



## **Decision-Making Practices of Urban Districts for Including and Accommodating English Language Learners in NAEP -- School-Based Perspectives**

A Study Sponsored by the National Center for Education Statistics,  
U.S. Department of Education

Lynn Shafer Willner, Ph.D., Charlene Rivera, Ed.D.,  
and Barbara D. Acosta, Ph.D.

The George Washington University  
Center for Equity and Excellence in Education  
1555 Wilson Boulevard, Suite 515  
Arlington, VA 22209  
Tel: 703.528.3588  
Fax: 703.528.5973

January 31, 2007

## **Acknowledgements**

The authors would like to thank the many individuals who provided guidance and assistance with this report. The project was funded by the National Center for Education Statistics (NCES), under the leadership of Arnold Goldstein, who offered invaluable guidance and feedback throughout the process. Holly Spurlock and Lisa Ward of NCES provided logistical support in determining and arranging contact with the schools that participated in the study. We are grateful to Megan Rice and Lori Nicholson of the American Institutes for Research who supported the project during the sample selection process.

George Washington University Center for Equity and Excellence in Education staff providing important support throughout the project included Eric Collum who managed the early phases of the project, including data collection and writing the first draft of the report. Carlos Astrada assisted the research team by building a database and Web interface for the collection of interview data. Charlotte Blane provided valuable assistance editing the final draft of the report. Cerissa Parrenas set up the initial interviews, and throughout the project, Roshuan Tyson provided administrative support.

Finally, special thanks to Louise Millikan, our intrepid interviewer who not only persisted in tracking down the decision makers at each of the schools in our sample, but also conducted the interviews in such a way that school decision makers were willing to discuss their decision making practices in detail.

**About The George Washington University  
Center for Equity and Excellence in Education**

The mission of The George Washington University Center for Equity and Excellence in Education (GW-CEEE) is to advance education reform so that all students achieve high standards. GW-CEEE provides evidence-based technical and professional development and conducts research and evaluation studies for state education agencies, local education agencies, and various offices of the U.S. Department of Education to facilitate education reform and school improvement initiatives. GW-CEEE was established in October 1991 and chartered by The George Washington University in January 1994.

**The George Washington University  
Center for Equity and Excellence in Education**

1555 Wilson Boulevard, Suite 515  
Arlington, Virginia 22209

Tel: 703.528.3588

[ceeeinfo@ceee.gwu.edu](mailto:ceeeinfo@ceee.gwu.edu)

<http://ceee.gwu.edu>

## **Executive Summary**

This report presents findings from a study conducted by The George Washington University Center for Equity and Excellence in Education (GW-CEEE) under the sponsorship of the National Center for Educational Statistics (NCES). The purpose of the study is to describe and analyze school-based decision-making practices relevant to the inclusion and accommodation of English language learners (ELLs) in the 2005 National Assessment of Educational Progress (NAEP), which was administered as part of the Trial Urban District Assessment (TUDA).

### *Research Questions*

This study investigated how school personnel in urban districts make decisions regarding the inclusion and accommodation of ELLs in NAEP. The study explored the following research questions for a sample of four of the 11 TUDA districts for the 2005 administration of Reading and Mathematics assessments in NAEP at Grades 4 and 8:

1. What factors influenced decisions of school personnel regarding whether to include or exclude ELLs in NAEP?
2. What factors influenced decisions of school personnel regarding the use of accommodations for those ELLs who were included in NAEP?
3. What was the relationship, if any, between the decision to include and the decision to accommodate?

### *Research Methods*

Researchers applied qualitative and case study methods to identify common patterns of decision-making in urban districts regarding inclusion and accommodation of ELLs. The research team collected data through telephone interviews with decision makers in 29 schools with high ELL student populations in four districts that had participated in the 2005 administration of NAEP as part of the TUDA program.

Data analysis indicated five themes (or dimensions) related to the decision-making process: (a) school personnel involved, (b) involvement of NAEP field staff, (c) NAEP tools used, (d) inclusion criteria used, and (e) accommodations criteria used. Three of the dimensions (school personnel, inclusion criteria, and accommodation criteria) coincided with the framework developed in Rivera et al. (2006). The other two dimensions emerged from discussions with NCES officials regarding the involvement of NAEP field staff and the use of NAEP tools in the decision-making process. Researchers analyzed data associated with the themes, to construct four district summaries, providing the reader with a synthesis of the findings in relation to the study's three research questions.

The district summaries of the decision-making practices used in schools were then analyzed as part of cross-district comparisons. To support the validity and reliability of these comparisons, the research team employed a case study strategy. Following methodology outlined in Yin (2004), the research team generated a theory of the core concepts being studied – i.e., the decision-making practices for including and accommodating ELLs. This theory was codified into three statements which functioned as the research propositions against which the data were examined.

### Research Propositions

Decision makers relied on one of the following sets of criteria for inclusion and/or accommodation decisions:

- A. ELL – responsive criteria

Data that supported Research Proposition A would indicate that decision makers consider ELLs’ linguistic and socio-cultural needs during high-stakes testing. Rivera et al. (2006) identified the following categories of ELL-responsive criteria: *language-related*, *academic-related*, *time-related*, and *opinion-related* criteria. “*Language-related* criteria ... included language proficiency in English and the native language as well as all educational context, such as program placement and language of instruction. *Time-related* criteria pertain to the length of exposure a student has to an English-speaking academic environment. *Academic-related* criteria relate to

students’ prior schooling and academic achievement as measured by test performance. *Opinion-related* criteria address teacher and parent inputs to the identification of eligible ELLs for accommodations on state assessments” (p. 82).

Data that supported Research Proposition B would indicate that decision makers did not consider criteria relevant to the specific linguistic needs of ELLs. In their study of state assessment policies, Rivera et al. (2006) found that ELLs and SDs were often grouped together within state policies as “special needs students,” “at-risk students,” or “special populations,” and that state policies often did not distinguish between those accommodations appropriate for ELLs and those appropriate for other students. An SD-responsive approach is based on the assumption that only ELLs who have Individualized Education Plans (IEPs) are eligible for accommodations. Decision makers applying this proposition use disability taxonomy for identifying appropriate accommodations, selecting from the following categories: (1) timing/scheduling accommodations, (2) setting accommodations, (3) presentation accommodations, or (4) response accommodations. An SD-responsive taxonomy does not include accommodations that address ELLs’ unique needs during assessment.

Data that supported Research Proposition C would indicate that decision makers interpreted the state assessment policy as requiring equality of treatment for ELLs and native English speakers.

### Findings

This study explored the factors that influenced inclusion and accommodation decisions for ELLs in NAEP and the extent to which there was a relationship between inclusion and accommodation decision-making practices. The research team first examined district-level decision-making patterns for correspondence with the three research propositions.

For ELL inclusion decisions, decision makers at 26 of the 29 sample schools across the four districts said they relied on ELL-responsive criteria – e.g., *language-related*, *academic-related*, *time-related*, and *opinion-related* criteria. Data further indicated that decision makers applying ELL-responsive criteria had selected from a wide range of criteria from within this category. For example, some decision makers said they had applied NAEP criteria as well as one or more ELL-responsive criteria, including amount of instruction in English (3 years or less), academic performance, and student background factors such as English language proficiency and time in the U.S. This variability of criteria may have influenced inclusion rates. For example, interview data indicated that ELLs were more likely to be excluded in District 3 than in District 1

or District 4 due to the application of the criterion “less than three years of instruction in English.”

For accommodations decisions, the application of SD criteria was equally as prevalent as the application of ELL-responsive criteria. A review of the accommodations decision-making practices for the 29 schools indicated that 13 schools relied on ELL-responsive criteria, while almost as many schools (12) relied on criteria intended for students with disabilities.

An analysis of the relationship between inclusion and accommodation decision-making practices in schools across the four sample districts identified four patterns of inclusion and accommodation decision-making among schools in the four districts:

- Pattern I: ELL-responsive criteria applied to both inclusion and accommodation decisions
- Pattern II: ELL-responsive criteria applied to inclusion decisions, but SD-responsive criteria (e.g., student had an IEP) applied to accommodation decisions
- Pattern III: ELL-responsive criteria applied to inclusion decisions, but no accommodations allowed
- Pattern IV: Blanket inclusion of ELLs, but applied SD-responsive criteria (e.g., student had an IEP) to accommodation decisions

Patterns I and II were the dominant patterns across most schools. Interview data indicated that decision makers in a large majority of sample schools (25 of 29) followed either Pattern I or Pattern II decision-making practices. The 13 school-based decision makers who followed Pattern I said they relied on ELL-responsive criteria for their inclusion and accommodation decisions. The 12 decision makers who followed Pattern II said they relied on ELL-responsive criteria when including and excluding ELLs from NAEP, but applied criteria reserved for students with disabilities (i.e., student had an IEP) when accommodating ELLs.

