

# NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

## **Technical Documentation for the NAEP Study of First-Year Texas Postsecondary Students 2010 Pilot Test**

*(NAEP SDC Special Study Task 7)*



Prepared for:  
U.S. Department of Education  
National Center for Education Statistics  
1990 K Street, NW # 8005  
Washington, DC 20006

Prepared by:  
Westat  
1600 Research Boulevard  
Rockville, Maryland 20850-3129  
(301) 251-1500

OMB# 1850-0803 v.33

Contract# ED-07-CO-0079

# Table of Contents

<b><u>Chapter</u></b>		<b><u>Page</u></b>
	<b>Acknowledgements .....</b>	<b>V</b>
	<b>Executive Summary .....</b>	<b>VI</b>
<b>1</b>	<b>Introduction.....</b>	<b>1-1</b>
	1.1 Background.....	1-1
	1.2 Study Design and Research Questions.....	1-2
	1.3 Role of the THECB and Pilot Colleges .....	1-5
	1.4 Role of the NAEP Alliance Members .....	1-5
<b>2</b>	<b>Phase 0 Activities .....</b>	<b>2-1</b>
	2.1 Literature Review and Expert Panel Meeting.....	2-1
	2.2 Selection of Pilot Colleges .....	2-2
	2.3 Westat IRB .....	2-3
<b>3</b>	<b>Phase 1 Activities.....</b>	<b>3-1</b>
	3.1 Exploratory Interviews with Pilot Colleges .....	3-1
	3.1.1 Telephone Interview Protocol.....	3-2
	3.1.2 Summary of Interview Findings.....	3-4
	3.1.3 Verification Site Visits.....	3-5
<b>4</b>	<b>Phase 2 Activities .....</b>	<b>4-1</b>
	4.1 OMB Clearance and IRB Approvals .....	4-1
	4.1.1 OMB Clearance .....	4-1
	4.1.2 College IRB Approvals.....	4-1
	4.2 Sampling.....	4-3
	4.2.1 Sampling Plan.....	4-3
	4.2.2 Student Sampling Frame Data.....	4-5
	4.2.3 Student Sample Selection .....	4-7
	4.3 College Data Collection Plans and Support.....	4-7
	4.3.1 NAEP Best Practices Manual.....	4-8
	4.3.2 NAEP FYPS Data Collection Schedule .....	4-10
	4.3.3 Obtaining Contact Information for Sampled Students.....	4-11

## **Contents (continued)**

4.3.4	Student Notification of Selection .....	4-11
4.3.5	Thank-You Gifts Provided by Colleges.....	4-12
4.4	Manual for Conducting Assessment Activities .....	4-13
4.4.1	NAEP FYPS Materials .....	4-14
4.4.1.1	Conducting the Student Scheduling Call .....	4-15
4.4.1.2	The Student Record of Contacts.....	4-20
4.4.1.3	Assessment Accommodations .....	4-22
4.5	Field Staff Recruitment and Training .....	4-25
4.5.1	Supervisor and Assessment Administrator Recruiting.....	4-25
4.5.2	Training .....	4-26
4.5.2.1	Supervisor Training.....	4-26
4.5.2.2	Assessment Administrator Training.....	4-26
4.6	Data Collection.....	4-27
4.6.1	Data Collection Schedule for Colleges.....	4-27
4.6.2	Pre-Assessment Site Visit to Colleges .....	4-28
4.6.3	Student Recruitment and Scheduling .....	4-29
4.6.4	Evaluation and Revision of Procedures.....	4-30
4.6.5	Quality Control.....	4-31
4.6.6	Make-Up Data Collection .....	4-31
4.6.7	Field Staff Debriefing .....	4-32
4.6.8	Sample Cleaning.....	4-32
4.7	Response Rates.....	4-33
4.8	Non-Response Bias Analysis.....	4-34
4.8.1	Overall Response Rates .....	4-36
4.8.2	Response Rates by Gender .....	4-37
4.8.3	Response Rates by Developmental Status.....	4-39
4.8.4	Response Rates by Race/Ethnicity.....	4-39
4.8.5	Response Rates by Achievement Stratum .....	4-41
4.8.6	Conclusions .....	4-42
<b>5</b>	<b>Lessons Learned .....</b>	<b>5-1</b>
5.1	Working with the Colleges and Universities.....	5-1

5.2	Recruiting the Students .....	5-3
-----	-------------------------------	-----

## **Contents (continued)**

<b><u>Exhibit</u></b>		<b><u>Page</u></b>
2-1	Introductory Letter from THECB Commissioner R. Paredes to Selected Texas Colleges .....	2-4
3-1	Introductory E-Mail to Phase 1 Contacts at Pilot Colleges .....	3-3
4-1	Cover Letter to College Coordinators Introducing the NAEP Best Practices Manual .....	4-9
4-2	Student Thank-You Gifts Provided by Colleges .....	4-13
4-3	Script for Scheduling Calls to Sampled Students .....	4-16
4-4	Frequently Asked Questions (FAQs) .....	4-17
4-5	Record of Contacts (ROC) Form .....	4-21
4-6	Accommodation Tracking Form.....	4-24
4-7	College Data Collection Schedules .....	4-27
 <b><u>Table</u></b>		 <b><u>Page</u></b>
4-3-1	Missing Student Data by School and Stratification Variable .....	4-6
4-7-1	Data Collection Summary Report.....	4-34
4-8-1	Overall Response Rates by College.....	4-36
4-8-2	Response Rate Multiple Comparison Tests.....	4-37
4-8-3	Student Response Rates and Sample Distributions, by Gender .....	4-38
4-8-4	Student Sample Distributions, by Gender .....	4-38

## Contents (continued)

<b><u>Table</u></b>	<b><u>Page</u></b>
4-8-5      Response Rates of Eligible Students, by College and Gender .....	4-38
4-8-6      Student Response Rates, by Developmental Status .....	4-39
4-8-7      Sample Distributions, by Developmental Status .....	4-39
4-8-8      Student Response Rates, by Race/Ethnicity .....	4-40
4-8-9      Sample Distributions, by Race/Ethnicity .....	4-40
4-8-10     Student Response Rates, by Achievement Stratum .....	4-41
4-8-11     Sample Distributions, by Achievement Stratum .....	4-41
4-8-12     Academic Stratum Multiple Comparison Tests .....	4-42
 <b><u>Appendix</u></b>	
A            NAEP FYPS Sample Questions Booklet .....	A-1
Texas Preparedness Study Literature Review.....	A-19
Notes from April 20, 2010 Expert Panel Meeting .....	A-31
 B            Phase 1 Telephone Interview Guide .....	 B-1
 C            Sample Data Collection Plan for Pilot Colleges .....	 C-1
NAEP FYPS Supervisor Training Agenda.....	C-14
NAEP FYPS Assessment Administrator Training Agenda.....	C-17
NAEP FYPS Field Staff Debriefing Questionnaire.....	C-20

## Acknowledgements



(This page intentionally left blank.)

# Executive Summary

## Highlights

From September 13-October 15, 2010, nine public colleges and universities in Texas participated in a pilot study designed to evaluate the operational feasibility of administering the 12<sup>th</sup> grade NAEP reading and math assessments to incoming first-year postsecondary students. Following are several highlights and key findings from the pilot:

- The participating institutions generally provided excellent cooperation and support throughout the pilot, especially given the short lead time for planning the data collection. They provided student lists required for sampling, student contact information, active recruitment of students through the college administration, on-campus facilities for conducting assessments, and other assistance.
- Experienced and specially trained NAEP Supervisors and Assessment Administrators performed data collection activities for up to two weeks at each campus. They conducted a multi-mode recruitment effort using landline and cell phone numbers, e-mail, regular mail and text messaging. At each campus, 15 or more assessment sessions were scheduled over five days to give students flexibility on when to attend.
- No monetary or in-kind incentives for students were provided by the project. Most of the colleges did provide some form of small thank-you gift and/or refreshments to students who attended the assessment.
- Despite extensive and collaborative efforts to recruit students for the pilot, the student response rate was 20.7 percent overall. The response rate for specific schools ranged from 7.2 percent at one 2-year community college to 33.3 percent at one 4-year university.

Given the research objectives and sampling requirements for a large scale, statewide administration of the NAEP assessments to first-year postsecondary students in Texas, the NAEP Alliance has recommended against proceeding with the planned main study unless significant enhancements are made to support higher response rates. Chief among these would be the addition of cash incentives



for students who complete the assessments, longer lead times to prepare for data collection at each campus and a comprehensive information campaign to increase awareness of NAEP and generate interest in the study among the sampled postsecondary students.

## Background

The NAEP Study of First-Year Texas Postsecondary Students was designed to help policymakers better understand the academic links between high school and college as they relate to preparedness for postsecondary instruction, as measured by the NAEP 12<sup>th</sup> grade math and reading assessments. Specifically, this study addresses two key research questions: What points on the NAEP reading and mathematics scales represent knowledge and skills required for entry-level, credit-bearing postsecondary coursework? And, conversely, what points on the NAEP scales are aligned with developmental or remedial instruction?

Establishing such statistical relationships would support and enhance the ability of the Governing Board and NCES to report on high school graduates preparedness for postsecondary education. This study is one of several the Governing Board is currently pursuing to evaluate the preparedness of high school graduates for successful entry into college, the military and the labor force. In order to facilitate timely implementation of this study, in late 2009 the Texas Higher Education Coordinating Board (THECB) volunteered to help recruit two-year and four-year public colleges and universities in Texas to participate.

Ultimately, this research could require administering the NAEP 12<sup>th</sup> grade math and reading assessments to a sample of 6,000 to more than 20,000 first-year Texas college students, depending on the final research questions and analytical objectives. Given the study's potential cost and complexity, it was organized into four progressive phases of investigation designed to evaluate feasibility and guide the final research design:

**Phase 0:** Literature Review and Expert Panel Meeting;

**Phase 1:** Exploratory Interviews with Texas College Administrators;

**Phase 2:** Pilot Study (n = 600 first-year Texas postsecondary students);

**Phase 3:** Main Study (n = 6,000 to 20,000 students).

Because NAEP had never previously been administered in a postsecondary setting, a small pilot test of the administration methods was fielded in Phase 2 prior to embarking on Phase 3, the main study. The purpose of Phase 2 was to evaluate the operational feasibility of administering NAEP in Texas to a large and representative sample of first-year postsecondary students. A primary goal of the pilot was to determine what response rate could be achieved, both overall and for the individual colleges and key subgroups. Other goals included: exploring the ability/willingness of colleges to provide essential sampling and contact information about eligible students, their ability to successfully recruit sampled students to participate in NAEP, and their ability to provide necessary logistical support to administer the assessments on their campuses.

Input obtained from a literature review and Expert Panel recommendations in Phase 0, and from interviews with the pilot colleges themselves in Phase 1, guided the design of final methods employed in the pilot. However, there were two notable exceptions:

- The pilot did not employ monetary incentives to students, which were recommended in the literature and cited by the Expert Panel as necessary to achieve acceptable response rates. The reason was that the use of incentives would be inconsistent with standard NAEP administration methods. Replicating the standard NAEP procedures as closely as possible was deemed essential by the Governing Board and NCES to ensure comparability of results between the 12<sup>th</sup> grade and first-year postsecondary student samples.
- Assessments were not conducted during the colleges' summer orientation sessions, as suggested by some of the literature and Expert Panel members. The data collection contract was awarded in mid-June and OMB clearance was received in mid-August, making this timing for the assessments impractical. Also, most of the pilot colleges rejected the idea of administering NAEP during the orientation sessions due to time limitations. They explained that their orientation schedules were already overburdened with other activities.

Otherwise, the methods employed in the pilot aligned with the advice of the postsecondary assessment experts and the participating schools.

## Methods

### Sampling

Nine (9) colleges and universities were purposively selected for the pilot by the Texas Higher Education Coordinating Board (THECB). Selection criteria included 2-year and 4-year degree programs and diversity with respect to size, selectivity, region, and racial/ethnic composition. The pilot colleges and universities were grouped into three data collection windows as follows:

#### Group 1 (September 13-17, 2010)

- University of Texas at Brownsville & Texas Southmost College\*
- Texas A&M University
- West Texas A&M University

#### Group 2 (September 27-October 1, 2010)

- Austin Community College (Cypress Creek Campus)
- El Paso Community College (Valle Verde Campus)
- Lone Star Community College (Montgomery Campus)

#### Group 3 (October 4-8, 2010)

- Prairie View A&M University
- Tyler Junior College
- University of Texas at San Antonio\*

(\*Make-up sessions conducted week of October 11-15, 2010.)

Altogether, 1,332 students (148 per college) were initially selected for the pilot and recruited to participate. The starting sample was equally split between males and females and was further stratified based on developmental enrollment status, race/ethnicity and academic achievement. Eligible students were defined as those who completed high school any time from January-June of 2010, at a high school located anywhere in the U.S., and registered to attend classes during the fall semester at any of the nine pilot college campuses. Foreign, GED and home-schooled students were ineligible, as were online students who do not regularly attend classes on campus, and students who were no longer officially enrolled at the time of data collection. Following sample adjustments to account for these factors, the final eligible sample size was 1,234 first-year students across the nine pilot colleges.

## **Data Collection**

Recruitment of students to attend the assessment sessions was a joint effort of the colleges and the NAEP field staff. The college's recruited sampled students by sending them a personalized letter from the college president describing the importance of this research and the key role played by NAEP in formulating national education policy (see Appendix C.) Hard-copy letters were mailed to students about 10-14 days prior to the assessment sessions. Included with the letter was a schedule showing the times and locations of NAEP sessions to be held on campus, a map of the campus showing the assessment locations, and an insert describing any thank you gifts offered by the college for students completing the assessment.

Colleges then sent the president's letter and enclosures to students via e-mail about one week before the NAEP sessions. For data collection groups 2 and 3, the colleges also e-mailed students the link to the "YouTube" site where they could view the latest version of the NAEP promotional video for high school students.

On the Wednesday through Saturday of the week preceding NAEP sessions at each college, the NAEP field staff called and e-mailed selected students to schedule them for a convenient session, answer questions and respond to any objections. Contact was attempted for all students with a phone number and/or e-mail address provided by the college. Multiple attempts were made to contact all students for whom contact information was provided, on different days and at different times of day. Staff left voice mail messages for students they could not reach. (Students were not re-contacted following a firm refusal.)

Assessments were conducted at central locations on-campus in facilities provided by the college. Generally 10-12 regular sessions plus 4-5 drop-in sessions were scheduled each week at each campus, covering all days Monday-Friday and different hours of the day. (Additional make-up sessions were also held at two of the colleges the week of October 11-15.) During the data collection week, the field staff called and/or e-mailed students to remind them to attend their scheduled appointments. These calls were timed for the evening before the scheduled sessions, or early on the day of the session, as appropriate. Students who missed scheduled sessions were reminded that they could attend any session later in the week, including several "drop-in" sessions. At their own expense, seven of the nine pilot colleges offered thank-you gifts to students who attended the assessments. Two offered USB computer "memory sticks"; one offered a memory

stick plus 1.5 hours of community service credit; one provided a \$10 Target gift card plus a chance to win a 24 inch flat-screen TV; one offered a free lunch from the “Subway” sandwich chain; four provided some type of snack foods and beverages. Two colleges offered no gifts to students.

