retest.

AUA Rajdamri is very fortunate to have an excellent **self-access center** housed in its library (renamed the Student Learning Center, or SLC). In fact, a number of branches are trying to put together small libraries and self-study materials. Supporting greater learner autonomy is an obvious consideration for students who have failed the test. The SLC is AUA’s greatest current resource for these students. Ironically, AUA’s SCL closes quite early, at 5:45 PM, just as the busy time of day gets underway. The peak study time is 7 to 9 PM, and students for those classes start arriving around 5:30. AUA should make the SLC available to these students, working through the constraints that have prevented later hours (the biggest of which is the uncooperative and territorial nature of the library manager). Further, if it has not already done so, the SLC should develop study modules specifically designed for students who have failed the Level 4 test.

AUA could also prepare **home-study modules** for students who cannot study at the SLC. These could be given or loaned to students at the conclusion of interviews in the
academic office, or distributed by teachers. AUA could also develop self-study modules on its website that could be linked to guided support for students who fail tests. If properly designed, these and other opportunities could also be designed as effective marketing tools.

Finally, if AUA were to make meaningful study options available to students who fail the Level 4 test, it should expand its retest options. As we explored in the literature review, external variances can affect test performance, and some students may perform better if given a second chance. Students should have an opportunity to take a retest during the interim week between terms (also cited above). For any student who passes the retest, s/he would be able to register for the next level and rejoin his or her classmates.

**Design a database of student profiles, test and enrollment statistics. Update AUA’s computerized registration system.** The information presented in Chapter 2 was generated by hand, and certain aspects of the analysis were labor intensive and time consuming. Yet, useful information was also revealed after only a few hours of research. AUA has no procedures in place to establish a database of students, test results and enrollment trends. Without this, the institution is unable to clearly identify who its students are, how to contact them, what they want, and how they behave. It is no wonder that enrollment is declining, as we don’t even know if we are giving our customers what they want. (Our director claims that students don’t know what they want and it is our job to make that decision for them.)
A database designer should be hired to design a useful database and train support staff to maintain it. The resulting data should be thoroughly analyzed from both an academic perspective and a marketing perspective. Further, suggested modifications to the registration system should be submitted to the database designer and implemented, so that data retrieval becomes an easy, routine process, rather than an intimidating, time consuming chore.

**Re-record the listening section.** I conclude with the very thing that first sparked my interest in the Level 4 test and resulted in this analysis. In Chapter 2, I described the high number of incorrect answers in the listening section, in contrast to the many successful interviews I have conducted with students who failed the test, and in particular the listening section. I have also described my own experience as a proctor, unable to understand entire phrases of the listening CD, despite being a native speaker of English. As inconceivable as it sounds, for approximately ten years, AUA has knowingly administered the Level 4 test with a listening component that is unacceptably distorted. It goes without saying that this should be immediately corrected.

**CONCLUSION**

My final suggestion — to re-record the listening portion of the Level 4 test – is an achievable task, but it takes motivation, coordination, and more importantly, a genuine concern for the well being of our students. Through this, my other recommendations, and countless other improvements yet to be identified, AUA can significantly improve the
testing and overall learning experience of its students. The institution can benefit from these enhancements through better student performance, greater and more meaningful involvement of teachers, better performance from supervisors, greater cooperation between teachers and supervisors, a healthier working environment, a clearer vision for the institution, and quite possibly, higher enrollment and retention rates.

Yet, this will not happen unless management embraces the concept of inclusiveness, cooperation, and cross-cultural sensitivity between teachers, managers, and departments. Only by clarifying management’s goals, and by igniting a passionate respect for everyone involved can significant results be achieved.
Appendix