A within-district analysis indicated that only one district applied inclusion and accommodations decisions consistently. A majority of school decision makers in District 3 reported that they had relied on ELL-responsive criteria for both inclusion and accommodations decisions. These decision makers reported, furthermore, that their ability to provide accommodations had influenced the decision to include or exclude ELLs in NAEP. In contrast, data for District 1, District 2, and District 4 decision-making patterns indicated little relationship between the provision of accommodations and the inclusion decision for ELLs. Although decision makers in District 2 said they relied on ELL-responsive criteria for inclusion decisions, accommodation decisions were split between reliance on ELL-responsive criteria and reliance on a disabilities framework. District 1 and District 4 data for accommodations decisions were highly variable, even though decision makers in both of these districts said they had relied on ELL-responsive criteria for their ELL inclusion decisions.

In sum, findings from this study reflect the complex interaction of decision-making practices with the school culture, district and state policies, and the pressures of daily existence in schools. These findings support the conclusion that (1) reliance on state assessment policy or classroom practice for NAEP inclusion and accommodations criteria may result in inconsistent and/or inappropriate decisions; (2) that school decision makers need a greater understanding of the needs of ELLs in relation to high stakes testing, in particular regarding the distinction between the needs of ELLs and those of students with disabilities; and (3) that stronger guidance may be needed from NAEP in order to assure the use of appropriate criteria to support the decision-making process.

### *Recommendations*

Findings from this study suggest a number of ways for NAEP to support improved decision-making for including and accommodating ELLs. This support can be provided at three moments: prior to the pre-assessment visit, during the pre-assessment visit, and during test administration.

1. Train NAEP staff to recognize common areas of misinterpretation of NAEP guidelines.
2. Provide explicit guidance to school decision makers regarding principles of fair and appropriate testing for ELLs. Make explicit the distinction between the needs of ELLs and those of students with disabilities for participation in high stakes testing.
3. Include a summary of important principles and common areas of misinterpretation in the letter sent to schools along with the new NAEP guidelines prior to the assessment visit.
4. Collaborate with school decision makers during the pre-assessment visit to review their decision-making process. Explicitly guide decision-makers to align their decisions with principles of fair and appropriate assessment of ELLs.
5. Prior to the day of NAEP administration, ensure that there are enough test administration materials for students who will receive accommodations and enough personnel to administer the test.

### *Conclusion*

A decade ago, school personnel reported that they were less likely to include ELLs in NAEP due to concerns that students might not be able to participate meaningfully due to their low levels of English language proficiency and because appropriate accommodations were not available (Olson & Goldstein, 1997). Findings from the current study suggest that decision makers' concerns about the meaningful participation of ELLs have not yet been adequately addressed. Although we found an increase in rates of inclusiveness for ELLs in our sample, the low rate of accommodations in some districts and schools and the selection of inappropriate accommodations in other districts and schools is cause for continuing concern. Withholding appropriate accommodations from students who need them not only reduces ELLs' opportunity to participate meaningfully in the assessment but also introduces new sources of error that jeopardize the validity of conclusions regarding the magnitude of achievement gaps. Until ELLs are provided consistent opportunities to participate meaningfully in large-scale assessments, valid and reliable assessment of ELLs is likely to remain one of unrealized promise.

To be effective, accommodations must address the unique needs of the students for whom they are provided. For ELLs, this means assistance in overcoming the linguistic and socio-cultural barriers that prevent them from demonstrating the academic knowledge and skills being tested. Without adequate accommodations, ELL achievement scores cannot accurately reflect what students know and can do (LaCelle-Peterson & Rivera, 1994). By knowing more about how school personnel make decisions about including and accommodating ELLs, NAEP field staff will be better able to provide the guidance necessary to increase the appropriate participation of ELLs in NAEP.

## Table of Contents

Introduction.....	1
Justification for the Study .....	1
<i>The Impact of the Inclusion of ELLs on NAEP Statistics</i> .....	1
<i>Background on NAEP Efforts to Increase the Inclusion of ELLs</i> .....	1
<i>The Trial Urban District Assessment Program</i> .....	4
<i>2005 ELL Inclusion and Accommodation Rates for NAEP TUDA Project</i> .....	5
Research Questions .....	9
Research Methods.....	9
<i>Sampling Plan</i> .....	9
<i>Data Collection</i> .....	12
<i>Data Analysis</i> .....	13
<i>Validity and Reliability</i> .....	15
<i>A Review of the NAEP Inclusion and Accommodation Process</i> .....	16
District Findings.....	17
<i>District 1</i> .....	17
<i>District 2</i> .....	19
<i>District 3</i> .....	20
<i>District 4</i> .....	23
Cross-Case Analysis of Decision-Making Practices in the Four Districts .....	25
<i>Comparison of Inclusion Decision-Making Practices Across the Four Districts</i> .....	26
<i>Comparison of Accommodation Decision-Making Practices Across the Four Districts</i> .....	28
<i>Cross-Case Analysis of the Relationship between Inclusion Decision-Making and Accommodation Decision-Making</i> .....	30
Discussion .....	34
Recommendations.....	35
<i>Future Research</i> .....	36
Conclusion .....	36
References.....	38
Appendix A: Interview Protocol for NAEP School Decision Makers .....	40
Appendix B: Analysis Codes .....	47
Appendix C: District Overviews and School Profiles for the Four TUDA Districts.....	52
District 1.....	52
<i>Fisk</i> .....	57



<i>Rice</i> .....	58
<i>Tiant</i> .....	59
<i>Yazstremski</i> .....	60
<i>Boomer</i> .....	60
<i>Lynn</i> .....	62
District 2 .....	63
<i>Sandburg</i> .....	67
<i>Middleton</i> .....	67
<i>Pronto</i> .....	68
<i>Day</i> .....	69
<i>Raleigh</i> .....	69
<i>Bacon</i> .....	70
<i>Drake</i> .....	71
District 3 .....	72
<i>Dickens</i> .....	77
<i>Hardy</i> .....	78
<i>Bronte</i> .....	80
<i>Bishop</i> .....	81
<i>Pynchon</i> .....	81
<i>Spenser</i> .....	82
<i>Chapman</i> .....	83
<i>Fletcher</i> .....	84
District 4 .....	85
<i>Tully</i> .....	90
<i>Tacitus</i> .....	91
<i>Livy</i> .....	92
<i>Homer</i> .....	93
<i>Smithers</i> .....	94
<i>Ovid</i> .....	94
<i>Josephus</i> .....	95
<i>Hesiod</i> .....	96

### Tables

Table 1. Frequently Provided Accommodations for ELLs for the 2005 NAEP .....	4
Table 2. 2005 NAEP ELL Inclusion Rates in TUDA Districts .....	6
Table 3. 2005 NAEP ELL Accommodation Rates in TUDA Districts .....	7
Table 4. Correlation Coefficients Between ELL Inclusion Rate and ELL Accommodations Rate on the 2005 NAEP .....	8
Table 5. ELL Participation in NAEP in the Four Sample TUDA Districts .....	10
Table 6. Operationalized Factors for Each of the Five Dimensions of Decision-Making .....	14
Table 7. Schools Within the Four Sample TUDA Districts .....	17
Table 8. Cross-Case Analysis of Inclusion Decision-Making Practices for ELLs .....	27
Table 9. Cross-Case Analysis of Accommodation Decision-Making for ELLs .....	29
Table E1. Enrollment per Grade Level(s) in District 1 Schools Participating in the Study .....	52
Table E2. Number of Individuals in the Decision-Making Unit in District 1 and Their Role/Expertise .....	53
Table E3. Involvement of NAEP Field Staff-District 1 .....	54
Table E4. NAEP Tools Used-District 1 .....	55
Table E5. Inclusion Criteria Used-District 1 .....	56
Table E6. Accommodations Criteria Used-District 1 .....	57
Table E7. Enrollment for Grade Level(s) in District 2 Schools Participating in Study .....	63
Table E8. School Personnel Involved-District 2 .....	63
Table E9. Dimension 2: Involvement of NAEP Field Staff-District 2 .....	64
Table E10. NAEP Tools Used-District 2 .....	65
Table E11. Inclusion Criteria Used-District 2 .....	65
Table E12. Accommodations Criteria Used-District 2 .....	67
Table E13. Enrollment for Grade Level(s) in District 3 Schools Participating in Study .....	72
Table E14. School Personnel Involved-District 3 .....	73
Table E15. Involvement of NAEP Field Staff-District 3 .....	74
Table E16. NAEP Tools Used-District 3 .....	75
Table E17. Inclusion Criteria Used-District 3 .....	76
Table E18. Accommodations Criteria Used-District 3 .....	77
Table E19. Enrollment for Grade Level(s) in District 4 Schools Participating in Study .....	85
Table E20. School-Based Decision Makers-District 4 .....	86
Table E21. Involvement of NAEP Field Staff-District 4 .....	87
Table E22. NAEP Tools Used -District 4 .....	88
Table E23. Inclusion Criteria Used-District 4 .....	89
Table E24. Accommodations Criteria Used-District 4 .....	90