### **Special Challenges for Postsecondary Sampling and Data Collection**

In preparing to administer NAEP in the postsecondary setting, it was important to recognize several key differences between the high school and college environments for data collection that would clearly impact both procedures and results. These included:

***School Administration:*** The college setting is much less hierarchical and routine compared to the high school setting. College administrators have less control and knowledge of student activities and schedules throughout the day, and they cannot be as directive towards either the instructors or the students.

***Sample Frames:*** The contents and availability of student lists for sampling vary widely from college to college, much more than we see across secondary schools within the same state.

***Attendance Schedules:*** For high school students, daily attendance is generally mandatory and most students are at school at the same times Monday-Friday. College freshmen are not legally required to attend classes, have highly variable schedules and may be on campus less than five days per week.

***Proximity to Assessment Site:*** Colleges and universities typically cover larger physical areas than high schools, with buildings and classrooms often far away from parking lots and public transportation. This can create a much greater burden on college students in terms of just getting to the assessment session, as compared to high school students.

***Living Arrangements:*** The vast majority of high school seniors live at home with parents, while college freshmen may live on campus, at home with parents, or in private apartments and houses off-campus.

***Telephone Status:*** An increasingly large percentage of college students do not have landline telephones where they live and can only be contacted via cell phone, e-mail, or regular mail. Some colleges are sensitive about releasing student cell phone numbers, personal e-mail addresses and

other contact information. And, cell phone users are often more wary about answering calls from numbers they don't recognize, making it harder to contact them during recruitment.

All of these predicted factors impacted our data collection experience in the Phase 2 pilot. Although we had prepared to encounter these problems, they were still disruptive at times and ready solutions were not always available. For example, we were unable to contact large numbers of students due to old or disconnected phone numbers or inactive e-mail addresses. The colleges were generally quite helpful but could only provide corrected contact data if they had received it from students. We know that some of the recruitment letters signed by college presidents were sent to students' home addresses, not to more recent in-town or on-campus addresses. Mail delivery to dorms was not always timely, nor was the students' attention to their "snail-mail" boxes. These factors affected both our success in contacting and recruiting sampled students and our ability to provide effective follow-ups and reminders to ensure they attended the assessment sessions.

## **Key Findings and Recommendations**

The following points are the most important findings and recommendations emerging from the Phase 2 pilot study:

- The participating institutions provided excellent cooperation and support throughout the pilot, especially given the short lead time for planning the data collection (early August through early September). They provided student lists required for sampling, student contact information, active recruitment of students through the college administration, on-campus facilities for conducting assessments, and other vital assistance. (See Chapter 3 and Sections 4.2 and 4.3).
- Experienced and specially trained NAEP Supervisors and Assessment Administrators performed data collection activities for up to two weeks at each campus (See Sections 4.5 and 4.6). They conducted an intense multi-mode recruitment effort using landline and cell phone numbers, e-mail, regular mail and text messaging. At each campus, 15 or more assessment sessions were scheduled over five days to give students flexibility on when to attend. However, these intense data collection efforts were not enough to overcome other priorities and time pressures faced by the sampled students (which in many cases included work and child care responsibilities in addition to schoolwork), the difficulties involved in

- contacting students, and/or their general lack of interest in the study. Even among students who agreed to participate, only about half attended the NAEP assessment sessions.
- No cash incentives were provided. The research design for this study required matching the standard NAEP 12<sup>th</sup> grade administration procedures, which do not employ student incentives. Therefore, the Governing Board and NCES directed that student incentives should not be provided as part of the postsecondary NAEP research process. (Note that most of the pilot colleges, at their own expense, did provide some form of small thank-you gift and/or refreshments to students who attended the assessment.)
- Despite extensive and collaborative efforts to recruit students, the combined student response rate across the nine pilot colleges was only 20.7%. The low response rate appeared to be caused by a combination of factors, but students' busy and varied schedules, the low salience or importance of the study for this population, the lack of lead time to promote awareness of NAEP, and the absence of cash incentives, all probably contributed.
- The response rate at individual campuses ranged from 7.2% at one 2-year community college to 33.3% at one 4-year university (see Section 4.8). In general, the 4-year universities exhibited a higher response rate than the 2-year community colleges. However, this was not a consistent pattern at the individual college level and one 2-year community college had the second-highest response rate at 33.0%, while one 4-year university had a response rate of 15.7%.
- A nonresponse bias analysis found that males and females and developmental and non-developmental students responded at similar levels to the assessment. Students at two year colleges had lower response rates overall than those at 4 year colleges, as did students with lower SAT/ACT scores compared to students with higher scores. While Hispanics showed higher absolute response rates than the non-Hispanic White and Black/African-American samples, the difference was not statistically significant.

Given the research objectives and sampling requirements for a large scale, statewide administration of the NAEP assessments to first-year postsecondary students in Texas, the NAEP Alliance has recommended against proceeding with the planned main study unless significant enhancements are made to support higher response rates. Chief among these would be the addition of cash incentives for students who complete the assessments, longer lead times to prepare for data collection at each

campus and a comprehensive information campaign to increase awareness of NAEP and generate interest in the study among the sampled postsecondary students.

We should note that the pilot also generated a number of positive findings about the design and objectives of the pilot. For example, relatively few students actively refused to participate. We found no evidence of *negative* student attitudes toward NAEP or the assessment process, per se. None of the contacted students requested accommodations or complained about the administration methods. While this does not indicate there is no need for accommodations among the postsecondary students, this finding does suggest that this issue was not a factor in the response rate results.

Also, none of the contacted students commented on the study's access to academic records, which was mentioned in the recruitment letter from college presidents. The NAEP supervisors and assessment administrators were specifically trained and instructed to actively listen for any student comments on this topic, as well as the accommodations issue, and carefully record any such comments. None were reported.

And finally, the degree of cooperation and engagement in the study provided by the pilot colleges and universities was exceptional. In part this was due to the high-level support from the THECB, but it was clear that the participating administrators had a strong interest in this research and the goal of better understanding the academic linkages between high school and college. The issue of academic preparedness among first-year postsecondary students was a high priority at all of the pilot colleges.



## 1.1 Background

The congressionally authorized National Assessment of Educational Progress (NAEP) is the only continuing source of comparable national and state data available to the public on the achievement of students at grades 4, 8, and 12 in core subjects. The National Assessment Governing Board (Governing Board) oversees and sets policy for NAEP. NAEP and the Governing Board are authorized under the National Assessment of Educational Progress Authorization Act (P.L.107-279).

Among the Governing Board’s responsibilities is “to improve the form, content, use, and reporting of [NAEP results].” Toward this end, the Governing Board established a national commission to make recommendations to improve the assessment and reporting of NAEP at the 12th grade. In its March 2004 report<sup>1</sup>, the commission noted the importance of maintaining NAEP at the 12th grade as a measure of the “output” of K-12 education in the United States and as an indicator of the nation’s human capital potential. The commission recommended that 12th grade NAEP be redesigned to report on the academic preparedness of 12th grade students in reading and mathematics for entry- level college credit coursework. The commission concluded that having this information is essential for the economic well being and security of the United States and that NAEP is uniquely positioned to provide such information.

As the Governing Board has been developing ways to implement the commission’s recommendations, there has been a wider recognition—among federal and state policymakers, educators, and the business community—of the importance of a rigorous high school program that results in meaningful high school diplomas and prepares students for college and for job training. The Administration has set the goal of ensuring that every high school graduate is college- and career- ready. Enabling NAEP to report on 12th grade preparedness would provide an indicator that can be used to monitor this goal.

---

<sup>1</sup> See [http://www.nagb.org/publications/12\\_gr\\_commission\\_rpt.pdf](http://www.nagb.org/publications/12_gr_commission_rpt.pdf).

As part of implementing the commission's recommendations, the Governing Board planned a program of research studies to support the validity of statements about 12th grade student preparedness in reading and mathematics.<sup>2</sup> Among the studies planned was a proposed study of first-year postsecondary student performance on the NAEP mathematics and reading assessments.

The data resulting from this study could be used, along with the results of the other planned studies, to help develop valid statements that can be made about the preparedness of 12th grade students in NAEP reports. While other studies, such as NCES high school longitudinal studies (e.g. NEELS, ELS, HSLS), provide information relating achievement on assessments and high school grades (assessments developed specifically for the study as well as AP Exams, ACT, and SAT scores) to college placement and success, the NAEP Study of First-Year Texas Postsecondary Students could provide valuable empirical linkages between NAEP achievement and college placement, which has not previously been examined.

However, the Governing Board and NCES also recognized that administering the NAEP assessments to postsecondary students could involve special challenges. These included: the willingness and ability of colleges and universities to participate and support the NAEP data collection; the various logistical hurdles such as obtaining appropriate sample frames for students, obtaining student contact information, contacting and recruiting students, obtaining appropriate space to conduct the assessments on campus; and of course, student response rates. Therefore, a comprehensive pilot study was viewed as essential before proceeding with the full-scale study. This report focuses primarily on results from the formal pilot study conducted in Phase 2 of this project.

## **1.2 Study Design and Research Questions**

The overall study design for the NAEP Study of First-Year Texas Postsecondary Students (or NAEP FYPS) encompasses four progressive phases of research, as follows:

- Phase 0: Literature Review and Expert Panel Meeting (Spring 2010)
- Phase 1: Exploratory telephone discussions with designated contact persons at nine Texas colleges and universities identified by the Commissioner of the THECB (Summer 2010)
- Phase 2: A small-scale pilot study of the procedures for and feasibility of administering the

---

<sup>2</sup> The full scope of the Governing Board's research agenda can be found on the Governing Board's website at <http://www.nagb.org/publications/PreparednessFinalReport.pdf>.

## NAEP assessments on campus (Fall 2010)

- Phase 3 (if Phase 2 successful): A full-scale, representative sample of postsecondary institutions randomly selected statewide in Texas for the fall 2011 semester.

This report focuses primarily on the implementation of the Phase 2 pilot study, but also briefly covers Phases 0-1.

The pilot study involved administering assessments to a planned sample of 600 incoming first-year students at nine Texas colleges and universities. The pilot institutions were purposively selected to represent different schools based on size, type (2-year and 4-year degree programs), geographic region within Texas, academic selectivity, and ethnic composition.

The main focus of the pilot was the operational experience of conducting NAEP assessments in the postsecondary setting, and it will evaluate aspects such as success with sampling, data collection, and response rates.

The pilot was intended to provide important insights about the operational feasibility of conducting NAEP assessments with first-year postsecondary students, including the willingness of students to participate (without monetary incentives) and the availability of records for creating sampling frames and collecting academic and course enrollment data on sampled students. Thus, the pilot would inform whether standard NAEP data collection methods could be successfully executed in the postsecondary setting.

The main research questions investigated in the pilot were:

- Based on the student participation in the pilot study, is a full-scale study feasible?
- What were the practical aspects of working with participating postsecondary institutions?
- How did the sampling and data collection activities vary across institutions?

Prior to performing a full-scale study of the relationship between NAEP performance and developmental/remedial versus credit-bearing placement in Texas postsecondary institutions, it was important that a pilot study be conducted to examine practical aspects of the planning and administration process, operational and logistical issues involved in the data collection, student response rates to the assessment, and the ability of the participating institutions to provide requested academic and demographic data for the study.

For this pilot, students were given two blocks of reading or mathematics cognitive questions followed by a short student survey consisting of four questions. The blocks of reading and mathematics cognitive questions were intact and unchanged from those that were administered to students across the nation at grade 12 in the winter of 2009, as part of the main NAEP assessment. The four student survey questions, about student's race/ethnicity and parents' highest education level, were a subset of the survey questions that were administered at grade 12 as part of the main NAEP assessment in 2009.

The **NAEP FYPS Sample Question Booklet** containing sample questions and selected responses for mathematics and reading, as well as the student survey questions, can be found in **Appendix A**.

Given that a primary goal of the Phase 3 main study was to determine how first-year postsecondary students perform on the 12th-grade NAEP reading and mathematics assessments, the standard NAEP 12th grade administration procedures needed to be followed as closely as possible in the postsecondary administration sessions. This meant using: the same assessment instruments with the fewest possible changes; trained NAEP supervisors (SVs) and Assessment Administrators (AAs); the same administration procedures and instructions for students; the same time limits for the assessment sessions; similar physical settings for the assessments; and so on. In order for the assessment results to be comparable across the 12th grade and first-year postsecondary student samples, the administration methods used in the pilot were kept as close as possible to standard NAEP procedures.

However, the same standard NAEP conditions could not be employed for the tasks of sampling and recruiting students. There are many important differences between the high school and postsecondary settings that impact how the data collection effort must be conducted at the postsecondary level. These differences include:

- **Attendance Schedules:** For high school students, daily attendance is generally mandatory and most students are at school at the same times Monday through Friday. College students are not legally required to attend classes, have highly variable schedules, and may be on campus fewer than five days per week. Many also work before or after classes and have limited free time to spend on other activities on campus.
- **Proximity to Assessment Site:** Colleges and universities typically cover larger areas than high schools, with buildings and classrooms often far away from parking lots and public transportation. This creates a bigger hurdle for postsecondary students in terms of getting to

the assessment session, as compared to high school students, and in many cases could become a “deal breaker” in terms of participation.

- **Living Arrangements:** The vast majority of high school seniors live at home with parents, while first-year postsecondary students may live on campus, at home with parents, or in private apartments and houses off-campus. This makes it more difficult to design contact and recruitment methods that will work effectively for students living in different situations. For the first-year students, many are living away from parents for the first time, class schedules and study responsibilities are new, and remembering secondary activities without reminders may be a problem.

We tried to account for these factors as much as possible in developing the methods we employed for the data collection effort.

### **1.3 Role of the THECB and Pilot Colleges**

To enable more rapid organization and execution of the study, the Texas Higher Education Coordinating Board (THECB) offered in late 2009 to assist the National Assessment Governing Board and NCES in conducting this study at public colleges and universities in Texas. In March of 2010, the Texas Commissioner of Higher Education, Raymund Paredes, sent a letter to the Presidents of 12 Texas 2- and 4-year public institutions asking for their support and cooperation in conducting the Phase 2 pilot. This request resulted in quick agreement from nine schools to participate in late March and April, making it possible to start planning the Fall 2010 data collection in June through August. Without the active support of the THECB it simply would not have been possible to field this study in 2010, and the earliest data collection window for the pilot would have been delayed to Fall 2011, at the earliest.

### **1.4 Role of the NAEP Alliance Members**

Following are the specific roles and activities performed by the NAEP Alliance members in conducting this project:

- **Educational Testing Service (ETS)** was responsible for identifying and convening the Expert Panel, preparing the OMB clearance submission (including the analysis plans for assessment results), revising the NAEP 12<sup>th</sup> Grade Background Questionnaire for use in the

pilot, and providing general coordination among the Alliance members working on the project.

- **Westat** was responsible for designing and implementing the Phase 1 interviews with pilot colleges, obtaining college IRB approvals, preparing field instruments and manuals, hiring and training the field staff, planning and conducting all Phase 2 sampling and data collection tasks, performing the nonresponse bias analyses, and preparing this report.
- **Pearson** was responsible for preparing and shipping the assessment booklets and related administration materials used in Phase 2, and scoring and processing the completed NAEP assessment booklets.