Table 3.1
Definition of the Research Period

<table>
<thead>
<tr>
<th>Term</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 5, 2006</td>
<td>Aug/Sep</td>
</tr>
<tr>
<td>Term 6, 2006</td>
<td>Sep/Oct</td>
</tr>
<tr>
<td>Term 7, 2006</td>
<td>Nov/Dec</td>
</tr>
<tr>
<td>Term 1, 2007</td>
<td>Jan/Feb</td>
</tr>
<tr>
<td>Term 2, 2007</td>
<td>Feb/Mar</td>
</tr>
<tr>
<td>Term 3, 2007</td>
<td>Apr/May</td>
</tr>
<tr>
<td>Term 4, 2007</td>
<td>June/July</td>
</tr>
</tbody>
</table>

*For retests, the Research Period includes Term 5, 2007, since that is when students who failed the Level 4 test in Term 4, 2007 would take the retest. Data from Term 5, 2006 is excluded, since those retests are for students who failed in Term 4, 2006.

Table 3.2
Information Used for this Analysis from AUA’s Computerized Registration System

<table>
<thead>
<tr>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student name</td>
</tr>
<tr>
<td>Student number</td>
</tr>
<tr>
<td>Terms studied</td>
</tr>
<tr>
<td>Sections studied</td>
</tr>
<tr>
<td>Hours absent</td>
</tr>
<tr>
<td>Final grade</td>
</tr>
<tr>
<td>Weak areas (for failures)</td>
</tr>
</tbody>
</table>
Table 3.3
Summary of Information Available from Answer Sheets, Grade Sheets and Retest Results

<table>
<thead>
<tr>
<th>Answer Sheets</th>
<th>Grade Sheets</th>
<th>Retest Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Name</td>
<td>Student Name</td>
<td>Student Name</td>
</tr>
<tr>
<td>Student Number</td>
<td>Student Number</td>
<td>Student Number</td>
</tr>
<tr>
<td>Section Number</td>
<td>Teacher Name</td>
<td>Current L2 Section</td>
</tr>
<tr>
<td>Test score</td>
<td>Section details (time, etc.)</td>
<td>Pass or Fail</td>
</tr>
<tr>
<td>Correct Answers</td>
<td>Student Grade</td>
<td>If fail, 2 problem areas</td>
</tr>
<tr>
<td>Incorrect Answers</td>
<td>(AA, A, BA, RA, RS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hours Absent</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.6
Students Who Failed the Level 4 Test and Took the Test the Following Term

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 5, 2006</td>
<td>28.6%</td>
</tr>
<tr>
<td>Term 6, 2006</td>
<td>47.6%</td>
</tr>
<tr>
<td>Term 7, 2006</td>
<td>21.7%</td>
</tr>
<tr>
<td>Term 1, 2007</td>
<td>22.2%</td>
</tr>
<tr>
<td>Term 2, 2007</td>
<td>10.3%</td>
</tr>
<tr>
<td>Term 3, 2007</td>
<td>33.3%</td>
</tr>
<tr>
<td>Term 4, 2007</td>
<td>57.1%</td>
</tr>
</tbody>
</table>
Table 3.7
Number of Failures by Section, Term 5, 2006
Number of Retests by Section, Term 6 2006

<table>
<thead>
<tr>
<th>Section</th>
<th>Failures</th>
<th>Registrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A302</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>A303</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>401</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MD401</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>MD402</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MD407</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TD404</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MD403</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>MD404</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>MD406</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TD401</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TD403</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TD405</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>TD406</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>35</strong></td>
<td><strong>10</strong></td>
</tr>
</tbody>
</table>

Table 3.8
Percentage of Students Taking the Level 4 Retest Who Passed in the First Attempt

<table>
<thead>
<tr>
<th>Term</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 5, 2006</td>
<td>70.0%</td>
</tr>
<tr>
<td>Term 6, 2006</td>
<td>70.0%</td>
</tr>
<tr>
<td>Term 7, 2006</td>
<td>40.0%</td>
</tr>
<tr>
<td>Term 1, 2007</td>
<td>50.0%</td>
</tr>
<tr>
<td>Term 2, 2007</td>
<td>66.6%</td>
</tr>
<tr>
<td>Term 3, 2007</td>
<td>28.6%</td>
</tr>
<tr>
<td>Term 4, 2007</td>
<td>75.0%</td>
</tr>
</tbody>
</table>
Table 3.9  
Level 4 Test Analysis, by Term  
Student Responses to Question #1 in the Listening Section  