### Figures

Figure 1. Variability in ELL Accommodation Rates in TUDA Districts on 2005 NAEP .....	8
Figure 2. Sampling Plan .....	12
Figure 3. Patterns of Decision-Making in Relation to Three Research Propositions. ....	31

## Introduction

This report presents findings from a study conducted by The George Washington University Center for Equity and Excellence in Education (GW-CEEE) at the request of the National Center for Educational Statistics (NCES). ELL inclusion and accommodation data for the 2005 TUDA trial for the Grades 4 and 8 NAEP reading and mathematics assessments indicates a pattern of high ELL inclusion rates, wide variability for ELL accommodation rates, and a general lack of correlation between accommodation and inclusion rates, suggesting that factors other than the provision of accommodations may have more explanatory power regarding inclusion decisions for ELLs. Thus, the purpose of the study is to describe and analyze school-based decision-making practices relevant to the inclusion and accommodation of English language learners (ELLs) in the 2005 National Assessment of Educational Progress (NAEP) which was administered as part of the Trial Urban District Assessment (TUDA).

## Justification for the Study

### *The Impact of the Inclusion of ELLs on NAEP Statistics*

Commonly referred to as “the Nation’s Report Card,” NAEP is a content-area assessment administered to a representative sample of K-12 public and private school students from across the nation. NAEP is intended to allow comparisons of student achievement from year-to-year, jurisdiction-by-jurisdiction, and among sub-groups (National Center for Education Statistics, 2006e, p. 1). In order to make these comparisons, the composition of the sample for NAEP is of prime importance. The overall sample size for NAEP has increased since 2002, which in turn has resulted in greater precision (in the form of smaller standard errors<sup>1</sup>) in NAEP data on student achievement (Haertl, 2003, p. 1). As a result, discussions concerning the composition of the NAEP sample have focused on “effects of exclusions [sic] [since they appear to] loom much larger relative to other sources of error” (Haertel, 2003, p. 7). Exclusion effects impact two sub-groups within the NAEP sample more than others, ELLs and students with disabilities (SDs). In particular, the improper exclusion of ELLs and students with disabilities can result in the presence of one of two factors that introduce a statistically significant degree of bias to all NAEP statistics: (a) A reduction in sample size or (b) an increase in variability among sampled students – either of which can “affect conclusions [concerning] changes over time, contrasts among jurisdictions, or gaps among subgroups” (Haertel, 2003, p. 15).

### *Background on NAEP Efforts to Increase the Inclusion of ELLs*

NAEP officials have been concerned with the inclusion of ELLs since the early 1990s when the rate of exclusions for ELLs in the 1992 and 1994 NAEP administrations were as high as 44 percent (Mazzeo, Carlson, Voelkl, & Lutkus, 2000, p. 5). In the mid-1990s the NAEP inclusion policies underwent significant modification to broaden participation among special needs students – i.e., ELLs and SDs (National Center for Education Statistics, 2006a). The new NAEP policy was designed to “ensure that all selected students who are capable of participating meaningfully in the assessment [NAEP] are assessed” (National Center for Education Statistics, 2006a, p. 3). As part of the strategy to raise inclusion rates among ELLs (and SDs), as well as to promote sample consistency of ELL subgroup samples, in the mid-1990s NAEP officials designed a number of administration procedures and supporting documents for school-based

---

<sup>1</sup> “The standard error summarizes the degree of uncertainty in the corresponding statistic” – e.g., mean scores or percentages of students at or above NAEP achievement levels (Haertl, 2003, p. 4).

decision makers when making decisions about including ELLs (National Center for Education Statistics, 2006a).

As a strategy to ensure inclusion, NAEP also began to allow test accommodations. Accommodations are widely used for the assessment of students who, because of limited proficiency in English, or physical or cognitive disabilities, are unable to participate meaningfully in state assessments without adjustments to the language or administration of the test (Rivera et al., 2006)<sup>2</sup>. The provision of accommodations to ELLs and SDs during NAEP were phased in using split administrations of the 1996 and 2000 NAEP in mathematics and the 1998 NAEP in reading – i.e., one portion of the sample of eligible ELLs and SDs were offered accommodations; the other portion of ELLs and SDs were not. The split administrations were given in order to maintain data trends from past administrations while introducing new trend lines for years in which accommodations were allowed and “to examine the feasibility of accommodating special-needs students” (National Center for Education Statistics, 2006a, p. 4). Beginning with the 2002 administration of NAEP, accommodations were offered for eligible ELLs and SDs for *all* content areas in which NAEP was administered (National Center for Education Statistics, 2006a). As a result, by 2005, the year in which this study took place, accommodations could be offered by decision makers to all eligible ELLs.

Between the 1998-2005 administrations of NAEP, NAEP officials used the following set of administration procedures and materials to assist school-based decision makers when making inclusion and accommodation decisions for limited-English-proficient (LEP) students – i.e., ELLs. Key text from the NAEP document *Inclusion of Limited-English-Proficient Students*<sup>3</sup> in NAEP follows<sup>4</sup>.

---

<sup>2</sup> In their review of state policies on inclusion and accommodation, Rivera et al. (2006) found that state assessment policies could be based on one of two frameworks: (1) an ELLs-responsive framework or (2) a disabilities framework.

<sup>3</sup> ELL is the term used in NAEP 2005 reports; LEP [Limited English Proficient] was used before 2005 (National Center for Education Statistics, 2006a, p. 1).

<sup>4</sup> NAEP field staff responsibilities are outlined as well in the *NAEP 2005 Assessment Administrator Manual* (Westat, 2004). Among the NAEP field staff’s responsibilities are selection of schools and students within these schools to participate in the NAEP, guidance of school personnel to ensure that students are included appropriately in the test, and administration of the NAEP.

### **NAEP 2005 Inclusion and Accommodation Guidelines for ELLs**

Decisions about inclusion and accommodations should be made in consultation with knowledgeable school staff, such as ESL teachers, and classroom teachers. The staff member most knowledgeable about each student's English proficiency should first complete the LEP Questionnaire and then return the completed LEP Questionnaire and recommendations for participation to the school coordinator [the primary NAEP contact at the school]. We ask this be done two weeks before the preassessment visit, which occurs about 2 weeks before the assessment date so the NAEP field staff [i.e., personnel from a private testing company hired by NCES] can make arrangements to provide the appropriate accommodations.

NAEP provides the following tools to help knowledgeable school staff decide whether to include a LEP student without accommodations, with accommodations, or to exclude the student: (1) the LEP Questionnaire (which collects information about the student's educational experiences, (2) Criteria for Inclusion (which describes NAEP's policy on inclusion), and (3) a Decision Tree (which is used in conjunction with the LEP Questionnaire to guide the inclusion decision).

NAEP strives to include as many students as possible with appropriate accommodations. Decisions about inclusion and accommodations should be made in consultation with knowledgeable school staff, such as ESL teachers, and classroom teachers. Decisionmakers [sic] should note that NAEP does not provide individual scores, and group groups do not have an effect on state-based performance initiatives or accountability standards. Decisionmakers should:

1. Consult decision tree in conjunction with Question 14 of LEP Questionnaire "when there is doubt about the assessability of a student."
2. Exclude only those students who (a) have less than 3 years of academic instruction in English and (b) cannot demonstrate their knowledge of subject being assessed without an accommodation.
3. Use Questions 12 and 13 of LEP Questionnaire to provide information on type of accommodation student receives on state or other assessments. [Each eligible ELL is then assigned accommodations that (a) the student has used on state assessment and (b) is approved for use on NAEP.]

Table 1 shows the accommodations allowed during the 2002-2005 Reading and Mathematics administrations of NAEP (National Center for Education Statistics, 2006f, 2006g). Indirect linguistic support accommodations were provided for both the Reading and Mathematics administrations of NAEP, while direct linguistic support accommodations were provided only for the Mathematics administration of NAEP. Indirect linguistic support accommodations involve adjustments to the conditions under which ELLs take an assessment. Direct linguistic

support accommodations involve adjustments to the language of the test (Rivera et al., 2006, p. 48).