### 2.1 Literature Review and Expert Panel Meeting

At the January 2010 meeting of the NAEP Quality Assurance Technical Panel (QATP), a subcommittee of the QATP members reviewed the preliminary research design and data collection plans for the NAEP FYPS study. In commenting on the initial plans for conducting Phases 1 through 3 as described above, the QATP recommended adding a more exploratory “Phase 0” to the project to include a literature review and formation of a panel of subject matter experts to advise the project. The goal was to strengthen the theoretical basis for the research design and to ensure that the project took full advantage of previous research on postsecondary assessments and relevant developmental instruction/readiness issues. The Governing Board and NCES accepted the QATP recommendations.

During Phase 0, the NAEP Education Statistics Services Institute (NESSI), prepared a literature review of other studies involving surveys of postsecondary students, which can be found in **Appendix A**.

ETS worked with NCES in March and early April to form an expert panel, which convened on April 20, 2010, in Washington, D.C.

The panel members in attendance were as follows:

- Maria Teresa Tatto, Associate Professor, Michigan State University
- Jennifer Sharp Wine, Project Director, RTI International
- David Gardner, Deputy Commissioner for Academic Planning and Policy, Texas Higher Education Coordinating Board (THECB)
- Geraldine Mooney, Vice President, Surveys and Information Services, Mathematica

In addition to the panel members, numerous representatives from NCES, the Governing Board, NESSI, Westat, and ETS were in attendance:

- Peggy Carr, NCES
- Tom Weko, NCES
- Ray Fields, NAGB
- Bill Ward, NCES
- Brenda Wolf, NCES
- Andy Kolstad, NCES
- Holly Spurlock, NCES
- Pat Etienne, NCES
- Amy Yamashiro, NESSI
- Bob Patchen, Westat
- Lou Rizzo, Westat
- Jay Campbell, ETS
- Robert Finnegan, ETS

Ray Fields of the Governing Board provided the panel with an overview of ongoing preparedness research underway or planned. Bill Ward of NCES presented a summary of the Texas pilot study and the goals for the meeting. In addition to the overviews from the Governing Board and NCES, Amy Yamashiro of NESSI presented a summary of the literature review on postsecondary survey practices.

The panel discussed a variety of issues regarding the design, sampling, data collection, recruitment of students, and feasibility criteria for the study. The memo from ETS to the National Center for Education Statistics entitled: Notes from April 20, 2010, Texas Pilot Expert Panel Meeting, presented in **Appendix A**, discusses each of these issues along with the panel's recommendations.

## **2.2 Selection of Pilot Colleges**

As noted earlier, the Texas Higher Education Coordinating Board (THECB) offered assistance by inviting a diverse mix of public colleges and universities to participate in the study -- based on two-year and four-year degree status, size (total enrollment), region of the state, race/ethnicity



composition of students, and selectivity. The letter from Commissioner Paredes to University/College Presidents requesting their participation is included as Exhibit 2-1 on the next page.

The nine colleges and universities that participated in the NAEP FYPS are listed below:

- Austin Community College, Austin, TX
- El Paso Community College (Valle Verde Campus), El Paso, TX
- Lone Star College System (Montgomery Campus), Conroe, TX
- Prairie View A&M University, Prairie View, TX
- Texas A&M University, College Station, TX
- Tyler Junior College, Tyler, TX
- University of Texas at Brownsville and Texas Southmost College, Brownsville, TX
- University of Texas at San Antonio, San Antonio, TX
- West Texas A&M University, Canyon, TX

## **2.3 Westat IRB**

In early June, 2010, the project submitted NAEP FYPS materials for review by Westat's IRB. Westat's IRB concluded that the NAEP FYPS activities were covered under that of main NAEP, and issued the project a memo providing an exemption from full IRB review.

**Exhibit 2-1: Introductory Letter from THECB Commissioner R. Paredes to Selected Texas Colleges**

Texas has been asked by the National Assessment Governing Board (Governing Board) to participate in a small pilot study that will lead to better information about the academic preparedness of 12<sup>th</sup> grade students for postsecondary education. The purpose of the pilot is to determine the feasibility of administering the 12<sup>th</sup> grade reading and mathematics tests of the National Assessment of Educational Progress (NAEP) to a sample of incoming first-year college students, to better understand the academic links between high school and college.

NAEP, as you know, is the only nationally representative source of data on student achievement. It is administered across the country to a large sample of students in grades 4, 8 and 12 and is regarded as the Nation's Report Card. The Governing Board oversees and sets policy for NAEP; the National Center for Education Statistics administers NAEP.

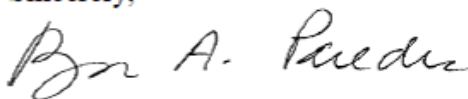
We believe that a successful pilot, followed in 2011 by a full-scale study, could provide us all with valuable empirical linkages between high school performance and new data on the necessary knowledge and skills needed to enter credit bearing, post-secondary education coursework.

The pilot is scheduled for the summer/fall of 2010. It will be conducted in Texas at up to 12 colleges, representing a mix of 2-year and 4-year institutions. Only 75 students will be needed per institution. An external contractor will be responsible for administering the tests. However, the contractor will need on-campus assistance in assessing feasibility, identifying and recruiting students, and in providing a location for the testing and other logistical support. The contractor will reimburse you for the administrative costs associated with your support.

Following NAEP guidelines, all information will be kept confidential. No individual student scores will be reported, nor will institution-specific scores be calculated or reported. This study has the full support of the Texas Higher Education Coordinating Board and your participation in the pilot study will be greatly appreciated.

If your institution is willing to participate in this study, please send me the name and contact information of an individual that can assist the contractor by Friday, March 26. If you have questions about the study please contact Ray Fields at [ray.fields@ed.gov](mailto:ray.fields@ed.gov) or 202-357-0395. Thank you for your support.

Sincerely,



Raymund A. Paredes

c: Ray Fields

## Phase 1 Activities

# 3

Phase 1 of this study was conducted in June, 2010 and involved gathering information from the nine Texas colleges and universities about the practical aspects of administering assessments with first-year (freshman) students on their campus, including:

- The number of first-year (freshman) students
- The capability (and timing) of conducting the NAEP assessments during freshman orientation or some other appropriate time and venue.
- The willingness of universities and colleges to participate, which included:
  - Providing lists of sampled students, including demographic information, from which to draw a representative sample; and
  - Assisting in identifying and recruiting students.

### 3.1 Exploratory Interviews with Pilot Colleges

Westat senior project staff completed Phase 1 exploratory telephone interviews with senior administration representatives at all nine participating Texas colleges during the week of June 21-25, 2010. The purpose of the interviews was to collect information about the willingness and capability of the participating colleges to provide the support needed to implement the Phase 2 pilot.

Prior to the start of Phase 1 interviews, the Commissioner of the Texas Higher Education Coordinating Board (THECB) asked the presidents of the participating colleges to designate a primary contact person to coordinate the NAEP pilot study sampling and data collection plans with Westat. The contact persons held a variety of positions at their respective colleges, but nearly all were executive-level staff in the provost's office (provosts or associate provosts), student affairs executives, research and evaluation directors or senior faculty members.

Interviews were conducted by senior Westat project staff familiar with NAEP and trained on the research design for the Phase 2 pilot study and the specific goals of the Phase 1 interviews.

A few days before calling began, each Westat interviewer sent the following introductory e-mail to the contact person at their assigned colleges, as shown in Exhibit 3-1, on the next page.

The purpose of the e-mail was to alert the contact to the upcoming call and also inform them about the general topics to be covered. After receiving the introductory e-mail, several of the contacts decided to invite other college staff to join the call, either to hear the discussion to be more informed about the goals of the study and the plans for data collection, or to help answer interview questions.

### **3.1.1 Telephone Interview Protocol**

During the calls, a formal interview guide was used to collect details on the following major topics:

- Availability of sampling lists;
- Ability to identify students enrolled in developmental courses;
- Appropriateness of the fall 2010 data collection schedule for the Phase 2 pilot study;
- Ability of NAEP staff to work on campus;
- Ability of college staff to assist in contacting and recruiting sampled students;
- Ability of the college to provide appropriate assessment locations;
- Ability of the college to provide academic records for sampled students; and,
- Ability of the college staff to assist with Institutional Review Board (IRB) clearance.

The interviews averaged about 60 minutes in duration. A copy of the Interview Protocol Guide can be found in **Appendix B**.

### **Exhibit 3-1: Introductory E-mail to Phase 1 Contacts at Pilot Colleges**

■ Thank you for agreeing to serve as the contact person for (name of college) as we prepare for the **National Assessment of Educational Progress (NAEP) Pilot Study of First Year College Students**. This important study is being conducted by the National Center for Education Statistics (NCES), U.S. Department of Education in conjunction with the Texas Higher Education Coordinating Board (THECB) and the National Assessment Governing Board (Governing Board), during the fall 2010 semester. Westat, a national research company working with NCES, will perform the sampling and data collection activities for the study.

The purpose of this pilot study is to determine the feasibility of administering the 12<sup>th</sup> grade reading and mathematics tests of NAEP to a representative sample of incoming first year college students. The overall goal is to help educators, researchers and policymakers better understand the level of preparedness among high school graduates for postsecondary instruction. Our immediate focus is on working with you to plan for a successful on-campus data collection effort this fall.

**I am writing to let you know that I will be calling your office on Monday, June 21, starting at 10:00 AM Central Time, to talk with you about our plans for administering the NAEP assessment at your college and to gather relevant background information.** The call should take approximately one hour and will cover the following topics:

**Availability of student lists for sampling:** We will need to determine the best way to assemble lists of incoming first-year students this fall who completed high school anywhere in the U.S. in the spring of 2010, and will be 18 or older by 8/31/10.

**Plans for conducting the NAEP assessments on campus:** We need to discuss the best times and locations for conducting the assessments on your campus, with the goal of making the process easy and convenient for students to participate.

**Maximizing student cooperation and response rates:** We will need your help to plan ways to contact the sampled students and recruit them to participate in this study. This may include gaining access to contact information such as names, addresses, telephone numbers (regular and mobile), e-mail addresses, etc.

**Institutional research policies and procedures:** We will need to discuss review and approval procedures for conducting research with students at your college, such as any Institutional Review Board (IRB) requirements.

**Involving appropriate faculty and college staff:** Achieving high student response rates in this study will require a team effort. Our hope is to work closely with you and other interested college faculty and staff to effectively communicate the importance of this study and encourage selected students to participate.

Westat has been conducting the NAEP sampling and data collection tasks nationwide since 1983. Our goal is to work closely with you and others at your college to ensure a smooth and successful pilot study.

Thanks again for your help, and I look forward to speaking with you.

### 3.1.2 Summary of Interview Findings

At the time of the Phase 1 interviews, all of the contacts were aware of the proposed NAEP pilot study and seemed highly engaged and cooperative. Several of the colleges invited additional key administrative staff to join the call to assist in providing information. All of the contacts indicated top-level support for the pilot from senior management and a strong interest in research projects addressing academic preparedness among first-year students.

The interview results were both consistently positive and notably uniform with respect to the feasibility of conducting the NAEP assessments on campus at these colleges.

When interviewed about the NAEP Pilot sampling and data collection plans, contact persons at all nine participating postsecondary institutions confirmed the following:

- Lists of first-year students covering virtually all of the eligible students<sup>3</sup> would be available for selecting a random, representative sample prior to the planned start of data collection in September.
- Sampling lists could identify students enrolled in developmental/remedial classes.
- The planned data collection window of September through October, 2010, was approved by all contacts.
- Westat field staff could review student lists for sampling purposes, help conduct recruitment of sampled students and administer the NAEP assessments on campus.
- The colleges could assist in student recruitment and would provide access to appropriate locations on campus to conduct the assessments.
- All of the contacts confirmed that student academic records and demographic data could be provided consistent with the requirements of the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. 1232g).

---

<sup>3</sup> Because colleges finalize their enrollment during the last few weeks of summer, and some students may drop out or be added, the lists were expected to be complete enough for sampling, but not yet 100% complete.

- All of the colleges agreed to assist Westat in understanding their specific IRB requirements and the timelines for submitting all necessary research application forms.

The Phase 1 interviews did not uncover any major logistical problems or concerns regarding Phase 2 data collection. However, the issues identified as likely to be the most problematic and require the most attention going forward were the Institutional Review Board (IRB) procedures and the difficulty of gaining student cooperation and achieving a high response rate.

While numerous specific plans and schedules needed to be resolved at each college, all of the Phase 1 contacts indicated they would assist in this process and involve other college staff as needed to support project activities.

### **3.1.3 Verification Site Visits**

To confirm the information collected from the Phase 1 telephone interviews and gain more first-hand knowledge of the on-campus conditions for conducting NAEP assessments, follow-up site visits were conducted by the Westat Project Director to the University of Texas at Brownsville / Texas Southmost College and Texas A & M University in July. The visits included inspections of proposed assessment sites, campus layout and conditions, student enrollment lists and telephone directories, campus e-mail systems and bulletin boards, and other aspects of the campus infrastructure that could have impacted pilot data collection outcomes.

Along with the Phase 1 verification site visits, Westat project staff conducted periodic telephone planning discussions with all of the school contacts throughout July and August to focus on IRB applications, student sampling, and other aspects of the campus infrastructure which could impact Phase 2 data collection outcomes.

During this period, schools with multiple campuses were requested to select one campus for sampling and participation in the pilot. Schools were also encouraged to consider providing a token of appreciation or other small thank-you gift to participating students.

## Phase 2 Activities

# 4

Phase 2 activities involved administering assessments to a target sample of 600 eligible, incoming first-year students at the nine Texas pilot colleges and universities. The main focus of Phase 2 was the operational experience in conducting NAEP assessments in the post-secondary setting, including the degree of success achieved with sampling, data collection, response rates, and other operational and logistical factors.

Phase 2 began with the submission of an OMB package requesting clearance for the data collection activities, as well as the preparation of individual IRB application packages for eight of the nine schools. These activities were followed by the creation of sampling and data collection plans for each school, field staff recruitment and training, and data collection.

### 4.1 OMB Clearance and IRB Approvals

#### 4.1.1 OMB Clearance

An Office of Management and Budget (OMB) Package was drafted and submitted for agency approval on July 14, 2010, and forwarded by the Agency to the OMB for review.

OMB clearance for this study was received on August 2, 2010; the OMB approval number was OMB# 1850-0803 v.33 with an expiration date of September 13, 2013.

#### 4.1.2 College IRB Approvals

Given that data collection was scheduled to begin in early September and conclude by mid-October, receiving Institutional Review Board (IRB) approval from each of the nine pilot colleges was a key step in the process as no sampling of students could occur until such approval was received. Timing was critical as some of the colleges held formal monthly IRB meetings to consider research applications received and all application materials had to be submitted by their deadline to be discussed at that month's IRB meeting. Missing the deadline for September's IRB meeting would



mean delay of discussion and possible approval of the application until October's meeting and therefore would preclude data collection at that college.

The IRB approval process for each pilot college began by searching each school's web site for information about its IRB process, including the application materials to complete and by identifying the school's requirement for human subject protection training, if specified. For each college, an IRB chair or contact person was identified and contacted by telephone or email so that the study could be introduced to that individual. A discussion of the process and deadlines for IRB approval and how it could be accomplished given the tight time schedule of the study followed.

Each of the nine pilot colleges did have an Institutional Review Board in place and all but one required submission of an IRB application package detailing the study's goals, protocols, informed consent procedures, benefits and risks. That one college's IRB administrator, based on a telephone conversation with the Project Director about the pilot, decided that the study was exempt from IRB review, and a full application package was not necessary. The IRB application packages were all different and ranged in length from two to 23 pages. The time to read and complete each package ranged from two to four hours.