**Term 5, 2007**

<table>
<thead>
<tr>
<th>Section</th>
<th>No. Taking Test</th>
<th>Incorrect Answer</th>
<th>Percent Incorrect</th>
<th>Form</th>
<th>Proctor</th>
<th>Teacher</th>
<th>Room Size: L / S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retest</td>
<td>13</td>
<td>11</td>
<td>84.6%</td>
<td>D</td>
<td>Unknown</td>
<td>Ginger</td>
<td>L</td>
</tr>
<tr>
<td>401</td>
<td>8</td>
<td>2</td>
<td>25%</td>
<td>G</td>
<td>Unknown</td>
<td>Pearsall</td>
<td>L</td>
</tr>
<tr>
<td>A302</td>
<td>19</td>
<td>17</td>
<td>89.5%</td>
<td>H</td>
<td>Unknown</td>
<td>Pearsall</td>
<td>L</td>
</tr>
<tr>
<td>MD401</td>
<td>9</td>
<td>9</td>
<td>100.0%</td>
<td>D</td>
<td>Unknown</td>
<td>Ginger</td>
<td>L</td>
</tr>
<tr>
<td>MD402</td>
<td>10</td>
<td>4</td>
<td>40.0%</td>
<td>E</td>
<td>Ham</td>
<td>Debbie P</td>
<td>L</td>
</tr>
<tr>
<td>MD403</td>
<td>9</td>
<td>7</td>
<td>77.8%</td>
<td>F</td>
<td>Unknown</td>
<td>McMurtry</td>
<td>L</td>
</tr>
<tr>
<td>MD404</td>
<td>22</td>
<td>7</td>
<td>31.8%</td>
<td>G</td>
<td>Ratnieks</td>
<td>Shelley</td>
<td>L</td>
</tr>
<tr>
<td>MD405</td>
<td>13</td>
<td>11</td>
<td>84.6%</td>
<td>A</td>
<td>Jones</td>
<td>Cowell</td>
<td>S</td>
</tr>
<tr>
<td>MD406</td>
<td>12</td>
<td>5</td>
<td>41.7%</td>
<td>B</td>
<td>Cornejo</td>
<td>Jones</td>
<td>L</td>
</tr>
<tr>
<td>MD407</td>
<td>12</td>
<td>9</td>
<td>75.0%</td>
<td>H</td>
<td>Cornejo</td>
<td>Ratnieks</td>
<td>S</td>
</tr>
<tr>
<td>TD401</td>
<td>12</td>
<td>8</td>
<td>66.7%</td>
<td>C</td>
<td>Unknown</td>
<td>Dujisik</td>
<td>L</td>
</tr>
<tr>
<td>TD402</td>
<td>13</td>
<td>9</td>
<td>69.2%</td>
<td>D</td>
<td>McMurtry</td>
<td>Shelley</td>
<td>L</td>
</tr>
<tr>
<td>TD403</td>
<td>16</td>
<td>6</td>
<td>37.5%</td>
<td>E</td>
<td>Cowell</td>
<td>Dupre</td>
<td>L</td>
</tr>
<tr>
<td>TD404</td>
<td>12</td>
<td>5</td>
<td>41.7%</td>
<td>F</td>
<td>Collins</td>
<td>Cowell</td>
<td>S</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>180</strong></td>
<td><strong>110</strong></td>
<td><strong>61.0%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Term 6, 2007**