Table 1. Frequently Provided Accommodations for ELLs for the 2005 NAEP

	Reading NAEP	Mathematics NAEP
Direct Linguistic Support	<ul style="list-style-type: none"> <li>• [none]</li> </ul>	<ul style="list-style-type: none"> <li>• Directions read aloud in English or presented by audiotape*</li> <li>• Bilingual dictionary without definitions**</li> <li>• Passages, other stimulus materials, or test questions read aloud in English or presented by audiotape</li> <li>• Bilingual version of test (Spanish/English)</li> </ul>
Indirect Linguistic Support	<ul style="list-style-type: none"> <li>• Small group</li> <li>• One-on-one (tested individually)</li> <li>• Extended time</li> <li>• Preferential seating</li> </ul>	<ul style="list-style-type: none"> <li>• Small group</li> <li>• One-on-one (tested individually)</li> <li>• Extended time</li> <li>• Preferential seating</li> </ul>

\*Standard NAEP practice, not considered an accommodation.

\*\* Not provided by NAEP, but school, district, or state may provide after fulfilling NAEP security requirements.

Source: National Center for Education Statistics (2006f, 2006g)

In summary, between 1998 and 2005 (with 2005 being the year in which this study is situated), NAEP offered a set of administration procedures and materials to assist school personnel when making inclusion and accommodation decisions for ELLs. Since 2005, NAEP officials have twice revised the ELL inclusion and accommodation decision-making procedures used by school-based decision makers. Thus this study offers a point-in-time perspective of the school-based inclusion and accommodations decision-making practices.

### *The Trial Urban District Assessment Program*

To examine issues related to the inclusion and accommodation of ELLs in NAEP, the current study analyzes a sample of districts who participated in the 2005 TUDA program for NAEP. Initiated in 2002, TUDA is a multi-year research project to examine how public school students in selected large urban districts participate in the administration of NAEP. Six school districts participated in the first NAEP TUDA trial in writing and reading: the Atlanta City School District, City of Chicago School District 299, District of Columbia Public Schools, Houston Independent School District, Los Angeles Unified School District, and New York City Public Schools. Ten districts participated in the second trial, including the original six districts as well as four new districts (Boston School District, Charlotte-Mecklenburg Schools, Cleveland Municipal School District, and San Diego Unified School District), each of which administered the NAEP reading and mathematics assessments. In 2005, the third trial of TUDA, 11 districts (the ten 2003 districts plus Austin Independent School District) participated, administering the NAEP reading, mathematics, and science assessments (National Center for Education Statistics,

2006b).<sup>5</sup> These 11 urban districts enroll, on average, nearly twice the average number of ELLs in Grades 4 and 8 (National Center for Education Statistics, 2006c, 2006d, 2006e), making them particularly suitable for exploring the inclusion and accommodation of ELLs in NAEP.

The following section describes our analysis of NCES (National Center for Education Statistics, 2006c, 2006d, 2006e) data from NAEP 2005. The section concludes with a discussion of NAEP efforts to include more ELLs and to reduce the variability associated with exclusion effects.

### *2005 ELL Inclusion and Accommodation Rates for NAEP TUDA Project*

In order to generate descriptive statistics for ELL inclusion and accommodation rates within the TUDA study, NCES (National Center for Education Statistics, 2006c, 2006d, 2006e) data from the 2005 NAEP Grade 4 and Grade 8 reading and mathematics assessments were examined. This analysis indicated ELL inclusion rates between 93% and 100%<sup>6</sup> in the 11 TUDA districts<sup>7</sup> (see Table 2). These rates indicate progress toward NAEP's goal to "assess all students selected as part of its sampling process" (National Center for Education Statistics, 2006a).

---

<sup>5</sup> For purposes of brevity, districts will be referred to by city name throughout this study.

<sup>6</sup> Cleveland, where 84% of students identified as ELLs were assessed on the Grade 8 NAEP reading assessment, is considered an outlier.

<sup>7</sup> Inclusion rates for ELLs were calculated by dividing the number of ELLs assessed by the number of students identified as ELLs [See pp. 24-26 in both the *Trial Urban District Assessment: Reading 2005* and the *Trial Urban District Assessment: Mathematics 2005* for a breakdown of the number of students sampled and the percentage of ELLs (National Center for Education Statistics, 2006c, 2006d, 2006e).

Table 2. 2005 NAEP ELL Inclusion Rates in TUDA Districts

District	Grade 4 Reading			Grade 4 Mathematics			Grade 8 Reading			Grade 8 Mathematics		
	Number of students identified as ELLs	Number of ELLs assessed	Of students identified as ELLs, percentage assessed	Number of students identified as ELLs	Number of ELLs assessed	Of students identified as ELLs, percentage assessed	Number of students identified as ELLs	Number of ELLs assessed	Of students identified as ELLs, percentage assessed	Number of students identified as ELLs	Number of ELLs assessed	Of students identified as ELLs, percentage assessed
Atlanta	12	11	93%	24	24	100%	10	10	100%	11	11	100%
Austin	324	314	97%	325	321	99%	192	184	96%	168	163	97%
Boston	266	261	98%	180	177	99%	99	96	97%	110	106	96%
Charlotte	126	125	99%	150	149	99%	112	111	99%	98	97	99%
Chicago	323	319	99%	378	376	99%	114	112	98%	114	111	98%
Cleveland	45	43	95%	40	39	98%	32	27	84%	27	25	93%
District of Columbia	126	124	99%	110	108	99%	57	51	90%	76	73	97%
Houston	612	596	97%	740	737	100%	238	230	97%	255	249	98%
Los Angeles	1176	1169	99%	1134	1128	99%	630	624	99%	646	642	99%
New York City	204	196	96%	240	235	98%	170	159	94%	180	174	97%
San Diego	468	464	99%	504	502	100%	312	305	98%	273	269	98%
Total	3682	3622	97%	3825	3797	99%	1966	1909	96%	1958	1920	97%

Note: Percentages in this table are rounded to the nearest whole number.

Data generated from National Center for Education Statistics (2006f, 2006g)

Notwithstanding the high *inclusion* rates for ELLs on 2005 NAEP reading and mathematics, data indicated that *accommodation* rates for these test administrations were highly variable across TUDA districts. The percentage of ELLs included with accommodations ranged from 0 to 92 % for Grade 4 reading, from 17% to 92% for Grade 4 mathematics, from 0 to 100% for Grade 8 reading, and from 18% to 100% in Grade 8 mathematics (see Table 3).



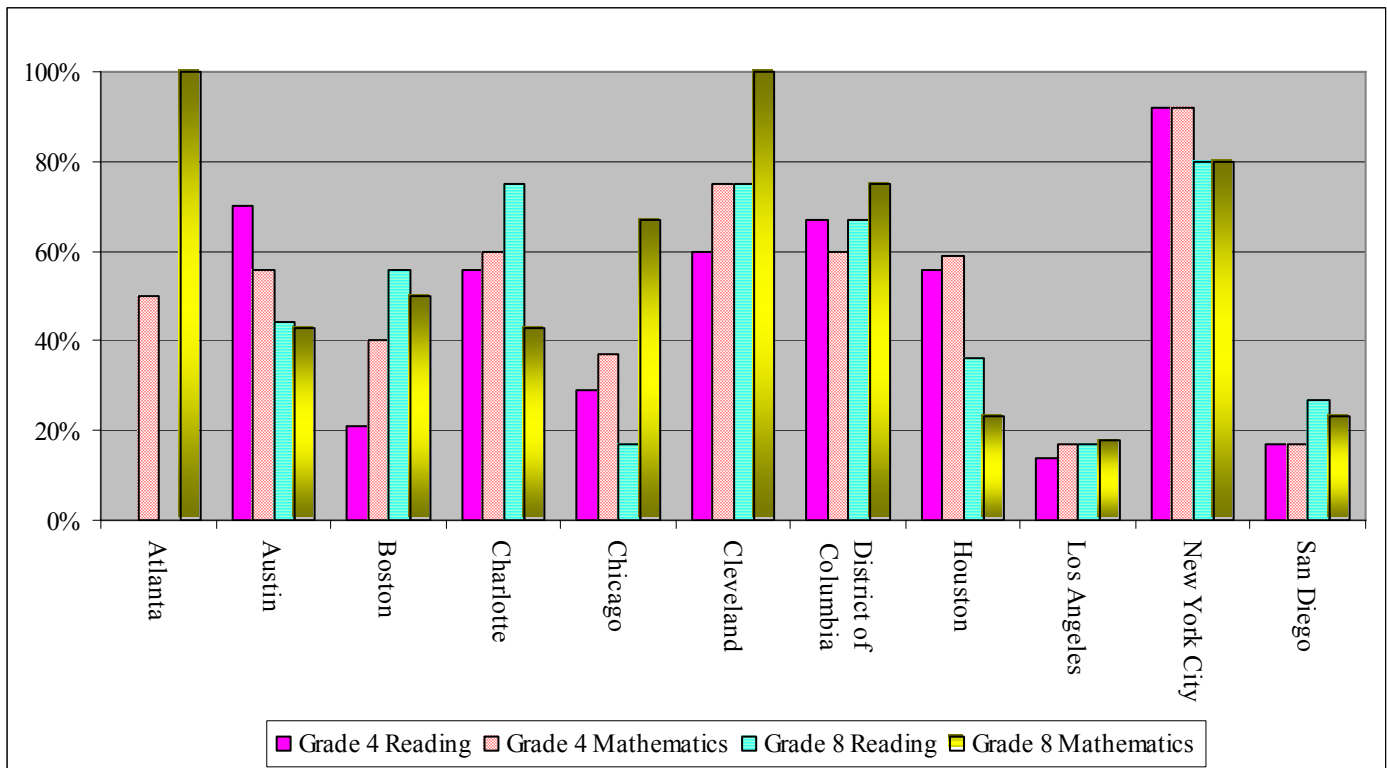
Table 3. 2005 NAEP ELL Accommodation Rates in TUDA Districts