In general, each of the IRB application packages covered the same information, including:

- Purpose and background of the study
- Subject population
  - Number in sample
  - Any vulnerable populations, such as pregnant women or minors
- Duration of the project
- Recruitment procedures
- Informed consent procedures
- Risks to participants
- Benefits to participants and
- Confidentiality protections

About half of the colleges specified certain human subjects protection training courses that key study personnel were required to have completed; these schools also required submission of a certificate of completion for these training courses. Two colleges required that the Principal Investigator and key personnel complete the Collaborative Institutional Training Initiative (CITI) Human Subjects Protection course which took approximately eight hours to complete. Two other

colleges required the CITI course or another similar course such as the National Institutes of Health (NIH) Human Subjects Protection course be completed by the Principal Investigator and key personnel; the NIH Human Subject Protection Course took approximately two hours to complete.

During the IRB review process, several colleges contacted Westat with specific questions or requests, including:

- Obtaining university counsel approval of the NCES FERPA exemption in order to release student academic data;
- Requesting to view the actual NAEP Reading and Mathematics Assessment Booklets being used for FYPS;
- Requesting a copy of Westat's internal IRB approval for NAEP.

In each case, Westat worked directly the college or university to respond to their requests for information.

Five of the colleges required IRB closeout forms which were to be completed and submitted at the end of the research project. Those forms requested information about the number of subjects initiated into the study, subjects' benefits, adverse reactions, and withdrawals from the study and the results of the research to date.

## **4.2 Sampling**

### **4.2.1 Sampling Plan**

The schools included in the sample for Phase 2 were purposively selected by the Commissioner of the Texas Higher Education Coordinating Board to represent a diverse mix of colleges based on two-year and four-year degree status, size (total enrollment), region of the state, race/ethnicity composition of students, and selectivity.

The eligible students in the nine schools listed in Section 2.2 were viewed as the universe. The first stage of sample selection then was at the student level within these schools. Students were eligible for the Texas FYCS if they met the following criteria:

- They completed high school sometime during the period January-June 2010 and were entering college for the first time in the fall 2010 semester;
- They were 18 years or older as of September 1, 2010 (i.e., their birth month was August 1992 or earlier);
- The place they completed high school was a high school in the United States which would be eligible for twelfth grade national NAEP.

The target sample size for completed assessments was 67 eligible students per school (600 total over the nine schools). The nominal (starting) sample size was 148, with 74 women and 74 men being sampled at each school. The total nominal sample size across the nine schools was 1,332 (148 times 9 schools). A stratified sample of students was drawn.

The first level of stratification (males and females) was explicit (assigned sample sizes). Within the two primary strata, stratification was implicit, carried out by an ordering of the students within gender in each school using a sort order following the hierarchy below, followed by a systematic sample using this sort order.

Within gender the first secondary stratum was Race/Ethnicity. Within Gender and Race/Ethnicity, the next stratifier was whether or not the student was assigned any developmental courses in math, reading, or writing. The final stratifier was a stratum defined by SAT/ACT score: 1—Lowest Tercile of SAT/ACT within school (0 to 33<sup>rd</sup> percentile); 2—No SAT/ACT score; 3—Middle Tercile of SAT/ACT score (34<sup>th</sup> to 66<sup>th</sup> percentile); 4—Highest Tercile of SAT/ACT score within school (67<sup>th</sup> to 100<sup>th</sup> percentile).

For the SAT/ACT stratum, the terciles were computed separately for eligible students with SAT scores in the school and eligible students with ACT scores in the school. We anticipate most students would have either an SAT score or an ACT score, or neither, but not often both<sup>4</sup>. Stratum 1 consisted then of students who have an SAT score and are in the lowest tercile of SAT scores among eligible students in the school with SAT scores, and students who had an ACT score and are in the lowest tercile of ACT scores among eligible students in the school with ACT scores. Stratum 3 and 4 were defined similarly.

The planned strata hierarchy then was as follows:

---

<sup>4</sup> If both, they were assigned based on the score more often found in the school (if the school has more SAT score students, students with both scores were assigned based on their SAT scores).

- Males or Females;
- Race/ethnicity (1—Hispanic; 2—Black nonHispanic<sup>5</sup>; 3—American Indian or Alaska Native nonHispanic; 4—White nonHispanic; 5—Asian nonHispanic; 6—Native Hawaiian or Pacific Islander nonHispanic; 7—Two or more races nonHispanic; 8—Unknown Race nonHispanic);
- Developmental student (1—Yes; 2—No);
- SAT/ACT stratum (1—Lowest Tercile of SAT/ACT within school (0 to 33<sup>rd</sup> percentile); 2—No SAT/ACT score; 3—Middle Tercile of SAT/ACT score (34<sup>th</sup> to 66<sup>th</sup> percentile); 4—Highest Tercile of SAT/ACT score within school (67<sup>th</sup> to 100<sup>th</sup> percentile).

The final sort order within SAT/ACT stratum was random.

### 4.2.2 Student Sampling Frame Data

As described in section 4.2.1, stratification was by gender, race/ethnicity, developmental status and ACT/SAT score terciles. Generally speaking, these variables were available for all nine colleges, with the exception of ACT/SAT scores, which were not available at all for three of the nine schools and were unavailable for large percentages of students at several schools.

Although the gender, race/ethnicity, developmental status and ACT/SAT score data were available in most schools, there was some missing data in these variables which varied by school and stratification variable. Table 4-1, shown on the next page, gives the count of records on the file provided by each school, the number of those records that were eligible for our sampling frame, and the count of eligible records with missing data for each of the four stratification variables.

---

<sup>5</sup> A student is in the Hispanic stratum if they were positively identified as Hispanic. They are in one of the nonHispanic stratum if they were positively identified as nonHispanic or Hispanicity status is missing.

**Table 4-1: Missing Student Data by School and Stratification Variable**

<b>School</b>	<b>GENDER (missing)</b>	<b>RACE (missing)</b>	<b>DEVELOPMENTAL (missing)</b>	<b>ACT / SAT (missing)</b>	<b>Records Eligible</b>	<b>Records Provided</b>
Austin Community College	6 (0.1%)	132 (20.7%)	216 (33.9%)	n/a	637	5,302
El Paso Community College	0 (0.0%)	89 (3.3%)	283 (10.6%)	n/a	2,667	2,689
Lone Star Community College	0 (0.0%)	46 (3.6%)	0 (0.0%)	919 (72.1%)	1,274	11,614
Prairie View A&M University	0 (0.0%)	4 (0.3%)	0 (0.0%)	586 (52.7%)	1,111	1,212
Texas A&M University	0 (0.0%)	21 (0.3%)	0 (0.0%)	0 (0.0%)	6,202	6,254
Tyler Junior College	8 (0.1%)	102 (5.0%)	0 (0.0%)	1795 (88.5%)	2,028	2,029
The University of Texas at Brownsville / Texas Southmost College	0 (0.0%)	6 (0.4%)	0 (0.0%)	n/a	1,456	1,571
The University of Texas at San Antonio	0 (0.0%)	41 (1.1%)	13 (0.4%)	0 (0.0%)	3,691	4,727
West Texas A&M University	0 (0.0%)	201 (18.2%)	0 (0.0%)	1 (0.1%)	1,103	1,103

Some observations are worth making. First, gender was seldom missing, followed in frequency by developmental status and race / ethnicity. Second, ACT / SAT score data was missing fairly frequently, and some interesting patterns emerged, including schools where practically all records had at least one of ACT / SAT, and other schools where large numbers of cases had neither ACT nor SAT scores, despite that fact that ACT / SAT scores were otherwise available within the school. Third, while most schools provided files where the number of records included was close to the number of records that were truly eligible, two schools provided files with numbers of records far in excess of the number of records that were truly eligible. Prior to sample selection, Westat removed the ineligible cases based on the data provided on the files. The largest numbers of exclusions were mostly due to high school completion years other than 2010.

### **4.2.3 Student Sample Selection**

Once the records for a given school were screened for eligibility (as per the eligibility criteria given in section 4.2.1) and stratified (as described in section 4.2.2), a sample of 148 student records, 74 women and 74 men, were drawn with equal probability within each explicit gender stratum within each school. A preliminary base weight was calculated for each sampled record as the reciprocal of its probability of selection. Once the sample was selected within each school and the preliminary base weight was available, weighted sample tabulations by gender, race / ethnicity, developmental status and ACT / SAT score terciles were compared to the corresponding full population (i.e., sampling frame) values as a check on the sample selection and distribution, which in all cases proved satisfactory.

## **4.3 College Data Collection Plans and Support**

This section provides a description of our approach to data collection for the 2010 NAEP Study of First-Year Texas Postsecondary Students, including the project's multi-pronged methods for working with sampled colleges to maximize survey response; develop data collection plans and schedules; obtain sampled student contact information, and; notify sampled students of their selection while adhering to strict confidentiality and security protocols.

Even though NAEP FYPS followed the standard NAEP 12<sup>th</sup> grade administration procedures as closely as possible, there were several important differences between the high school and college settings which required that NAEP FYPS Westat staff work closely with college administrators at each of the nine colleges prior to receiving OMB clearance, and concurrent with the submission of IRB applications. For example:

- **Attendance Schedules:** For high school students, daily attendance is generally mandatory and most students are at school at the same times Monday through Friday. First-year postsecondary students are not legally required to attend classes, have highly variable schedules, and may be on campus fewer than five days per week.
- **Proximity to Assessment Site:** Colleges and universities typically cover larger areas than high schools, with buildings and classrooms often far away from parking lots and public

transportation. This may pose a greater challenge for postsecondary students in terms of getting to the assessment session, as compared to high school students.

- **Living Arrangements:** The vast majority of high school seniors live at home with parents, while first-year postsecondary students may live on campus, at home with parents, or in private apartments and houses off-campus.

As a result, creative strategies for communicating with sampled students and maximizing their participation in the NAEP FYPS were needed. In addition, data collection was scheduled to occur on each campus over a 5-day period, with a flexible assessment schedule so students could attend an assessment session on the day and time that worked best for them, or choose to “drop-in” to complete the assessment.

#### **4.3.1 NAEP Best Practices Manual**

As part of Westat’s ongoing collaboration with schools to maximize student participation in the assessment, each University/College Coordinator was sent a copy of the National Assessment of Educational Progress (NAEP) Best Practices Manual along with a cover letter. The cover letter, Exhibit 4-1, is shown on the next page.

A copy of the NAEP Best Practices Manual can be found in the NAEP SDC Special Study Task 5 Deliverable dated August 20, 2010.

## **Exhibit 4-1: Cover Letter to College Coordinators Introducing the NAEP Best Practices Manual**

■ [NAME]  
[TITLE]  
[COLLEGE/UNIVERSITY]  
[ADDRESS]

July 27, 2010

Dear [NAME],

I'm forwarding a copy of the National Assessment of Educational Progress (NAEP) Best Practices Manual for your review and use. This manual was designed by the NAEP survey team for use at high schools across the U.S. as they prepare for the 12<sup>th</sup> Grade NAEP assessment data collection each year. It contains many ideas that you might find useful as we prepare to conduct the NAEP assessments on your campus this fall.

Historically, the NAEP assessment surveys have achieved consistently high response rates nationally. However, starting in the late 1990's, response rates among high school seniors began to show a troubling downtrend. In response to this problem, the National Center for Education Statistics convened a Secondary School Principals Working Group in 2005. This Working Group generated a number of innovative and effective ideas for improving response rates among 12<sup>th</sup> graders, and the ones reflected in this manual were implemented during the 2007 NAEP data collection cycle. That year, response rates among 12<sup>th</sup> grade students nationwide increased by 13 points or nearly 20 percent, to 79 percent compared to 66 percent in 2005. In the field of survey research, this represents a huge gain in response rate and provides compelling evidence that older students can be encouraged to participate given the right strategies.

In this manual you will find information about the strategies used by high schools and data on student participation rates. Obviously, the college setting is very different from high school, and not every strategy for improving response rates as covered in the manual will work on your campus. But several of them might be very helpful in getting the best possible response from your first-year students to the NAEP pilot this fall. Please review this manual and feel free to adopt any of these suggested strategies for use on your own campus. If you have questions or comments, please let me know.

Sincerely,

Bob Patchen  
Westat Project Director, NAEP Study of First-Year Texas Postsecondary Students  
[bobpatchen@westat.com](mailto:bobpatchen@westat.com)  
301.610.5113



### 4.3.2 NAEP FYPS Data Collection Plans

After receiving OMB clearance in early August, project staff began constructing data collection plans for each participating school. Due to the compressed schedule for Phase 2 activities, it was necessary to complete data collection plans for several schools with IRB approval still pending.

In order to accurately assess the knowledge and skills the students have upon beginning their postsecondary education, it was critical that the sampled first-year college students took the NAEP assessment as early in their first semester of college as possible. With input from the primary contacts, each of the nine schools was placed into one of three groups assigned to data collection either the week of September 13 – 19, 2010, September 27 – October 1, 2010 or October 4 – 8, 2010. The week of October 11 – 15, 2010 was reserved for make-up data collection on a campus as needed.

Along with identifying when data collection would occur on their campus, the plan also provided schools with:

- Background information on the pilot study;
- Westat's role in data collection;
- The overall NAEP FYPS project schedule for their campus (key dates and deadlines);
- Information on Federal OMB clearance, the date of IRB application and expected approval and;
- Sampling procedures including sampling frame data and sample selection.

In addition, the plan covered the school's role in:

- Assisting Westat in obtaining contact information for sampled students;
- Notifying sampled students of their selection by both mail and email (with provided templates), and;
- Participating in a pre-assessment site visit with the NAEP FYPS Supervisor.

NAEP FYPS reimbursed schools up to \$1,000.00 to help defray some of the administrative costs related to participation. Project staff notified the schools that the administrative reimbursement could not be used to provide student incentives.

A sample Data Collection Plan for Schools can be found in **Appendix C**.

### **4.3.3 Obtaining Contact Information for Sampled Students**

In order for Westat to contact and schedule sampled students for the NAEP assessment, participating schools needed to provide contact information for the 148 sampled students. Contact data was requested, but not required for the creation of the initial sampling frame. After the initial sampling frame was created, a file containing the names of the sampled students was uploaded to Westat's secure FTP link and made available to the school. If the school had not previously provided contact information, they were requested to return the following information (if available) for each sampled student:

- Home Address
- Home Phone Number
- Local Address
- Local Phone Number
- Cell Phone Number
- Primary E-Mail Address
- Secondary E-Mail Address, and
- Current Living Arrangement (on campus, off-campus, commuter, unknown)

All nine schools provided Westat with some form of contact information for the sampled students. In most cases, schools provided a combination of phone numbers and email addresses. This information was then used to create the Record of Contacts (ROC) form used by the data collection team to contact and schedule students for the assessment. Two schools provided limited contact information -- Austin Community College declined to provide email addresses for students and Prairie View A & M only provided email addresses.

### **4.3.4 Student Notification of Selection**

As part of the data collection plan, Westat provided participating schools with a template for the contents of the student notification packet, including:

- A letter notifying sampled students of their selection and describing the importance of the research;
- A schedule showing the times and locations of the NAEP Assessment sessions
- A map of the campus, highlighting the location of the assessment; and

- A flier describing any “thank-you gift” the school was providing to participating students (optional).