<table>
<thead>
<tr>
<th>Section</th>
<th>No. Taking Test</th>
<th>Incorrect Answer</th>
<th>Percent Incorrect</th>
<th>Form</th>
<th>Proctor</th>
<th>Teacher</th>
<th>Room Size: L / S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retest</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>A302</td>
<td>11</td>
<td>4</td>
<td>36.4%</td>
<td>A</td>
<td>Crosthwaite</td>
<td>Pearsall</td>
<td>S</td>
</tr>
<tr>
<td>A303</td>
<td>7</td>
<td>4</td>
<td>57.1%</td>
<td>B</td>
<td>Pearsall</td>
<td>Thorpe</td>
<td>S</td>
</tr>
<tr>
<td>MD401</td>
<td>6</td>
<td>2</td>
<td>33.3%</td>
<td>G</td>
<td>Unknown</td>
<td>Debbie P</td>
<td>S</td>
</tr>
<tr>
<td>MD402</td>
<td>13</td>
<td>9</td>
<td>69.2%</td>
<td>H</td>
<td>Unknown</td>
<td>Upson</td>
<td>L</td>
</tr>
<tr>
<td>MD403</td>
<td>5</td>
<td>3</td>
<td>60.0%</td>
<td>A</td>
<td>Unknown</td>
<td>Cortan</td>
<td>S</td>
</tr>
<tr>
<td>MD404</td>
<td>23</td>
<td>8</td>
<td>34.8%</td>
<td>B</td>
<td>Raftery</td>
<td>Greynolds</td>
<td>L</td>
</tr>
<tr>
<td>TD401</td>
<td>10</td>
<td>8</td>
<td>80.0%</td>
<td>E</td>
<td>Debbie P</td>
<td>Raship</td>
<td>S</td>
</tr>
<tr>
<td>TD402</td>
<td>19</td>
<td>11</td>
<td>57.9%</td>
<td>F</td>
<td>Wilde</td>
<td>Manley</td>
<td>L</td>
</tr>
<tr>
<td>TD403</td>
<td>14</td>
<td>8</td>
<td>57.1%</td>
<td>G</td>
<td>Collins</td>
<td>Ebaugh</td>
<td>L</td>
</tr>
<tr>
<td>TD404</td>
<td>10</td>
<td>10</td>
<td>100.0%</td>
<td>H</td>
<td>Ebaugh</td>
<td>Richardson</td>
<td>L</td>
</tr>
<tr>
<td>TD405</td>
<td>11</td>
<td>3</td>
<td>27.3%</td>
<td>C</td>
<td>Raftery</td>
<td>McMurtry</td>
<td>S</td>
</tr>
<tr>
<td>TD406</td>
<td>15</td>
<td>12</td>
<td>80.0%</td>
<td>D</td>
<td>McMurtry</td>
<td>Jones</td>
<td>L</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>144</strong></td>
<td><strong>82</strong></td>
<td><strong>56.9%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

72
### Table 3.9 (continued)
#### Level 4 Test Analysis, by Term
Student Responses to Question #1 in the Listening Section