District	Grade 4 Reading				Grade 4 Mathematics				Grade 8 Reading				Grade 8 Mathematics			
	Number of ELLs assessed with accommodations	Percentage ELLs assessed with accommodations	Number of ELLs assessed without accommodations	Percentage ELLs assessed without accommodations	Number of ELLs assessed with accommodations	Percentage ELLs assessed with accommodations	Number of ELLs assessed without accommodations	Percentage ELLs assessed without accommodations	Number of ELLs assessed with accommodations	Percentage ELLs assessed with accommodations	Number of ELLs assessed without accommodations	Percentage ELLs assessed without accommodations	Number of ELLs assessed with accommodations	Percentage ELLs assessed with accommodations	Number of ELLs assessed without accommodations	Percentage ELLs assessed without accommodations
Atlanta	0	0%	12	100%	12	50%	12	50%	0	0%	10	100%	11	100%	0	0%
Austin	218	70%	96	30%	178	56%	143	44%	76	44%	108	56%	67	43%	96	57%
Boston	52	21%	209	79%	69	40%	108	60%	52	56%	44	44%	51	50%	55	50%
Charlotte	69	56%	56	44%	89	60%	60	40%	83	75%	28	25%	41	43%	56	57%
Chicago	91	29%	228	71%	124	37%	252	67%	17	17%	95	83%	73	67%	38	33%
Cleveland	25	60%	18	40%	29	75%	10	25%	19	75%	8	25%	25	100%	0	0%
District of Columbia	82	67%	42	33%	64	60%	44	40%	32	67%	19	33%	54	75%	19	25%
Houston	324	56%	272	44%	437	59%	300	41%	77	36%	153	64%	79	23%	170	67%
Los Angeles	161	14%	1008	86%	183	17%	945	83%	102	17%	522	83%	110	18%	532	82%
New York City	179	92%	17	8%	215	92%	20	8%	125	80%	34	20%	138	80%	36	20%
San Diego	74	17%	390	83%	82	17%	420	83%	110	27%	195	63%	87	23%	182	67%
Total	1275	44%	2348	56%	1483	51%	2314	49%	693	46%	1216	54%	736	58%	1184	42%

Note: Percentages in this table are rounded to the nearest whole number.

Data generated from National Center for Education Statistics (2006c, 2006d, 2006e)

Figure 1 illustrates the variability in the accommodation rates for ELLs across the 11 districts for the 2005 NAEP in reading and mathematics.



Data generated from National Center for Education Statistics (2006c, 2006d, 2006e)

Figure 1. Variability in ELL Accommodation Rates in TUDA Districts on 2005 NAEP

The ELL inclusion and accommodation data from the 2005 administration of NAEP for Grade 4 and Grade 8 NAEP in Reading and Mathematics revealed a puzzling picture: high inclusion rates in combination with widely variable accommodation rates for ELLs. In order to investigate whether the use of accommodations was associated with the increase in ELL inclusion during the 2005 administration of NAEP, correlation coefficients were computed between the rate of accommodations and the rate of inclusion of ELLs on the NAEP reading and mathematics assessments at Grades 4 and 8. Only the correlation for Grade 8 Reading is significant (see Table 4). This finding contradicts the assumption that the use of accommodations for ELLs influences most decisions to include ELLs in NAEP.

Table 4. Correlation Coefficients Between ELL Inclusion Rate and ELL Accommodations Rate on the 2005 NAEP

Grade	Reading	Mathematics
4	.148	.000
8	.622*	.360

\*  $p < .05$

### **Research Questions**

This study is designed to investigate how school personnel in urban districts make decisions regarding the inclusion and accommodation of ELLs in NAEP. The study explores the following research questions in four of the 11 TUDA districts for the 2005 Grade 4 and Grade 8 reading and mathematics NAEP:

1. What factors influenced decisions of school personnel regarding whether to include or exclude ELLs in NAEP?
2. What factors influenced decisions of school personnel regarding the use of accommodations for those ELLs who were included in NAEP?
3. What was the relationship, if any, between the decision to include and the decision to accommodate?

### **Research Methods**

The Methods section of this report is divided into three sections. The first section outlines the sampling plan. This is followed by a description of data collection processes. The third section details the data analysis procedures.

#### *Sampling Plan*

The 11 districts who participated in the 2005 TUDA trial have almost twice the average ELL population in Grade 4 and Grade 8 (National Center for Education Statistics, 2006c, 2006d, 2006e), making them ideal sites for exploring issues related to inclusion and accommodation. As shown in Table 5, three of the four districts selected for this study had relatively high ELL student populations: District 2 (18,000 ELLs for Grade 4 Reading), District 3 (36,000 ELLs for Grade 4 Reading), and District 4 (63,000 ELLs for Grade 4 Reading). The fourth district, District 1, enrolled a proportionally smaller ELL student population (5,000 ELLs for Grade 4 Reading). Other districts with higher ELL populations were already participating in other TUDA studies.

Table 5. ELL Participation in NAEP in the Four Sample TUDA Districts

District	Grade 4 Reading			Grade 4 Mathematics			Grade 8 Reading			Grade 8 Mathematics		
	Total student population	Percentage of students who participated in NAEP	Ells as a percentage of all students who participated in NAEP	Total student population	Percentage of students who participated in NAEP	Ells as a percentage of all students who participated in NAEP	Total student population	Percentage of students who participated in NAEP	Ells as a percentage of all students who participated in NAEP	Total student population	Percentage of students who participated in NAEP	Ells as a percentage of all students who participated in NAEP
District 1	5000	94%	14%	5000	93%	15%	5000	91%	9%	5000	91%	10%
District 2	36000	95%	17%	36000	95%	18%	35000	95%	6%	35000	93%	6%
District 3	18000	95%	36%	18000	96%	37%	14000	88%	14%	14000	88%	15%
District 4	63000	93%	56%	63000	93%	54%	50000	89%	35%	50000	89%	34%
Total (TUDA)	250000	93%	20%	250000	94%	20%	210000	88%	12%	210000	88%	11%

Source: National Center for Education Statistics (2006c, 2006d, 2006e)

Sample districts provided a variety of language support programs for ELLs. District 1 and District 4, in accordance with state law, provided one year of English as a second language support before moving ELLs into the mainstream classroom. District 2 and District 3 provided bilingual instruction for ELLs who were native Spanish speakers. Limited English proficient students in District 2 were enrolled in a transitional bilingual education program for three years, or until the student had achieved a level of English language proficiency necessary for successful performance in classes in which instruction is offered only in English. District 3 provided several models of bilingual education, in which students received academic instruction primarily in Spanish during the early grades, with a gradual transition to English by Grade 5 for most subjects.

To select schools for the study, the GW-CEEE research team obtained sampling data for the four districts selected from the company hired to serve as NAEP field staff during the 2005 NAEP. Using the data, the team ranked schools according to the number of ELLs sampled. Based on the ranking, initial lists of 30 schools with the highest number of students identified as ELLs were compiled from which the team selected the sample for the study. (Fifteen of the 30 schools enrolled Grade 4 students and the other 15 enrolled Grade 8 students)<sup>8</sup>.