The schools were asked to personalize the notification letter, print it on letterhead and have it signed by the college president. The letter provided students with the NAEP FYPS supervisor’s cell phone number and email address to contact with questions or concerns, or requests to set up or reschedule assessment appointments.

Westat staff worked directly with the University/College Contacts to assist them in developing the NAEP Assessment schedule, campus map and flier in advance of the student notification mailing. The template for the contents of the student notification packet can be found in the sample Data Collection Plan for Schools in **Appendix C**.

Each of the schools was asked to mail the notification packet to sampled students approximately two weeks prior to the start of data collection on their campus. Students residing in on-campus housing received the notification packet at their campus address by mail or hand-delivery.

To help ensure that sampled students received notification of their selection prior to the NAEP team’s initial scheduling call, the schools were also asked to email the same notification materials to students approximately one week prior to data collection.

In addition, the participating schools in groups two and three were encouraged to send an additional email to sampled students containing a link to the Department of Education’s YouTube page, and the “What is NAEP” video (<http://www.youtube.com/watch?v=udkIgBzs2TM>) which provided further information about the assessment.

#### **4.3.5 Thank-You Gifts Provided by Colleges**

The NAEP FYPS project provided no incentive to participating students other than the NAEP Certificate of Community Service if requested by the school. Many schools chose to use their own funding to provide refreshments and/or a small “token of appreciation” to students who completed the NAEP assessment. Schools providing “thank-you gifts” other than snacks and/or bottled water included a flier describing the gift in the notification packet and E-mail.

As shown in Exhibit 4-2 below, seven of the nine schools provided refreshments and/or a small gift to students who participated in the assessment.

**Exhibit 4-2: Student Thank-You Gifts Provided by Colleges**

	Bottled Water and/or Bagged Snack	Subway Sandwich	USB Flash Drive	\$10.00 Target Gift Card	Raffle Ticket for 24" Flat-Screen TV	No "Thank-You" Gift
Austin Community College						✓
El Paso Community College	✓	✓				
Lone Star Community College			✓			
Prairie View A & M University			✓			
Texas A & M University	✓					
Tyler Junior College	✓			✓	✓	
The University of Texas at Brownsville/Texas Southmost College	✓		✓			
The University of Texas at San Antonio						✓
West Texas A & M University	✓					

## 4.4 Manual for Conducting Assessment Activities

Due to the importance of following the standard NAEP 12<sup>th</sup> grade administration procedures as closely as possible in the postsecondary administration sessions, the 2010 Manual for Conducting Assessment Activities for main NAEP was used as the template for creating the manual for NAEP FYPS.

A copy of the 2010 NAEP FYPS Assessment Manual can be found in the NAEP SDC Special Study Task 6 Deliverable dated August 20, 2010.

The manual highlighted the important differences between the 2010 NAEP FYPS and regular NAEP, including:

- The NAEP assessment team was responsible for contacting sampled students by phone and/or email to assist in scheduling the assessment sessions and answer any questions the students had about participation.
- The number and types of testing accommodations offered to students on FYPS was limited compared to what is typically offered on regular NAEP. Students who requested an accommodation not offered were encouraged to participate in the study without the accommodation, if possible. Students could decide in advance that they did not wish to take the assessment without a requested accommodation, but no student was considered “excluded” from the assessment.
- There were fewer general student background questions in the assessment booklet, and no subject area background questions.
- There were no teacher or school questionnaires.
- The Administration Schedule, NAEP’s central record-keeping document, was not preprinted with student information or booklet numbers; FYPS field staff entered this information by hand.
- Students were free to choose to attend any session offered during the assessment week at their school. All 148 sampled students in a college/university were listed alphabetically across seven Administration Schedules, so the students listed on a given Administration Schedule were not necessarily tested all together in one assessment session.

#### **4.4.1 NAEP FYPS Procedures and Materials**

Based on these differences, a number of forms and procedures used on regular NAEP were modified for use on NAEP FYPS, and several new forms were developed.

##### **4.4.1.1 Conducting the Student Scheduling Call**

Unlike in regular NAEP, students at each of the nine participating schools had to be recruited individually to participate in the NAEP FYPS study and scheduled to attend one of the many assessment sessions offered during assessment week at the school.

Student recruitment and scheduling was conducted by the assessment team, primarily by telephone. These recruitment and scheduling contacts were referred to as “scheduling calls”, and took place the Wednesday through Sunday prior to the assessment week at the school.

The purpose of the scheduling call was to:

- introduce the NAEP FYPS study;
- remind students of the upcoming assessment week;
- ensure that the student received the advance materials (letter from school, assessment schedule, map) and arrange to have an additional copy sent if necessary;
- identify and quickly respond to issues that could prevent student participation; and
- schedule the student for an assessment appointment.

To ensure that the assessment team provided the same information to each student, the Script for Scheduling Calls to Sampled Students and the Frequently Asked Questions (FAQ) handout, shown as Exhibits 4-3 and 4-4 beginning on the next page, were developed. The assessment team was instructed to use both documents when making scheduling calls to students.

## Exhibit 4-3: Script for Scheduling Call to Sampled Students

■

### Script for Scheduling Call to Sampled Students (revised 9-1-10)

1. Hello, this is (NAEP REPRESENTATIVE NAME) from Westat, a national research company working with the U.S. Department of Education. I'm calling to follow-up with you about the letter sent to you by your college regarding the National Assessment of Educational Progress, Study of First-Year Texas Postsecondary Students. Did you receive that letter?

This research is being conducted with the help of (NAME OF COLLEGE/UNIVERSITY) to learn more about how high schools prepare students for college, and to find ways to improve the experience of students who are, like you, entering college for the first time.

You have been chosen at random to represent thousands of other new college students across Texas. Your participation in this study is voluntary. If you join in this study, your part will be to attend an on-campus NAEP session and fill out a randomly-assigned booklet of math or reading questions, plus several background questions.

The entire process lasts about one hour and 15 minutes. You don't need to study or prepare anything in advance – just pop in and take a seat. Your results and all other personal information about you will be kept strictly confidential, and your grades will not be affected in any way.

2. Do you have any questions? (RESPOND TO QUESTIONS.)

3. Can we count on you take part in this important study?

Yes (CONTINUE TO Q4)

No (RESPOND TO OBJECTIONS)

4. The NAEP sessions will be held at (name of college) during the week of (dates) at (locations). For your convenience we've scheduled several sessions each day so you can choose the time that's easiest for you to attend. Please tell me which of the following would work best for you:

(READ ASSESSMENT DATES/TIMES AND LOCATIONS AND SCHEDULE STUDENT. ASK TO MAKE SURE THE STUDENT IS FAMILIAR WITH THE TEST LOCATION AND PROVIDE DIRECTIONS AS NEEDED.)

5. We'd like to send you a reminder the day before your NAEP session. What is your preferred e-mail address? May I also have your preferred phone number?

E-MAIL: \_\_\_\_\_

PHONE NUMBER: \_\_\_\_\_

Thanks for your time. If you have any questions, or if you need to change your NAEP session, please call me at \_\_\_\_\_ or e-mail me at \_\_\_\_\_.

## Exhibit 4-4: Frequently Asked Questions (FAQs)

### **FAQ's for Texas First-Year College Student Pilot Study**

#### **What's this study about?**

The purpose of this study is to learn more about how high schools prepare students for college and to find ways to improve the college experience for new students.

#### **Who is sponsoring this study?**

The National Center for Education Statistics, part of the U.S. Department of Education, is conducting this study in cooperation with your college, the Texas Higher Education Coordinating Board (THECB) and the National Assessment Governing Board.

#### **Who is Westat?**

Westat is a national employee-owned research company based in Rockville, Maryland. We perform research projects for many government agencies, including the U.S. Department of Education. We have performed the sampling and data collection tasks for the National Assessment of Educational Progress (NAEP) continuously since 1983.

#### **How/Why was I chosen?**

You have been chosen at random to represent thousands of other new college students across Texas.

#### **Why do you want to test me?**

We want to learn more about how high schools prepare students like you for college. We also want to find ways to improve the experience of students who are, like you, entering college for the first time.

#### **I don't think I want to do it, but my friend might. Can he/she volunteer?**

Sorry, the results would not be statistically valid. The goals of this study require a representative, random sample which cannot be obtained using volunteers.

#### **I already took the SAT/ACT. Can I just tell you my scores?**

We need you to take this test which is different from the SAT or ACT. It has been specially designed to meet the research needs of this study.

#### **How long will this take?**

The testing will take about one hour to one hour and 15 minutes.

#### **Do I have to do this all today?**

Yes, the test is administered to you in one session.



#### Exhibit 4-4: Frequently Asked Questions (FAQs) (cont.)

**Do I have to study for this?**

No, you don't need to study or prepare anything in advance. Just come to the designated location at the scheduled time and take a seat. We'll even give you a pencil!

**I'm a bad test taker/I hate to take tests.**

Well, you don't have to worry about that for this test. This test can have no impact on your grades. There's no pressure because we won't tell anyone how you did, even you. Your test results will be combined with those of other students and used for research purposes only. All we ask is that you do your best.

**What kinds of things will you ask me?**

You will be randomly assigned to answer either some math or some reading questions. You will only answer questions on one subject, not both. The test is in two parts, each lasting about 25 minutes. After you finish both parts of the test, you will be asked to answer four background questions to provide a little information about yourself.

**Will you tell me how I do on the test?**

No, none of the students will receive their scores on this assessment. Your test results will be combined with those of other students.

**Will the college get my score?**

No, all results will remain confidential and will not be reported to your college, or to anyone else. Only the aggregate or combined test scores will be used.

**Will my parents get my score?**

No, all results will remain confidential and will not be reported to your parents, or to anyone else.

**Will this affect my grades?**

No, it will have no impact on your grades.

**Will I get paid to take the test?**

No, we can't pay you anything for taking the test. [Describe any thank-you gifts to be provided by the college.]

**Why should I do this?**

This is an opportunity for you to be involved in research that could improve the way high school students prepare for college, both in Texas and nationwide. Every student selected for this study is important, and you represent thousands of other students who were not selected.

## Exhibit 4-4: Frequently Asked Questions (FAQs) (cont.)

Revised 9-1-10

### **Is this mandatory? Do I have to do it?**

Participation in this study is voluntary. Your decision to take part is completely up to you. However, we do hope you will help us with this important project.

### **Who else did you ask to do this?**

We selected 148 first-year students on your campus to take part in this study. Each sampled student represents hundreds of others who were not selected.

### **I usually have/use (TESTING ACCOMMODATION) when I take tests like this. Can I have/use/bring (ACCOMMODATION) for this test?**

Some accommodations are permitted for this test. If it is one of the permitted accommodations, AND if the accommodation is officially approved by the college/university for your use, we will provide you with that accommodation. [REFER TO APPENDIX SECTION OF MANUAL FOR LIST OF PERMITTED ACCOMMODATIONS.] [CONFIRM WITH STUDENT THAT HE/SHE HAS UNIVERSITY APPROVAL FOR ACCOMMODATION.] IF APPROPRIATE: If it is not an accommodation permitted on this test, you are welcome to come view some sample test questions to see if you would like to try to take the test without the accommodation.

### **When will this happen?**

We will be on campus during the week of (dates) at (locations). We've scheduled several sessions each day so that you can choose the time that's best for you to attend. I'll be glad to work with you to find a good time. There will also be several "drop-in" sessions if you miss your scheduled session for any reason. All of the times and locations are listed in the schedule you received with your notification letter from the President of your college several days ago.

### **Do I need to bring anything?**

No, we will supply everything you will need.

### **How long will this take? OR How long do I have to stay here (if already at the assessment location)?**

The NAEP booklets contain two parts. Each part is timed for 25 minutes. To meet the requirements of this research study, students should remain seated with the test in front of them for the full 25 minutes allotted for each section. If you finish early you may go back and check your work.

#### **4.4.1.2 The Student Record of Contacts (ROC)**

The Record of Contacts (ROC) was the key record-keeping document used to track and detail all scheduling call attempts and outcomes with sampled students for the FYPS pilot. Using information from the student sampling lists, a ROC was created for each of the nine schools, pre-printed with each sampled student's name and contact information as well as their demographic information.

Because the Record of Contacts contained student names and contact information, all FYPS staff observed all NAEP rules for handling secure documents when working with this form.

Students were listed on the ROC three to a page, as shown in Exhibit 4-5, on the following page. The remainder of the form was for the Assessment Administrator to complete to assist in making and documenting the scheduling and reminder calls to students.

Approximately two weeks prior to the start of data collection at a school, two copies of the ROC were sent to the supervisor via FedEx, signature required. The supervisor kept one copy of the ROC for her records and assigned approximately one-third of the remaining ROC to each of her AAs assigned to that school. Each AA received just their assigned portion of the ROC via FedEx, signature required.

The AAs used the ROC to contact the students they were assigned to recruit and schedule for the assessment, and returned the ROC to the supervisor on the first day of data collection. The ROC was used by the team during the assessment week to continue scheduling calls as needed, and make reminder calls to students with scheduled appointments.

At the end of assessment week, student names and contact info were removed from both copies of the ROC and left at the school in the NAEP Storage Envelope. The supervisor returned the remaining portion of the ROC to the home office with the School Folder.

2010 NAEP First-Year Texas Postsecondary Study															School Name: BEDROCK COMMUNITY COLLEGE						
Scheduling Call-Record of Contacts															School ID: 483-000-0						
Line #	Student Name and Contact Info.	Line #	Birth date		Sex	Race/Ethnicity	Attempt # 1-3 Sched. 4-5 Remind.	Date	Day of Week	Time	Contact Type						Comments	Planning Code	Admin. Code from A.S.		
			Mo.	Yr.							Cell Phone	Other Phone	Text Msg.	Campus Email	Home Email	Other					
1	Name: Abbott, Angelica S.	1	1	1992	2	1	1				am										
	Cell Phone: 555-555-1234																				
	Local Phone: 555-555-3876																				
	Home Phone: 555-555-9420																				
	Primary Email: primary@gmail.com																				
	Secondary Email: secondary@yahoo.com																				
2	Name: Aguilar, Noe	2	2	1992	1	3	1				am										
	Cell Phone: 555-555-8975																				
	Local Phone: 555-555-7780																				
	Home Phone: 555-555-1004																				
	Primary Email: primary@gmail.com																				
	Secondary Email: secondary@yahoo.com																				
3	Name: Alcantar, Hubert	3	9	1991	1	3	1				am										
	Cell Phone: 555-555-9763																				
	Local Phone: 555-555-4237																				
	Home Phone: 555-555-1829																				
	Primary Email: primary@gmail.com																				
	Secondary Email: secondary@yahoo.com																				

Disposition Codes	1. Agree-appt. set	3. Refusal-firm	5. Refusal-soft	7. Left/sent msg.	9. Busy signal	11. Disconnected phone	13. Withdrawn/not enrolled
	2. Agree-drop in	4. No answer	6. Refusal-parent	8. Ans. machine no msg.	10. Bad/failed email	12. Call back/not home	14. Other

#### 4.4.1.3 Assessment Accommodations

Main NAEP assessment procedures encourage inclusion of students with special needs, and permits most commonly offered testing accommodations for special needs students during the assessment. Based on guidance from NCES, discussions with the participating schools about relevant testing accommodation policies and practices, and given the schedule limitations, NAEP FYPS offered the following 10 testing accommodations to participating students:

- Extended Time
- Small Group
- One-On-One
- Directions Only Read Aloud in English
- Breaks During Testing
- Large Print Version of the Test
- Magnification Equipment
- Uses Template/Special Equipment/Preferential Seating
- Bilingual Dictionary without Definitions (Math only)
- Read Aloud – Occasional (Math only)

Even though the FYPS pilot was not able to offer the full set of accommodations offered on main NAEP, the ten accommodations offered represented the most common accommodations provided to students. Since the NAEP FYPS Reading and Mathematics Assessments would only be administered in English, other than the use of a bilingual dictionary without definitions for the math assessment, no other language accommodations were offered to students.