**Term 7, 2007**

<table>
<thead>
<tr>
<th>Section</th>
<th>No. Taking Test</th>
<th>Incorrect Answer</th>
<th>Percent Incorrect</th>
<th>Form</th>
<th>Proctor</th>
<th>Teacher</th>
<th>Room Size: L / S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retest</td>
<td>18</td>
<td>11</td>
<td>61.1%</td>
<td>B</td>
<td>Unknown</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>A301</td>
<td>9</td>
<td>6</td>
<td>66.7%</td>
<td>E</td>
<td>Savitri</td>
<td>Ham</td>
<td>S</td>
</tr>
<tr>
<td>A302</td>
<td>10</td>
<td>6</td>
<td>60.0%</td>
<td>F</td>
<td>Unknown</td>
<td>McBride</td>
<td>S</td>
</tr>
<tr>
<td>401</td>
<td>11</td>
<td>6</td>
<td>54.5%</td>
<td>D</td>
<td>Unknown</td>
<td>Filicetti</td>
<td>L</td>
</tr>
<tr>
<td>MD402</td>
<td>11</td>
<td>4</td>
<td>36.4%</td>
<td>C</td>
<td>Dupre</td>
<td>Thorneycroft</td>
<td>L</td>
</tr>
<tr>
<td>MD403</td>
<td>22</td>
<td>11</td>
<td>50.0%</td>
<td>D</td>
<td>Scholz</td>
<td>Glynn</td>
<td>S</td>
</tr>
<tr>
<td>MD404</td>
<td>17</td>
<td>11</td>
<td>64.7%</td>
<td>E</td>
<td>Cowell</td>
<td>White</td>
<td>L</td>
</tr>
<tr>
<td>MD405</td>
<td>17</td>
<td>7</td>
<td>41.2%</td>
<td>F</td>
<td>White</td>
<td>Oltmanns</td>
<td>L</td>
</tr>
<tr>
<td>TD401</td>
<td>16</td>
<td>6</td>
<td>37.5%</td>
<td>G</td>
<td>Walker</td>
<td>Ginger</td>
<td>L</td>
</tr>
<tr>
<td>TD403</td>
<td>17</td>
<td>9</td>
<td>52.9%</td>
<td>H</td>
<td>Collins</td>
<td>Jones</td>
<td>L</td>
</tr>
<tr>
<td>TD404</td>
<td>18</td>
<td>12</td>
<td>66.7%</td>
<td>B</td>
<td>Dupre</td>
<td>White</td>
<td>S</td>
</tr>
<tr>
<td>TD405</td>
<td>18</td>
<td>5</td>
<td>27.8%</td>
<td>C</td>
<td>White</td>
<td>Powell</td>
<td>L</td>
</tr>
<tr>
<td>TD406</td>
<td>16</td>
<td>15</td>
<td>93.8%</td>
<td>A</td>
<td>Jones</td>
<td>Glynn</td>
<td>S</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>109</td>
<td>54.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRAND TOTAL</td>
<td>524</td>
<td>301</td>
<td>57.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 3.10
#### Level 4 Test Analysis, by Term
Student Responses to Question #1 in the Listening Section
**Classes With 18 or More Students**

<table>
<thead>
<tr>
<th>Term</th>
<th>Section</th>
<th>No. Taking Test</th>
<th>Incorrect Answer</th>
<th>Percent Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 5</td>
<td>A302</td>
<td>19</td>
<td>17</td>
<td>89.5</td>
</tr>
<tr>
<td>Term 5</td>
<td>MD404</td>
<td>22</td>
<td>7</td>
<td>31.8</td>
</tr>
<tr>
<td>Term 6</td>
<td>MD404</td>
<td>23</td>
<td>8</td>
<td>34.8</td>
</tr>
<tr>
<td>Term 6</td>
<td>TD402</td>
<td>19</td>
<td>11</td>
<td>57.9</td>
</tr>
<tr>
<td>Term 7</td>
<td>MD403</td>
<td>22</td>
<td>11</td>
<td>50.0</td>
</tr>
<tr>
<td>Term 7</td>
<td>TD404</td>
<td>18</td>
<td>12</td>
<td>66.7</td>
</tr>
<tr>
<td>Term 7</td>
<td>TD405</td>
<td>18</td>
<td>5</td>
<td>27.8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>141</td>
<td>71</td>
<td>50.4</td>
</tr>
</tbody>
</table>
Table 3.11
Level 4 Test Analysis, by Term
Student Responses to Question #1 in the Listening Section
Sections with 75% or More Students Answering Incorrectly

<table>
<thead>
<tr>
<th>Term</th>
<th>Section</th>
<th>Percent Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 5</td>
<td>A302</td>
<td>89.5</td>
</tr>
<tr>
<td>Term 5</td>
<td>MD401</td>
<td>100.0</td>
</tr>
<tr>
<td>Term 5</td>
<td>MD403</td>
<td>77.8</td>
</tr>
<tr>
<td>Term 5</td>
<td>MD405</td>
<td>84.6</td>
</tr>
<tr>
<td>Term 5</td>
<td>MD407</td>
<td>75.0</td>
</tr>
<tr>
<td>Term 6</td>
<td>TD401</td>
<td>80.0</td>
</tr>
<tr>
<td>Term 6</td>
<td>TD404</td>
<td>100.0</td>
</tr>
<tr>
<td>Term 6</td>
<td>TD406</td>
<td>80.0</td>
</tr>
<tr>
<td>Term 7</td>
<td>TD406</td>
<td>93.8</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