After the lists of 30 schools per district were generated, the research team narrowed the selection further by consulting NAEP administration schedules, which provide information regarding students' ELL status, NAEP test booklet ID number, and, where applicable, type of accommodation used during assessment. Schools with large numbers of ELLs and which, according to the NAEP administration schedule, provided accommodations to ELLs were favored. The goal of the sampling plan was to examine inclusion and accommodation decision-making for ELLs under four conditions: (1) exclusion of ELLs from NAEP reading, (2) inclusion of ELLs (with or without accommodation) in NAEP reading, (3) exclusion of ELLs from NAEP mathematics, and (4) inclusion of ELLs (with or without accommodation) in NAEP mathematics.

Finally, the research team identified 12 schools in each district for possible participation in the study: four elementary schools and two alternate elementary schools (Grade 4); four middle schools and two alternate middle schools (Grade 8). Letters were sent to the principals of these schools requesting the participation of their staff in the study. The research team obtained permission to conduct interviews at 29 schools enrolling students at Grade 4 and/or Grade 8 (two schools enrolled students at both grade levels). The team was able to conduct interviews at all but one Grade 8 school in District 1. Figure 2 illustrates the sampling plan for the study.

---

<sup>8</sup> A comparison of the data for the five dimensions of inclusion and accommodation decision-making practices for ELLs for the Grade 4 and Grade 8 schools does not reveal a clear pattern of difference in the decision-making practices for Grade 4 and Grade 8 ELLs. The Grade 4 and Grade 8 school data can be examined by reviewing Tables E1-E6, Tables E8-E12, Tables E14-E18, and Tables E20-E24 in Appendix C.

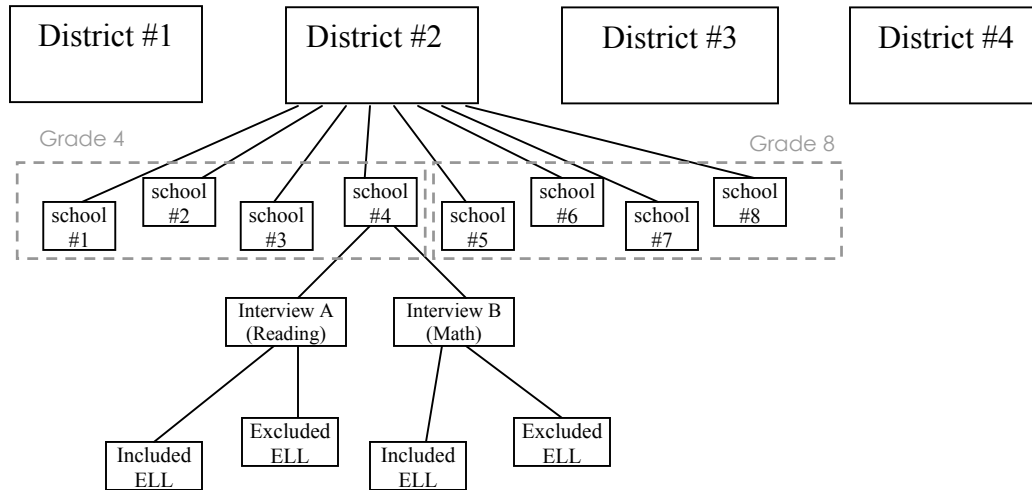


Figure 2. Sampling Plan.

The small sample size for both districts and schools is not statistically representative of all of the schools within the 11 TUDA districts and therefore, does not allow for statistical generalization. However, because the sample was selected purposively, it does allow for *thematic generalization* based on patterns found in the data.

### Data Collection

The primary source of data for this study was phone interviews with school-based decision makers. To develop questions for these interviews and to analyze responses NAEP testing materials, researchers used the student administrative schedule and NAEP test administration schedule (Westat, 2004), and NAEP tools (the LEP Questionnaire, the NAEP decision tree, and the NAEP Criteria for Inclusion).

The school principal was contacted via telephone and asked to briefly describe the decision-making process for including and accommodating ELLs at his or her school and to help the researchers identify the actual decision makers who had made inclusion and accommodation decisions for which ELLs would participate in the Reading and Mathematics NAEPs. After the principal identified the decision makers for each school, interviews were scheduled with these individuals. It should be noted that in almost every school, personnel interviewed were school coordinators or administrators who oversaw decision-making and administration of NAEP rather than teachers who worked closely with ELLs and completed the LEP Questionnaire. The research team used each school's NAEP administration schedule to select ELLs about whom the respondents would be interviewed. The selected ELLs satisfied one of four conditions examined in the study: (1) included in NAEP reading, (2) excluded from NAEP reading, (3) included in NAEP mathematics, and (4) excluded from NAEP mathematics. NAEP test booklet numbers, rather than student names, were used during interviews to identify students. Finally, to ensure that the respondent could discuss decision-making regarding accommodations, the research team attempted to identify from the NAEP administration schedule ELLs who had been included with accommodations.

The interview protocol (see Appendix A) consisted of four parts: (1) a series of introductory questions to identify the role of the decision maker at the school, (2) questions regarding the decision-making process for including a selected student in NAEP, (3) questions

regarding the decision-making process for excluding a selected student in NAEP, and (4) concluding questions regarding differences in the process of including and accommodating ELLs in NAEP as opposed to state or district assessments. Each interview required approximately 40 minutes to complete. Respondents were asked to have at the ready completed copies of the administration schedule as well as NAEP testing materials and policy guidelines used to support the decision-making process. Following Maxwell (1996), hypotheses that emerged during interviews were noted in interviewer memos to be included in the analysis.

### *Data Analysis*

Data analysis proceeded in three stages. **Stage One** focused on the organization and packaging of interview data so that it could be accessed in any one of following three formats: Access database, Excel spreadsheet, and single document transcripts (for NVivo analysis). The unit of analysis for the study, i.e., the case, was defined as a single district. The team explored the data in order to understand the extent to which contextual factors (e.g., instructional programs in which ELLs were enrolled; district and state decision-making policies) influenced differences among the decision-making practices employed across the four districts.

During **Stage Two**, the single-document interview transcripts for each respondent were labeled with the appropriate respondent, school, and district attributes. Transcript data were entered into NVivo software and coded based on codes identified by Rivera et al. (2006). These codes included types of decision makers, inclusion/exclusion criteria, accommodations criteria, and use of NAEP or other guidelines (see Appendix B).

Data analysis indicated five themes (or dimensions) related to the decision-making process: (a) school personnel involved, (b) involvement of NAEP field staff, (c) NAEP tools used, (d) inclusion criteria used, and (e) accommodations criteria used. Three of the dimensions (school personnel, inclusion criteria, and accommodation criteria) coincided with the framework developed in Rivera et al. (2006). The other two dimensions emerged from discussions with NCES officials regarding the involvement of NAEP field staff and the use of NAEP tools in the decision-making process.

Using within-case analysis methodologies, researchers constructed individual school profiles by categorizing data and displaying relationships among five dimensions of the decision-making process. Within each of these five dimensions, researchers then identified two or more inter-related, operationalized factors associated with inclusion and accommodations decisions. The five dimensions of decision-making and associated operationalized factors which were used to build the district profiles are summarized in Table 6.

Table 6. Operationalized Factors for Each of the Five Dimensions of Decision-Making

Dimension	Operationalized factors
1. School personnel involved	<ul style="list-style-type: none"> <li>• School-based decision-making unit</li> <li>• Decision maker's role or job expertise</li> </ul>
2. Involvement of NAEP field staff	<ul style="list-style-type: none"> <li>• Timing of the decision</li> <li>• Influence of NAEP Field Staff</li> </ul>
3. NAEP tools used	<ul style="list-style-type: none"> <li>• Central document(s) used with decision-making</li> <li>• NAEP tool used</li> </ul>
4. Inclusion criteria used	<ul style="list-style-type: none"> <li>• English language proficiency</li> <li>• Academic performance</li> <li>• Time in the U.S.</li> <li>• Amount of instruction in English</li> </ul>
5. Accommodations criteria used	<ul style="list-style-type: none"> <li>• English language proficiency</li> <li>• Academic performance</li> <li>• Time in the U.S.</li> <li>• Amount of instruction in English</li> <li>• Individual Education Program (IEP)</li> </ul>

The codification of operationalized factors within each of the five dimensions allowed the research team to explore the relationships among factors within each of the five dimensions in order to identify patterns. For example, researchers examined the following questions for each of the five dimensions:

- Dimension 1: School personnel involved. To what extent were decision makers' job roles or areas of expertise associated with similar kinds of decision-making processes (e.g., language acquisition specialist, test coordinator, or school administrator). To what extent did group versus individual decision-making influence this process?
- Dimension 2: Involvement of NAEP field staff. To what extent did the timing of the decision influence the use of NAEP and/or the involvement of NAEP field staff in the decision-making process?
- Dimension 3: NAEP tools used. To what extent were NAEP documents or documents from other sources (e.g., state documents) used as the primary guidance for decisions?
- Dimension 4: Inclusion criteria applied. To what extent were NAEP criteria (e.g., amount of instruction in English and academic performance) applied?
- Dimension 5: Accommodation criteria applied. To what extent were criteria commonly used for students with disabilities applied to ELL decisions?