To adequately inform the development of procedures for future studies, it was important that information was collected about students who requested accommodations, whether they used one or not. NAEP FYPS staff were trained to respond to any student request, inquiry or expressed need for a testing accommodation with the following response from the FAQs:

Some accommodations are permitted for this test. If it is one of the permitted accommodations, AND if the accommodation is officially approved by the college/university for your use, we will provide you with that accommodation. [REFER TO APPENDIX SECTION OF MANUAL FOR LIST OF PERMITTED ACCOMMODATIONS.] [CONFIRM WITH STUDENT THAT HE/SHE HAS UNIVERSITY APPROVAL FOR ACCOMMODATION.] IF APPROPRIATE: If it is not an accommodation permitted on this test, you are welcome to come view some sample test questions to see if you would like to try to take the test without the accommodation.

Two forms were used to document any request by a sampled student for a testing accommodation:

- the Record of Contacts, and
- the Accommodation Tracking Form

Because the NAEP FYPS pilot did not collect accommodation use on the assessment booklet covers or on the Administration Schedule, the Accommodation Tracking Form, shown as Exhibit 4-6 on the next page, was the primary mechanism for tracking information about accommodation request and use on this study. All of the accommodations provided in regular NAEP are listed on this form. The accommodations provided for the FYPS pilot are the first 10 listed, and printed in bold. The form was designed to document the following:

- Scheduling call outcomes where a student requested or inquired about using an accommodation to take the assessment;
- Any request or inquiry about using an accommodation to take the assessment made by a student at the assessment location; and
- Any student who received one of the ten accommodations offered on the FYPS pilot (in order for a student to receive one of these accommodations the student must have official approval from the college/university to use it in testing situations).

Across all nine participating schools, data collection teams reported that no students inquired about, or requested a testing accommodation for NAEP FYPS.

At each of the nine participating schools, in order to request an accommodation, a student must come forward, identify and document their disability and the need for accommodation, and then specifically request the provision of such accommodation. The difference in procedures for obtaining accommodations in the post-secondary versus high school environment may have accounted for the lack of student requests for accommodations on the pilot.

## Exhibit 4-6: Accommodation Tracking Form

<b>2010 NAEP Study of First-Year Postsecondary Students</b> <b>ACCOMMODATION TRACKING FORM</b>																								
Region: _____	School Name: _____	School ID: 483- _____																						
AA or Supervisor completing form: _____		Date: _____																						
<b>Instructions:</b> Any time a student requests or inquires about, or uses a testing accommodation for the assessment, use this form to record the specific accommodation and to provide details about the request.																								
<b>Note:</b> Use the Session Debriefing Form to document issues regarding a specific assessment session.																								
<u>If applicable:</u> Student line number: _____ Booklet ID: _____ Admin. Code: _____																								
<p>Circle the code(s) for the accommodation(s) requested, whether or not the accommodation was received. If the accommodation is not listed, circle the code for "Other" and specify the accommodation in the space below. If the accommodation circled needs clarification, use the space below. Accommodations in bold print can be provided for the NAEP FYPS. Refer to the Appendix section in your manual for details.</p> <table style="width: 100%; border: none;"><tbody><tr><td style="width: 50%; vertical-align: top;">01. Extended Time</td><td style="width: 50%; vertical-align: top;">12. Must have Aid to Administer Test</td></tr><tr><td style="vertical-align: top;">02. Small Group</td><td style="vertical-align: top;">13. Responds Orally to a Scribe</td></tr><tr><td style="vertical-align: top;">03. One-on-One</td><td style="vertical-align: top;">14. Uses a Calculator for All Sections</td></tr><tr><td style="vertical-align: top;">04. Directions Only Read Aloud in English</td><td style="vertical-align: top;">15. Cuing to Stay on Task</td></tr><tr><td style="vertical-align: top;">05. Breaks During Testing</td><td style="vertical-align: top;">16. Presentation and/or Response in Braille</td></tr><tr><td style="vertical-align: top;">06. Large Print Version of the Test</td><td style="vertical-align: top;">17. Presentation and/or Response in Sign Language</td></tr><tr><td style="vertical-align: top;">07. Magnification Equipment</td><td style="vertical-align: top;">18. Inclusion Booklet</td></tr><tr><td style="vertical-align: top;">08. Uses Template/Special Equipment/Preferential Seating</td><td style="vertical-align: top;">19. Directions Only Read Aloud in Spanish</td></tr><tr><td style="vertical-align: top;">09. Bilingual Dictionary without Definitions (<b>math only</b>)</td><td style="vertical-align: top;">20. Test Items Read Aloud in Spanish</td></tr><tr><td style="vertical-align: top;">10. Read Aloud – Occasional (<b>math only</b>)</td><td style="vertical-align: top;">21. Spanish/English Version of the Test</td></tr><tr><td style="vertical-align: top;">11. Read Aloud – Most or All</td><td style="vertical-align: top;">22. Other, specify below</td></tr></tbody></table>			01. Extended Time	12. Must have Aid to Administer Test	02. Small Group	13. Responds Orally to a Scribe	03. One-on-One	14. Uses a Calculator for All Sections	04. Directions Only Read Aloud in English	15. Cuing to Stay on Task	05. Breaks During Testing	16. Presentation and/or Response in Braille	06. Large Print Version of the Test	17. Presentation and/or Response in Sign Language	07. Magnification Equipment	18. Inclusion Booklet	08. Uses Template/Special Equipment/Preferential Seating	19. Directions Only Read Aloud in Spanish	09. Bilingual Dictionary without Definitions ( <b>math only</b> )	20. Test Items Read Aloud in Spanish	10. Read Aloud – Occasional ( <b>math only</b> )	21. Spanish/English Version of the Test	11. Read Aloud – Most or All	22. Other, specify below
01. Extended Time	12. Must have Aid to Administer Test																							
02. Small Group	13. Responds Orally to a Scribe																							
03. One-on-One	14. Uses a Calculator for All Sections																							
04. Directions Only Read Aloud in English	15. Cuing to Stay on Task																							
05. Breaks During Testing	16. Presentation and/or Response in Braille																							
06. Large Print Version of the Test	17. Presentation and/or Response in Sign Language																							
07. Magnification Equipment	18. Inclusion Booklet																							
08. Uses Template/Special Equipment/Preferential Seating	19. Directions Only Read Aloud in Spanish																							
09. Bilingual Dictionary without Definitions ( <b>math only</b> )	20. Test Items Read Aloud in Spanish																							
10. Read Aloud – Occasional ( <b>math only</b> )	21. Spanish/English Version of the Test																							
11. Read Aloud – Most or All	22. Other, specify below																							
Describe the details about the accommodation used or requested: <div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px; margin-bottom: 5px;"></div> <div style="border-bottom: 1px solid black; height: 15px;"></div>																								

## **4.5 Field Staff Recruitment and Training**

This section describes the overall plan for recruiting and training experienced NAEP supervisors and Assessment Administrators to conduct the NAEP FYPS assessments during the fall of 2010.

### **4.5.1 Supervisor and Assessment Administrator Recruiting**

Due in part to the compressed schedule for Phase 1 and 2 activities, only experienced NAEP field staff were recruited to work on the pilot study. Three experienced NAEP Supervisors (SVs) who lived in Texas were hired in mid-August and assigned to one of three regions, each containing three of the nine sampled schools:

- Region 1: West Texas A & M University, El Paso Community College, and Tyler Junior College
- Region 2: The University at Brownsville and Texas Southmost College, Austin Community College and The University of Texas at San Antonio
- Region 3: Texas A & M University, Lone Star Community College and Prairie View A & M University

Each of the supervisors was provided with a project laptop with wireless data card and a portable printer, as well as a dedicated project email address. Over the course of the study, the laptops were used to communicate with home office staff, school contacts and sampled students.

The supervisors recruited experienced NAEP Assessment Administrators (AAs) who lived near the sampled schools in Texas to assist in contacting the sampled students and administering the assessments. If a region's schools were spread out geographically, supervisors attempted to hire a team of three local AAs per school to minimize field costs. Each supervisor hired the following number of AAs for their region:

- Region 1: Eight Assessment Administrators
- Region 2: Nine Assessment Administrators
- Region 3: Five Assessment Administrators



Since the NAEP Reading and Mathematics Assessments would only be administered in English, supervisors did not specifically recruit bilingual Assessment Administrators.

### **4.5.2 Training**

Because all field staff were experienced with previous NAEP high school data collection cycles, the training for NAEP FYPS focused primarily on procedures for contacting and recruiting sampled students. All field staff received eight hours of in-person training, the NAEP FYPS Manual for Conducting Assessment Activities, and a NAEP FYPS cell phone to use when contacting sampled students. During training, staff received a refresher on the actual administration of the NAEP assessment, since, for the pilot, the actual administration sessions were designed to replicate the 12<sup>th</sup> grade process as closely as possible.

The NAEP FYPS Manual for Conducting Assessment Activities can be found in the NAEP SDC Special Study Task 6 Deliverable dated August 20, 2010.

Detailed agendas for the NAEP FYPS Supervisor and Assessment Administrator trainings are presented in **Appendix C**.

#### **4.5.2.1 Supervisor Training**

Training for the three NAEP FYPS Supervisors (SVs) occurred on August 25, 2010 in Rockville, Maryland. SV training focused primarily on procedures and forms exclusive to supervising NAEP FYPS – specifically, completing Administration Schedules, contacting and recruiting sampled students and working directly with the sampled schools to schedule Pre-Assessment visits. At the conclusion of their in-person training, supervisors received the materials necessary to train their AAs.

#### **4.5.2.2 Assessment Administrator Training**

The NAEP FYPS supervisors conducted training for their Assessment Administrators in Texas near their assigned locations between August 27 and September 1, 2010. Once again, training focused primarily on contacting and recruiting sampled students and included several role-play exercises. In addition to completing eight hours of in-person training, each AA completed four hours of home study to practice making scheduling calls to sampled students before they began data collection activities.

## 4.6 Data Collection

The primary activities for NAEP FYPS data collection included:

- Determining the data collection schedule for each participating school;
- Conducting a Preassessment site visit to each school;
- Mailing and emailing notification letters to prepare sampled students for the assessment;
- Contacting sampled students and securing cooperation; and
- Administering the NAEP Assessment.

### 4.6.1 Data Collection Schedule for Colleges

Westat worked closely with each of the pilot schools to plan appropriate assessment dates for each campus. Data collection occurred on each campus over a 5-day period and was conducted in three waves using three groups of colleges each, as shown in Exhibit 4-7 below.

#### Exhibit 4-7: College Data Collection Schedules

<u>Group 1 Schools:</u> <ul style="list-style-type: none"><li>• Texas A &amp; M University</li><li>• The University of Texas at Brownsville / Texas Southmost College</li><li>• West Texas A &amp; M University</li></ul>	<u>Data collection:</u> 9/13 – 9/19/2010
<u>Group 2 Schools:</u> <ul style="list-style-type: none"><li>• El Paso Community College (Valle Verde Campus)</li><li>• Lone Star Community College (Montgomery Campus)</li><li>• Austin Community College (Cypress Creek Campus)</li></ul>	<u>Data collection:</u> 9/27 – 10/1/2010
<u>Group 3 Schools:</u> <ul style="list-style-type: none"><li>• The University of Texas at San Antonio</li><li>• Tyler Junior College</li><li>• Prairie View A &amp; M University</li></ul>	<u>Data collection:</u> 10/4 – 10/8/2010

Data collection began for Group 1 schools the week after the Labor Day holiday. No data collection was scheduled for the week of September 22, 2010, to allow Westat Home Office staff to reflect on the first week of data collection, debrief the supervisors, and make any needed adjustments to field procedures before continuing with data collection for Group 2 schools.

#### **4.6.2 Preassessment Sit Visit to Colleges**

Between September 2 – 22, 2010, and prior to the start of data collection on each campus, the NAEP FYPS Supervisor conducted a Preassessment Visit (PAV) to each of their three sampled schools. To prepare for the visit, each Supervisor received a Preassessment Visit Folder which provided:

- Contact information for the University/College Coordinator;
- A script to schedule the PAV – either via phone or by email;
- A list of materials they needed to bring with them during the PAV;
- A check-list of topics to discuss with the University/College Coordinator during the PAV; and
- A script for a “thank-you” email to be sent to the University/College Coordinator after the visit was complete.

College Contacts were asked to allot three to four hours for the preassessment site visit. This allowed the NAEP FYPS Supervisor enough time to:

- Establish contact with the University/College Coordinator and additional campus staff, and review the details of plans for conducting NAEP assessments on campus;
- Resolve any questions/problems with the student sample (for example students who have dropped since sampling occurred or obtain additional student contact information to assist in making scheduling calls);
- Tour the assessment location(s);
- Finalize logistical arrangements for the NAEP team during assessment week, such as parking, security passes, procedures for gaining access to assessment locations; etc.;
- Finalize plans for any “thank-you” gifts the university/college was providing to students upon completion of the assessment; and
- Answer questions as needed.

### **4.6.3 Student Recruitment and Scheduling**

Using the Record of Contacts (ROCs) log, cell phones, contacting script and FAQs provided by the project, each NAEP FYPS data collection team began making scheduling calls to sampled students the Wednesday prior to the start of data collection at the school.

Assessment Administrators recorded the details of each scheduling contact attempt on the Record of Contacts (ROC). To increase the chance of contacting a sampled student, AAs were encouraged to attempt contact at all phone numbers provided by the college, and to vary the days and times of their calls. Since many of the participating colleges IRB's required that the study must be truly voluntary for selected students, AAs were instructed to record all student refusals on the ROC, and subsequently, terminate all contact with the student.

Initially, the scheduling calls were to be conducted Wednesday through Saturday between 9:00 a.m. and 8:00 p.m. and the AAs were trained to limit their scheduling call attempts to three, at which point they entered one of the first four disposition codes in the "Planning Code" column of the ROC.

The planning code reflected the summary outcome of scheduling calls, and was transcribed onto the Administration Schedule on the first day of the assessment week. Students who were found to be ineligible for data collection, or those who declined to participate in the assessment because a requested testing accommodation was not offered, did not receive a planning code.

Based on feedback from field staff during the first round of scheduling calls, the scheduling call window was extended to 9:00 p.m., and calls could be made to students on Sunday between 12:00 and 5:00 p.m. In addition, the number of scheduling call attempts was extended when there was no answer and no opportunity to leave a voice mail message.

In data collection weeks two and three, the planning code was no longer assigned after the third call attempt, but rather assigned based on the outcome of the scheduling attempts through Sunday.

During the data collection week, field supervisors reviewed the ROCS daily to assist in devising strategy for the next contact attempt.