The exploration of the relationships among the operationalized factors contributed to individual school profiles reflecting the various factors associated with each districts' decision-making process. Researchers then cross-checked the findings from the school profiles with district summaries. The decision-making practices for each school were then summarized in an Excel spreadsheet to construct district summaries (see Appendix C).



For **Stage 3** of the study, researchers followed a three-step case study procedure. These procedures contributed to the validity and reliability of findings by analyzing evidence for each of the research propositions. Following Yin's (2003) methodology, the research team (1) outlined a theory of the concepts being studied prior to data analysis; (2) codified this theory into three research propositions (shown below); and then (3) examined district-level decision-making patterns for correspondence with each of the propositions. Yin suggests that if the data provides strong evidence to support one or more propositions, sufficient strength for the broader theory is ascertained and, the concept may be generalized to a larger universe.

### **Research Propositions**

Decision makers relied on one of the following interpretations of state assessment policy for inclusion and/or accommodation decisions:

- A. ELL-responsive criteria
- B. Students with disabilities (SD)-responsive criteria
- C. Treatment of ELLs similarly to native English-speaking students

Data that supported Research Proposition A would indicate that decision makers consider ELLs' linguistic and socio-cultural needs during high-stakes testing. Rivera et al. (2006) identified the following categories of ELL-responsive criteria: *language-related*, *academic-related*, *time-related*, and *opinion-related* criteria. "*Language-related* criteria ... included language proficiency in English and the native language as well as all educational context, such as program placement and language of instruction. *Time-related* criteria pertain to the length of exposure a student has to an English-speaking academic environment. *Academic-related* criteria relate to students' prior schooling and academic achievement as measured by test performance. *Opinion-related* criteria address teacher and parent inputs to the identification of eligible ELLs for accommodations on state assessments" (p. 82).

Data that supported Research Proposition B would indicate that decision makers did not consider criteria relevant to the specific linguistic needs of ELLs. In their study of state assessment policies, Rivera et al. (2006) found that ELLs and SDs were often grouped together within state policies as "special needs students," "at-risk students," or "special populations," and that state policies often did not distinguish between those accommodations appropriate for ELLs and those appropriate for other students. An SD-responsive approach is based on the assumption that only ELLs who have Individualized Education Plans (IEPs) are eligible for accommodations. Decision makers applying this proposition use disability taxonomy for identifying appropriate accommodations, selecting from the following categories: (1) timing/scheduling accommodations, (2) setting accommodations, (3) presentation accommodations, or (4) response accommodations. An SD-responsive taxonomy does not include accommodations that address ELLs' unique needs during assessment.

Data that supported Research Proposition C would indicate that decision makers interpreted the state assessment policy as requiring equality of treatment for ELLs and native English speakers.

### *Validity and Reliability*

Validity of the data is supported by the sampling and data collection procedures. Inductive analyses (Maxwell, 1996; Miles & Huberman, 1994) provided a means of identifying dimensions of decision-making and generating research questions and hypotheses from the data.

The sampling procedures used in this study followed a case study research strategy and relied on *analytical generalization* to develop or test some broader theory using the data (Yin, 2003, p. 28 and p. 37).

To ensure validity and reliability of the findings, at least three members of the research team triangulated all data analyses. The research team reviewed and compared their notes on (a) participants' exact wording from the taped interviews, (b) the codes and coding constructed during the Stage Two analysis, and (c) the findings generated during the Stage Three analyses. Data and findings that could not be verified across sources were not included.

### *Limitations of the Study*

Due to limitations of the data, it is not possible to determine whether there is a causal association between a school's decision-making practices identified in the current study and the ELL inclusion and accommodation rates for a particular school or district, nor with ELL NAEP outcomes for that school or district. NAEP data on ELL inclusion and accommodation rates are collected at the district level, rather than the school level (National Center for Education Statistics, 2006b). However, the findings from this study can be used to identify general patterns of behaviors that occurred within the four districts. NAEP officials and NAEP field staff representatives may thus find results of this study useful for recognizing possible decision-making patterns they encounter during the ELL inclusion and accommodation process and for developing procedures to improve test administration.

### *A Review of the NAEP Inclusion and Accommodation Process*

For each administration of NAEP within the 11 TUDA districts, a random sample of students, including ELLs, are selected for participation. For each of the schools at which randomly-selected students are located, school coordinator(s) are selected to serve as the primary contact with NAEP field staff and to determine the final list of students who will be participating in NAEP. School coordinator(s) receive an administrative schedule that contains a list of students who have been selected to participate in NAEP. Using NAEP criteria, school coordinators are asked to determine which ELLs are eligible for inclusion and of those ELLs included, which are eligible for accommodations.

To assist with the decision, school coordinators receive three NAEP tools to use as guides: (1) the LEP Questionnaire, (2) the NAEP decision tree, and (3) the NAEP Criteria for Inclusion document. The LEP Questionnaires are to be completed by knowledgeable school staff, such as ESL teachers and classroom teachers. The ESL teacher and/or classroom teachers return the completed LEP Questionnaire and recommendations for participation to the school coordinator. Approximately two weeks before the assessment date, NAEP field staff arranges a time to meet with the school coordinator(s) to go over the recommendations for inclusion and accommodations, answer any questions, and make provisions for any accommodations to be offered. NAEP guidelines suggest the use of the following inclusion and accommodations criteria as part of the decision-making process:

Decision makers should:

1. Consult decision tree in conjunction with Question 14 of LEP Questionnaire "when there is doubt about the assessability of a student."
2. Exclude only those students who (a) have less than 3 years of academic instruction in English and (b) cannot demonstrate their knowledge of subject being assessed without an accommodation.

3. Use Questions 12 and 13 of LEP Questionnaire to provide information on type of accommodation student receives on state or other assessments. [Each eligible ELL is then assigned accommodations that (a) the student has used on state assessment and (b) is approved for use on NAEP.]

### District Findings

This section presents summative data regarding inclusion and accommodation decision-making for the 29 schools in the four districts studied. As shown in Table 7, in 14 of the 29 schools, decision makers were asked to make decisions about Grade 4 ELLs; in 13 of the 29 schools, decision makers were asked to make decision about Grade 8 ELLs. Decision makers at two schools made decisions about both Grade 4 and Grade 8 ELLs.

Table 7. Schools Within the Four Sample TUDA Districts

District	District 1	Grade	District 2	Grade	District 3	Grade	District 4	Grade
School	1. Fisk	4	7. Sandburg	4	14. Dickens	4	22. Tully	4
	2. Rice	4	8. Day	4	15. Hardy	4	23. Tacitus	4
	3. Tiant	4	9. Pronto	4	16. Bronte	4	24. Livy	4
	4. Yazstremski	4 + 8	10. Raleigh	4 + 8	17. Bishop	4	25. Homer	4
	5. Boomer	8	11. Bacon	8	18. Pynchon	8	26. Smithers	8
	6. Lynn	8	12. Drake	8	19. Spenser	8	27. Ovid	8
			13. Middleton	8	20. Chapman	8	28. Josephus	8
					21. Fletcher	8	29. Hesiod	8

The four district summaries provide the reader with a synthesis of the findings in relation to the study's three research questions:

1. What factors influenced decisions of school personnel regarding whether to include or exclude ELLs in NAEP?
2. What factors influenced decisions of school personnel regarding the use of accommodations for those ELLs who were included in NAEP?
3. What was the relationship, if any, between the decision to include and the decision to accommodate?

The findings for each district end with a description of additional procedural factors that influenced both the inclusion and accommodation decision. Appendix C provides detailed district summaries and school profiles.

#### *District 1*

In 2005, District 1 had 57,742 students, 136 schools, and 9,789 students were identified as ELL students (National Center for Education Statistics, 2005). Within this district, ELLs received sheltered English immersion instruction during a temporary transition period of one year, after which all of their academic instruction was given in English.