#### **4.6.4 Evaluation and Revision of Field Procedures**

As mentioned previously, no data collection was scheduled for the week of September 20, 2010 so that the Project and Field Director could conduct an initial debriefing of the field supervisors on their experiences during the first week of data collection. During the debriefing call, the supervisors brought up the following issues:

- AAs felt that many students were not responding to the team's scheduling calls due in part to the Maryland area codes assigned to the cell phones;
- Contacting the sampled students proved to be more challenging and require more time from the AAs than initially expected.
- Using the ROC to keep track of reminder calls to students with scheduled assessments was difficult.

To address these issues prior to the start of week 2 of data collection, home office staff:

- Obtained cell phones with local Texas area codes through the wireless company, Cricket, and sent them to AAs via FedEx so they could use them for the scheduling calls for week two.
- Obtained an additional two laptops with wireless data cards for each data collection region and sent them out to the field via FedEx so they could be used to send scheduling and reminder e-mails to sampled students for weeks two and three data collection.
- Extended the scheduling call hours from 9:00 a.m. until 9:00 p.m. and Sunday's from noon to 5:00, and also expanded the number of scheduling calls from three to six. However, staff were instructed to reduce the number of voicemail messages left to students.
- Developed a Daily Scheduling Log so that Supervisors and AAs could keep better track of student appointments and reminder calls.

#### **4.6.5 Quality Control**

NAEP FYPS followed the standard NAEP 12<sup>th</sup> grade administration and quality control procedures as closely as possible. All NAEP FYPS field staff understood the importance of maintaining data security and the highest standards of confidentiality. As part of the hiring paperwork for NAEP FYPS, all field staff completed the Westat Code of Conduct Ethics Training and signed the Westat Data Collector Code of Conduct and Assurance of Confidentiality, an Affidavit of Nondisclosure and a Conflict of Interest form.

To help ensure the safe transport of materials, hard copy documents containing respondent information were only exchanged between the home office and field via Federal Express. Regardless of the package and regardless of who received the package, the project required that packages be sent using the Direct Signature Required option.

As part of the overall quality control procedures, supervisors observed each of their AAs over the course of each data collection week worked, and completed an Assessment QC Record to document whether the AA administered and monitored the session according to the project's standards and procedures.

In addition, the Project and Field Director conducted additional field observations.

#### **4.6.6 Make-Up Data Collection**

The week of October 11 – 15, 2010 was reserved for make-up data collection at a school if deemed necessary by NAEP FYPS project staff and NCES. Based on discussions with NCES, Westat requested two schools participate in make-up data collection that week – the University of Texas at Brownsville/Texas Southmost College (UTB/TSC) and the University of Texas at San Antonio (UTSA).

During initial data collection at UTB/TSC, several special circumstances occurred which may have impacted the number of scheduled students who actually showed up for the assessment – very heavy rain on more than one day, and a reporting of shots fired from across the Mexico border which landed on campus. For these reasons, a full week of make-up assessment sessions was conducted.

During the first two days of data collection at UTSA, very few students showed up for their scheduled assessment. Since the study was already in week three of data collection, the decision was made to ask UTSA to allow us to conduct make-up assessments on October 13-14, 2010, to try and increase response rates at the school.

The make-up assessment sessions yielded eight additional assessments at UTB/TSC and an additional three assessments at UTSA.

#### **4.6.7 Field Staff Debriefing**

The Project and Field Director held two telephone debriefing sessions with 11 Assessment Administrators on November 3, 2010. The purpose of the debriefing was to have AAs provide their observations and opinions on all aspects of the pilot study, including feedback on how procedures could be improved if future iterations of NAEP FYPS were conducted.

To help facilitate the discussion, a copy of the debriefing questionnaire was emailed to AAs several days before the telephone sessions were held, so they could review the discussion topics in advance. A copy of the debriefing questionnaire can be found in **Appendix C**.

#### **4.6.8 Sample Cleaning**

During data collection, several key issues with the sample came to light which required Westat to re-contact each of the nine colleges to verify student eligibility. Across many of the colleges, during the scheduling calls, several students mentioned that they:

- Had completed GEDs rather than high school in 2010;
- Had lost their student status due to non-payment of tuition;
- Had withdrawn from the college;
- Did not attend classes on the sampled campus; or
- Attended on-line only classes.

During the Scheduling calls to students attending El Paso Community College – Valle Verde Campus, NAEP Field Staff discovered that the original sampling frame erroneously included students who attended classes on EPCC's other campuses. After consulting with NCES, it was determined that students who did not have at least one class on the sampled campus (Valle Verde) would be excluded from the sample.

Following the regular and make-up data collection weeks, Westat re-contacted each of the schools to verify the final eligibility of the sample, and removed any ineligible students from the sampling frame as needed.

## **4.7 Response Rates**

During data collection, preliminary Data Collection Summary Reports were provided to NCES twice per week. Following the regular and make-up data collection weeks, the NAEP FYPS preliminary overall response rate was 20.5%. Once data cleaning was conducted with each of the nine colleges to remove ineligible students from the overall sample, the overall response rate across the nine schools rose to 20.7%.

Recruitment was attempted for 100 percent of the eligible sample at each pilot college. Recruitment methods included advance mailings and telephone calls, advance e-mails, and reminder calls/e-mails during the assessment week. Students' cell phone numbers, school and personal e-mail addresses were used when provided by the college. All contacted students were offered a range of assessment days and times throughout the week. However, sampled students could attend ANY assessment session, regardless of which session they requested in advance. Multiple drop-in sessions were also offered for students who missed scheduled sessions.

The response rate at individual campuses ranged from 6.9% at one two-year community college to 33.3% at one four-year university. Table 4-7-1 on the following page, presents the overall summary data collection report as of October 15, 2010 as well as the cumulative response rates by school group.

Overall, the four-year universities exhibited a higher response rate (24.0% combined average) than the two-year community colleges (15.7% combined average). However, this was not a consistent pattern at the individual college level and one, two-year community college had the second-highest response rate at 32.7%, while one four-year university was at 15.5%



**Table 4-7-1: Data Collection Summary Report**

Survey Totals	All Groups	Group 1	Group 2	Group 3
	9/13 - 10/8	9/13 - 9/17	9/27 - 10/1	10/4 - 10/8
Cumulative Response Rate <sup>1</sup>	20.7%	27.1%	14.9%	19.4%
Eligible Sample (n) <sup>2</sup>	1234	425	376	433
Number of Students Attending	255	115	56	84
Percent Males	48.2%	45.2%	62.5%	44.0%
Percent Females	51.8%	54.8%	37.5%	56.0%
<sup>1</sup> <b>Cumulative Response Rate:</b> By end of day on Friday, a combined total of 27.1% of Group 1 students, 14.9% of Group 2 students and 19.4% of Group 3 students had attended a NAEP session.				
<sup>2</sup> <b>Eligible Sample (n):</b> sample base applies for the entire week. Make-up exam sessions were conducted at UTB/TSC (Group 1) and UTSA (Group 3) from 10/11/10 - 10/15/10. A total of 11 additional assessments were completed on these dates - 8 assessments at UTB/TSC and 3 assessments at UTSA.				

## 4.8 Non-Response Bias Analysis

Following completion of data collection for the Phase 2 pilot study for the NAEP Study of First-Year Texas Postsecondary Students, Westat conducted a nonresponse bias analysis to evaluate differences of response rates across important categories of students which are likely to be related to achievement differentials. The most important of these are the developmental categories: whether or not a first-year student was enrolled in any developmental courses in their first semester. Other categories of interest are gender, race/ethnicity, and academic achievement levels. All of these characteristics were collected on the sample frames provided by the colleges and used to form strata in the sampling process.

All response rates were calculated among eligible students only. We present response rates over all eligible students in the study, and by each school individually. We performed significance tests of response rate differences using binomial distribution tests and logistic regression. A *nonsignificant* result means that the difference in response rates across schools or groups of students may have occurred simply by chance, and does not necessarily indicate an actual difference in the study population. In other words, the differences might be explained simply as an artifact of the ‘luck of the draw’, and the apparent difference observed may likely not be observed in a second, new sample of the same size from that population. A *significant* result means that the probability that the apparent difference can be explained entirely as ‘the luck of the draw’ is low, and the difference as observed would likely persist in future samples from the same population. The observed difference can, *with high confidence*, be inferred to be a real one in the population.

We also provide distributions across subgroups for all sampled students and assessed students. Differences in these distributions are driven in fact by the response differences, and thus in a sense

are presenting the same results as the response rate tables in a different form. But seeing the results in this form is useful in evaluating the potential impact of response rate differences in estimates of average achievement levels, by showing how the distribution in the responding sample differs from that of the population across subgroups that can be expected to have differing achievement levels. If the percentage distribution of the responding samples closely resembles the distribution of the study population, then the risk of bias in the final estimates is generally lower than it would be with large differences. The distributions for the sampled eligible students are very close to the corresponding frame distributions, as the sampled eligible students are a stratified random sample from the corresponding set of eligible students on the nine school frames from which we sampled. Thus, for all practical purposes, the sampled eligible distribution can be viewed as a proxy for the population distribution.

Included with the sample and assessed student distributions are indices: ratios between the distributional percentages for the assessed students and the distributional percentages for the sampled students. A ratio or index of 100 indicates a perfect match between the cell distribution for the responding sample and the frame. A number greater than 100 for the index corresponds to overrepresentation of that particular subgroup in the assessed sample, and an index less than 100 corresponds to underrepresentation, as compared to the frame. These indices provide the reader an easy point of reference for identifying groups that were over- or underrepresented in the responding sample.

It should be noted that the nine schools in this study are very different: we expect considerable differences across schools. These nine schools were selected purposively; that is, they were “hand-picked” to represent a wide variety of school types, student types, regions of the state, and selectivity in admissions standards. Purposive sampling is a useful way to ensure representation of selected characteristics in the study sample, especially when sample size must be limited. But unlike random sampling, purposive sampling makes it impossible to easily generalize to the entire state. The actual mix and proportions of types of schools and types of students in the total Texas postsecondary population is not the same as our pilot sample of nine schools. Therefore, the pilot results, strictly speaking, are only representative of these nine schools and should not be viewed as projectable statewide in Texas.

Also note that the response rates within each school and within various subgroups are treated as if they are sample percentages from a simple random sample from a homogeneous population. In particular, the binomial distribution will be utilized to evaluate significance of response rate differences.

### 4.8.1. Overall Response Rates

The overall response rate across the nine schools was 20.7%. Table 4-8-1, shown below, presents the overall response rates by school. There were considerable differences across the schools and several of these are highly statistically significant. The school-level response rates ranged from a high of 33.3% at Texas A&M University to a low of 7.2% at Austin Community College (Cypress Creek Campus). A logistic regression analysis of response propensity was done with a separate intercept for each school. The p-value is less than 0.0001 for an 8-degree of freedom chi-square test of the null hypothesis of no differences across schools, which means we can, with great confidence, reject the null hypothesis that there are no systematic response rate differences between the schools.

To identify where these differences lie, we conducted multiple comparisons tests of response rates of all school pairs using a Bonferroni adjustment as shown in Table 4-8-2, on page 4-37. Most pairs are not significant, but a number of pairs are significant (consistent with the chi-square test result). Austin Community College (Cypress Creek) has a significantly lower response rate than all five schools with the highest response rates, and Texas A & M University has a significantly higher response rate than the four schools with the lowest response rate. El Paso Community College (Valle Verde Campus) has a significantly higher response rate than the three schools with the lowest response rates.

The four-year colleges have higher observed response rates than the community colleges except for the two outliers, Prairie View A & M University (a four-year college but with a low response rate) and El Paso Community College (a community college but with a higher response rate). We tested the contrast between four-year colleges and two-year colleges as combined samples; this test yields a p-value of 0.0005, thus there is a significant finding of higher response rates among four-year colleges as opposed to two-year colleges in the pilot.

**Table 4-8-1. Overall Response Rates by College**

School Name	# of Eligible Sampled Students	Response Rate
Austin Community College (Cypress Creek)	139	7.2%
Lone Star Community College (Montgomery)	140	10.0%
Prairie View A & M University	147	15.6%
Tyler Junior College	140	18.6%
West Texas A & M University	139	23.0%
University of Texas at San Antonio	146	24.0%
University of Texas at Brownsville/Texas Southmost College	139	24.5%
El Paso Community College (Valle Verde)	97	33.0%
Texas A & M University	147	33.3%
All Phase 2 Schools	1234	20.7%

**Table 4-8-2. Response Rate Multiple Comparison Tests**

School	Response Rate	Austin	Lone Star	Prairie View A&M	Tyler	West Texas A&M	San Antonio	Brownsville	El Paso	Texas A & M
Austin Community	7.2%		NS	NS	NS	*	*	*	*	*
Lone Star	10.0%	NS		NS	NS	NS	NS	NS	*	*
Prairie View A&M	15.6%	NS	NS		NS	NS	NS	NS	*	*
Tyler Junior	18.6%	NS	NS	NS		NS	NS	NS	NS	*
West Texas A&M	23.0%	*	NS	NS	NS		NS	NS	NS	NS
San Antonio	24.0%	*	NS	NS	NS	NS		NS	NS	NS
Brownsville	24.5%	*	NS	NS	NS	NS	NS		NS	NS
El Paso	33.0%	*	*	*	NS	NS	NS	NS		NS
Texas A & M	33.3%	*	*	*	*	NS	NS	NS	NS	

A 'NS' means that there is no significant difference between the response rates of schools in that pair.

An '\*' means that there is a significant difference between the response rates of schools in that pair.

## 4.8.2. Response Rates by Gender

Differences in response rate by gender are commonly seen in surveys, with females generally having higher response rates than males. However, in this study we did not in general see large gaps between males and females in response. The overall response rate of all eligible male students is 19.8%, slightly lower than the 21.5% for all eligible female students, as seen in Table 4-8-3 on the next page. Table 4-8-4 shows the distribution differences between sampled and assessed students. The chi-square test of differences between males and females gives a p-value of 0.4540: not significant. Therefore we can conclude that the small gender differences in response did not significantly affect the distribution of the final assessed sample. We also tested the effect of gender in logistic regression models along with a school effect (an intercept for each school) or 4-year college vs. 2-year college. In all these tests, there is no significant difference in the response rates of male and female students. This is an encouraging result, suggesting that future assessment surveys of this type with this population, conducted with similar methods, could expect to achieve roughly equivalent response from male and female students.

**Table 4-8-3. Student Response Rates and Sample Distributions, by Gender**

Subgroup	Eligible Sample Students	Assessed Total	Response Rate
Males	621	123	19.8%
Females	613	132	21.5%
Total	1234	255	20.7%

**Table 4-8-4. Student Sample Distributions, by Gender**

Subgroup	Eligible Sample Total	Eligible Sample Percent	Assessed Total	Assessed Percent	Index*
Males	621	50.3%	123	48.2%	96
Females	613	49.7%	132	51.8%	104
Total	1234	100.0%	255	100.0%	

\*The index is the ratio of the assessed percentage and the sample percentage.

Table 4-8-5 below presents response rates by gender for each school individually. One can see that Lone Star is a definite outlier: a considerably lower female response rate than male response rate (4.4% vs. 15.3%). This reduced what otherwise may have been the usual significant gap in the direction of higher response rates for women. Childcare and work concerns were one factor given by sampled women at Lone Star accounting for this low female response rate.

**Table 4-8-5. School Response Rates of Eligible Students, by College and Gender**

School Name	Gender	# of Eligible Sampled Students	Response Rate
Austin Community College Cypress Creek	Male	68	7.4%
	Female	71	7.0%
Lone Star Community College Montgomery Campus	Male	72	15.3%
	Female	68	4.4%
Prairie View A & M University	Male	73	13.7%
	Female	74	17.6%
Tyler Junior College	Male	67	14.9%
	Female	73	21.9%
West Texas A & M University	Male	68	22.1%
	Female	71	23.9%
University of Texas at San Antonio	Male	73	23.3%
	Female	73	24.7%
University of Texas at Brownsville/Texas Southmost College	Male	69	21.7%
	Female	70	27.1%
El Paso Community College Valle Verde	Male	57	33.3%
	Female	40	32.5%
Texas A & M University	Male	74	28.4%
	Female	73	38.4%

### 4.8.3. Response Rates by Developmental Status

Developmental status (enrollment in remedial courses) is one of the most important student characteristics to examine in this nonresponse analysis given the purposes of this study. This was a stratification variable because we were concerned that response rates might be lower for students who were judged to be in need of developmental instruction by their college. In the event, the difference in response rates was not large, and was not significant<sup>6</sup>. The response rate for students defined as developmental is 19.2%, slightly lower than the 21.9% for the non-developmental students. Table 4-8-6 shows the school response rates by developmental status. Table 4-8-7 provides the comparisons of assessed sample versus eligible sample distributions. The assessed sample of developmental students indexed at 93 compared to the total eligible sample, indicating only a slight underrepresentation of this group in the final sample.

**Table 4-8-6. Student Response Rates, by Developmental Status**

Subgroup	Eligible Sample Students	Assessed Total	Response Rate
Non- Developmental	652	143	21.9%
Developmental	582	112	19.2%
Total	1234	255	20.7%

**Table 4-8-7. Sample Distributions, by Developmental Status**

Subgroup	Eligible sample total	Eligible sample percent	Assessed total	Assessed percent	Index*
Non- Developmental	652	52.8%	143	56.1%	106
Developmental	582	47.2%	112	43.9%	93
Total	1234	100.0%	255	100.0%	

### 4.8.4. Response Rates by Race/Ethnicity

There were seven race/ethnicity categories which we tested for significant differences in response rate. The six degree of freedom chi square test of race/ethnicity effect alone is significant with a p-value 0.023, indicating that response rates for some of the race/ethnicity groups were significantly higher or lower than those for other groups. We also conducted multiple comparisons tests of response rates of all race/ethnicity pairs. There is a borderline significant difference (Bonferroni adjusted p-value 0.053) between Hispanics (24.9% response rate) and nonHispanic Whites (16.7%

---

<sup>6</sup> This nonsignificance result occurred in the simple one degree of freedom chi square test, and in logistic regression models including school intercept parameters.

response rate). We are somewhat surprised at the comparatively low response rate for nonHispanic Whites, which is not commonly seen in surveys. Generally speaking, nonHispanic Whites are among the higher responding race/ethnicity groups, so this result seems anomalous.

However, the race/ethnicity effect is only borderline significant (p-value 0.072) when the school effect (eight degrees of freedom) is included into a logistic regression model. This indicates some confounding between school effect and race/ethnicity effect. In this case, one or two schools with a high percentage of Hispanics, such as El Paso Community College (roughly 90% Hispanic), may be driving a higher response rate for Hispanics, but the higher response may be partially driven by other factors impacting response rates at the school itself. Given the small number of schools and small sample sizes, it is not possible to resolve this confounding completely. However, we know that EPCC provided small “Thank-You Gifts” to their students to participate, including a free lunch from Subway, which could have driven higher response rates compared to other colleges with lower percentages of Hispanic students.

Tables 4-8-8 and 4-8-9 present response rates and distributions for Hispanics, nonHispanic Whites, nonHispanic Blacks, and Other nonHispanics<sup>7</sup>.

**Table 4-8-8. Student Response Rates by Race/Ethnicity**

Subgroup	Eligible Sample Students	Assessed Total	Response Rate
Hispanic	421	105	24.9%
Black or African-American, not Hispanic	237	45	19.0%
White, not Hispanic	461	77	16.7%
Other, not Hispanic	115	28	24.3%
Total	1234	255	20.7%

**Table 4-8-9. Sample Distributions by Race/Ethnicity**

Subgroup	Eligible Sample Total	Eligible Sample Percent	Assessed Total	Assessed Percent	Index*
Hispanic	421	34.1%	105	41.2%	121
Black or African-American, not Hispanic	237	19.2%	45	17.6%	92
White, not Hispanic	461	37.4%	77	30.2%	81
Other, not Hispanic	115	9.3%	28	11.0%	118
Total	1234	100.0%	255	100.0%	

<sup>7</sup> The four categories Asians and Pacific Islanders, American Indians and Alaskan Natives, Multi-race persons, and Other/Unknown Race/Ethnicity are collapsed into a single “Other” category because of small sample sizes..

### 4.8.5. Response Rates by Achievement Stratum

The achievement strata were defined using four levels (1—Low Tercile<sup>8</sup>; 2—Missing Achievement<sup>9</sup>, 3—Middle Tercile<sup>10</sup>; 4—High Tercile<sup>11</sup>). For all schools, these strata were based on SAT/ACT scores. As shown in Table 4-9-10 the response rates increase with achievement stratum. The significance test showed that achievement stratum is significant when it is alone in the model (p-value 0.002). It is also significant with school effect included in the model (p-value 0.022), or with the 4-year college indicator in the model (p-value 0.025). We also did multiple comparison tests of all achievement strata pairs, as shown in Table 4-8-12, on the following page. We can see that the low and unknown achievement strata have significantly lower response rates than the middle and high achievement strata.

**Table 4-8-10. Student Response Rates, by Achievement Stratum**

Subgroup	Eligible sample students	Assessed total	Response rate
Low Achievement Tercile	206	36	17.5%
Missing Achievement	652	119	18.3%
Middle Achievement Tercile	196	45	23.0%
High Achievement Tercile	180	55	30.6%
Total	1234	255	20.7%

**Table 4-8-11. Sample Distributions by Achievement Stratum**

Subgroup*	Eligible Sample Total	Eligible Sample Percent	Assessed Total	Assessed Percent	Index*
Low Achievement Tercile	206	16.7%	36	14.1%	85
Missing Achievement	652	52.8%	119	46.7%	88
Middle Achievement Tercile	196	15.9%	45	17.6%	111
High Achievement Tercile	180	14.6%	55	21.6%	148
Total	1234	100.0%	255	100.0%	

<sup>8</sup> 0<sup>th</sup> to 33<sup>rd</sup> percentile in achievement within each school (for SAT, or ACT).

<sup>9</sup> Neither SAT nor ACT score present.

<sup>10</sup> 33<sup>rd</sup> to 67<sup>th</sup> percentile in achievement within each school (for SAT or ACT).

<sup>11</sup> 67<sup>th</sup> to 100<sup>th</sup> percentile in achievement within each school (for SAT or ACT).



**Table 4-8-12. Academic Stratum Multiple Comparison Tests**

Academic Stratum	Response Rate	Low Tercile	Missing	Middle Tercile	High Tercile
Low Achievement Tercile	17.5%	NS	NS	NS	*
Missing Achievement	18.3%	NS	NS	NS	*
Middle Achievement Tercile	23.0%	NS	NS	NS	NS
High Achievement Tercile	30.6%	*	*	NS	NS

A 'NS' means that there is no significant difference between the response rates of schools in that pair.

An '\*' means that there is a significant difference between the response rates of schools in that pair.

### 4.8.6. Conclusions

The overall student response rate in this pilot study was much lower than expected (20.7%), despite the extensive efforts to contact and recruit all eligible students and excellent cooperation from the pilot college administrations. The low overall response rate presents two critical challenges: it increases the risk of nonresponse error in the survey estimates, and it also reduces the operational feasibility of the study design.

Viewed on the whole, the results of the nonresponse bias analyses presented here suggest a surprisingly mild risk of nonresponse error. Despite the low overall response rate, differences across most key subgroups were small and nonsignificant. For the most part, the response rate differences observed in the pilot could, in practice, be ameliorated through post-stratification weighting. But attempting to conduct the main study of first-year postsecondary students statewide in Texas would probably be cost-prohibitive if only one in five eligible students can be expected to take part.

Several findings have important implications for future similar research. Student response rates differed considerably across schools, with two-year colleges showing lower response rates in general than four-year colleges, but with differences across schools within these two broad groupings. On the whole it appears that future studies should focus extra attention, preparation and respondent treatments on the two-year colleges in an effort to bring their overall response rates more in line with the four-year schools.

The considerable difference in response rates across individual schools makes it difficult to resolve confounding issues for subgroups which are more heavily represented in some schools than in others. Race/ethnicity is a particular example of this. Response rate differences in these cases could be ascribed to school-level differences (e.g., specific respondent treatments used, level of interest in

the study, school loyalty) or to differences across the race/ethnicity groups. This confounding is not resolvable without further new data.

There did not appear to be a gap between male and female response rates, though there was a pattern of nonsignificant differences suggesting slightly higher response rates for females in most schools. Hispanics showed a higher response rate than nonHispanic Whites, though the evidence for this is obscured by confounding with school-level differences. In this case we need to ask, would Hispanics have shown such comparatively high response rates as a group if not for the participation of El Paso Community College? Or, did a concentration of Asian students at Texas A&M University drive higher response rates for that college? While such interaction effects are probably real, they are a constant part of understanding nonresponse effects in surveys and are not particularly problematic. It's just important to understand that apparent differences across the race/ethnicity groups could be rooted as much in differences between the colleges as between the students themselves.

There is some evidence that developmental students have slightly lower response rates. The evidence of a response differential by achievement tercile is very strong. This achievement score relationship to response propensity is of considerable concern, as achievement score is highly related then to the characteristics of interest in this study. Future studies of this kind should collect this information and make weighting adjustments to reduce the potential for bias from this achievement/response propensity relationship.

This section of the report summarizes the most important lessons we learned from the Phase 2 pilot study as they relate to the prospects for administering NAEP in the postsecondary setting going forward. We also provide recommendations for improving implementation of similar research in the future.

## 5.1 Working with the Colleges and Universities

**College and university administrations are keenly interested in the college preparedness issue and are willing and able to participate in administering NAEP to support research efforts in this area.** We found broad-based interest and engagement in this research across multiple functional areas and administrative levels at the pilot colleges, not just staff implementing a directive from the university president. This suggests that future efforts to study academic linkages between the secondary and postsecondary student populations would also be well-received.

**The role of the THECB was vital to keeping the pilot on schedule and helping to cut through “red tape”.** The endorsement of this study by the THECB and the strong appeal for cooperation and support provided in the letter sent by Commissioner Paredes to college presidents proved invaluable. This would be an essential element for future similar research in Texas or other states.

**The role of the contact person/study liaison assigned by each college was also vital to the successful planning and execution of sampling and data collection.** It is critical that this person have good working relationships with and knowledge of the registrar’s office, provost, student affairs, director of research & evaluation, director of IT/information systems, and the facilities manager.

**There is a direct tension between the sampling and data collection tasks for this project and the need to assess students early in their first college semester.** A key goal of the research design was to administer 12<sup>th</sup> grade NAEP to first-year college students before they received much postsecondary instruction. Otherwise, the test results could provide invalid information on the students’ level of preparedness before arriving at college. But for some colleges it was a challenge to

provide complete student characteristics and contact information needed for sampling and recruitment as early as August-September. Enrollment lists change as students are added and deleted, students who recently moved to the area have not had time to update their contact information, and some academic records and other student characteristics data have not yet been entered into the college information systems. It would be useful to continue to explore ways to extend the data collection window or sample a subset of students for whom the needed data can be made available earlier in the semester.

**More work is needed to compensate for the absence of the NAEP infrastructure and long history of operations in the K-12 setting, which benefit and support NAEP data collection today but do not exist in the postsecondary world.** This means that far more lead time (than was available in the pilot) is needed to prepare postsecondary administrators and students for the sampling and data collection process, and to promote awareness of, and interest in, NAEP among students and faculty. In order to effectively prepare for a fall semester data collection window, it appears that formal planning and coordination activities with the colleges should begin the preceding January, in parallel with planning for the summer orientation sessions.

**Summer orientation sessions are not useful for data collection, but are useful for promoting awareness of NAEP.** During Phase 1, we learned that while all of the pilot colleges provide some form of summer orientation for first-year students, none of them were willing to schedule NAEP assessments during these programs. All explained that the orientation schedules were already overloaded, and students had very little free time available during the sessions. Instead, all of the pilot schools indicated that using the orientation programs to promote awareness of NAEP, explain the benefits of the postsecondary research, and advertise the school's support for the project would be the better strategy.

**The process of applying for IRB approvals at each individual college or university is schedule and labor intensive.** In a large scale, statewide study involving several dozen colleges it would either be necessary to add a large home-office staff dedicated to this task, or work with the THECB or other appropriate state agency to obtain a coordinated approval accepted by all the participating schools.

## 5.2 Recruiting the Students

**The logistical differences in conducting NAEP data collection in high schools versus the postsecondary setting will require using incentives to obtain comparable response rates.**

The fact that colleges and universities cannot provide the “captive audience” that NAEP generally encounters in the high school setting means that some aspects of the NAEP data collection protocol should be re-engineered and tailored to the postsecondary setting. It appears that acceptable response rates will not be obtained unless monetary incentives are used.

**Cell phones, texting and e-mail are the most popular forms of communication for postsecondary students, but each presents some unique problems.** This information is protected by some colleges and could not be provided to the project in all cases. Advance negotiations are needed with some colleges and universities to gain the release of this contact information. Also, cell phones present pros and cons for the recruiters. While having access to the student numbers is essential, relying only on cell phone contact is often not effective. Virtually all cell phones include caller ID displays, so users can check the incoming number before answering. Unfamiliar numbers may not be answered. Also, students often turn off their cell phones while in class or at the library, making it difficult to reach them with reminder calls during the day. Solutions for this include using cell phones for the field staff that have local area codes, leaving appropriate voice mail messages to inform students about the purpose of the call, and promoting awareness of NAEP in advance to increase the likelihood of callbacks.

**Separate, dedicated recruitment and administration teams may be useful in raising student response rates, and could be more efficient for large-scale postsecondary surveys.** During the pilot, the same teams of field staff that administered NAEP at each college also performed the recruitment and reminder phone calls to students. As we increased efforts to contact more of the sampled students and added more reminder calls during the assessment week, this became burdensome for the assessment administrators and supervisors in some cases. Separate teams dedicated to either contacting/recruiting students or administering the NAEP sessions would be able to specialize and focus all their attention on the one task area, by working in waves through the data collection period. While one team administers the assessments, the other could move on to begin recruitment at the next college, while also handling reminder calls for the first school. This would also allow us to assign the best telephone recruiters and the best administrators to the applicable teams.

**The NAEP Best Practices Manual provides valuable guidance to administrators for improving response rates at colleges.** It can be revised to better reflect the postsecondary setting and provided to study coordinators earlier in the year for use in planning response rate strategies. It could also be used as background material for focus groups and brainstorming sessions with coordinators from multiple colleges to generate ideas for improving student response rates. (However, this would require allowing the participating school contacts to know each other's identity.)

**Multiple, convenient assessment locations are needed to optimize student participation.** However, space for conducting assessments on campus is at a premium, so advance reservations are needed. With more advance lead-time most schools would be able to ensure space in several locations rather than just one or two as was the case in the pilot.