For NAEP inclusion decisions, decision makers at five of the six District 1 schools sampled for this study reported that they relied on criteria such as the student's English language proficiency, academic performance, and amount of instruction in English. However, due to the complexity of the decision-making process and the variety of information collected through the LEP Questionnaire, no decision makers were able to report exclusive reliance on the two main criteria outlined in the NAEP guidelines (amount of instruction in English and academic

performance). Rather, decision makers said they attempted to use a variety of ELL-responsive criteria, e.g., *language-related*, *time-related*, or *academic-related*. Additionally, because the LEP Questionnaire was based on teacher input, *opinion-related* criterion was used by all decision makers who used the completed LEP Questionnaire as part of their decision-making process. Only one of the six schools (Tiant) included all ELLs, as the decision maker explained, based on the state assessment policy that “everyone must be tested.”

The three schools that reported following the state inclusion criteria had different interpretations. The Tiant decision maker thought the state criteria for inclusion was to include all ELLs. The Boomer decision makers thought the state inclusion criteria called for the use of test results and English language proficiency. A third school, Fisk, who made the inclusion decision during the NAEP pre-assessment visit also relied on state inclusion criteria, but explained, “NAEP criteria are the same as those used for MCAS testing anyway, so of course we did [follow NAEP guidelines].”

Decision makers at two Grade 8 schools (Yazstremski and Lynn) used ELL-responsive criteria when determining whether or not to provide ELLs with accommodations. Three schools (Fisk, Tiant, and Boomer) provided accommodations only to students with IEPs. Finally, at the sixth school (Rice), the decision maker explained that she did not give any accommodations to ELLs because she was following the NAEP requirements listed in the LEP Questionnaire - which asked decision makers to consult state guidelines. According to her, state guidelines did not offer “Spanish help.” Therefore, the bilingual NAEP accommodation was not used. The two schools (Yazstremski and Lynn) that used ELL-responsive criteria for accommodations decisions also tied the inclusion decision to the provision of accommodations, while the other four schools in District 1 did not.

In addition to the inclusion and accommodation criteria used, decision makers reported a number of procedural factors that influenced *both* the inclusion and accommodation decisions for ELLs. These factors reveal the complexity and multiple procedures used among all of the schools sampled, even those following ELL-responsive practices. In District 1, the additional factors that framed the context behind *both* the inclusion and accommodation decisions for ELLs were (1) the involvement of the school administrator in making inclusion and accommodation decisions, (2) the role of NAEP field staff, and (3) the use of NAEP tools.

At the school in which all ELLs were included, the administrator did the decision-making without consultation. At the other five schools, the administrator worked with one or more individuals on staff to make the inclusion decision. At three schools (Rice, Yazstremski, and Boomer), the administrator worked with another person – either a test coordinator or a specialist with background in language acquisition issues. At two schools (Fisk and Lynn), the school administrator set up a committee.

In four of the six schools, NAEP field staff did not play an important role in guiding the decision – except to adjust the inclusion decision on the day of the test due to limited numbers of testing personnel on the day of NAEP administration. At two schools (Fisk and Lynn) ELL inclusion and accommodation decisions were made during the NAEP field visit, and the NAEP field staff simply confirmed the decisions made by the school-based decision makers.

All three NAEP tools were used fully by only two of the six schools sampled in District 1 (Lynn and Rice). At one of the six schools (Yazstremski), the LEP Questionnaire was used to shape the final inclusion and accommodation decision. At the other three schools (Boomer, Tiant, and Fisk); NAEP tools were not used in the final decision – although Boomer decision makers reported relying on the NAEP criteria, but not the LEP Questionnaire.

### *District 2*

In 2005, District 2 enrolled 426,812 students in 634 schools<sup>9</sup> (National Center for Education Statistics, 2005). ELLs in this district who were native Spanish-speakers received academic instruction in their native language for three years (or fewer if the student's English language proficiency was deemed sufficient to allow them to perform successfully in classes in which instruction is conducted only in English).

Decision makers at four District 2 schools said they applied the NAEP criterion of academic performance, and at one school the decision was also based on amount of instruction in English. However, in addition to the NAEP criteria, decision makers at five of the seven schools sampled also relied on students' English language proficiency and time in the U.S. These decision makers told us that they relied on a variety of criteria to determine if it would be appropriate for individual ELLs to be included and/or accommodated in NAEP. In other words, decision makers were aware of and attempting to base their inclusion decisions on ELL-responsive criteria, e.g., language-related, time-related, or academic-related. Additionally, because the LEP Questionnaire was based on teacher input and was used extensively by District 2 decision makers, opinion-related criterion was used by all decision makers who used the completed LEP Questionnaire as part of their decision-making process.

District 2 decision makers did not seem to be influenced by the state inclusion policy (nor did the decision makers report that they thought all ELLs needed to be included). Neither did District 2 decision makers report that the list of pre-selected students was the final list of students for the test. The Day respondent summarized, "The NAEP process differs from the state. For the state, the student must be in program year 3 or higher to take a [standardized] test in English, but we also give it in Spanish. For NAEP, we were given a choice."

At two of the seven schools sampled in District 2 (Bacon and Day), accommodation decisions were made based on the students' English language proficiency. In three schools (Raleigh, Day, and Bacon) respondents indicated they were able to include students because they were able to provide them with accommodations. Because of the availability of the bilingual test booklet, no student was excluded at Day from the mathematics NAEP. However, other students at Day were excluded from taking the reading test because a dual-language booklet was not available. It should be noted that, in at least one instance, the decision maker (at Raleigh) stated that a student who was in the U.S. and bilingual program for less than 2 years "could manage the test with the accommodation of extended time." Because NAEP offered this accommodation, the teacher wanted students to attempt the test. The respondent reiterated, however, that, ordinarily, LEP students do not receive any accommodation unless they are also identified as special education. However, according to District 2 decision makers, the state policy on the use of accommodations is only for students with IEPs. This factor exerted an influence on whether students were accommodated at all. In discussing decision-making for accommodating students participating in NAEP, the most common response (at four schools: Raleigh, Middleton, Pronto, and Sandburg) was that, generally, only special education students were allowed to use accommodations per state guidelines. While explaining that the NAEP and state accommodations guidelines were the same (both intended accommodations only for students with disabilities), the Drake decision maker explained that students were not given any accommodations because NAEP and state guidelines on accommodations only allowed accommodations to students with disabilities. The Drake decision maker added, "However if this

<sup>9</sup> The number of students identified as ELL in 2005 in this district was not available.

were a mandatory assessment, I would look at lower-level testing. But in this case, if I thought the student was capable, the student would take the on-level test.”

Three of seven schools in District 2, Raleigh, Bacon, and Day, linked the accommodation decision to the inclusion decision. Raleigh based its decision on the availability of extended time, while Bacon and Day based their decision to accommodate on the availability of the bilingual test booklet.

In addition to the inclusion and accommodation criteria, decision makers reported a number of procedural factors that influenced both the inclusion and accommodation decisions for ELLs. These factors reveal the complexity and multiplicity of procedures implemented among all of the schools sampled, even those following ELL-responsive practices. In District 2, the additional factors that framed the context behind both the inclusion and accommodation decisions for ELLs were (1) the use of NAEP tools, and (2) the role of NAEP field staff.

District 2 decision makers appear to have read and used the NAEP tools provided to them before the NAEP pre-assessment visit. All three NAEP tools were used at three schools (Sandburg, Raleigh, and Drake). Two of the three NAEP tools were used at Middleton and Day. It appears that only one school, Pronto, used only one NAEP tool, the LEP Questionnaire. Of the three tools distributed by NAEP field staff to inform decision-making in District 2, the LEP Questionnaire was the only one to have been used in all District 2 schools. At each school the questionnaire was completed by the teachers most familiar with the selected students. At only one school (Bacon) the extent to which decision makers relied on the completed LEP Questionnaire, Decision Tree, and NAEP Criteria of Inclusion during decision-making was unclear.

In every school in District 2 except Bacon, respondents indicated that inclusion decisions were made prior to the “visit from NAEP field staff.” NAEP field staff, in general, tended to confirm the decisions that had already been made when they arrived for the pre-assessment visit. At Bacon, once the school received the list of students selected for participation in the NAEP, the principal and “NAEP administrator” (presumably, the NAEP field staff representative) made decisions regarding inclusion of students. It appears that because District 2 decision makers had relied on NAEP tools (i.e., the written directions for completing and applying the LEP Questionnaire, the decision tree and the criteria for inclusion), they were less dependent on the NAEP pre-assessment visit and verbal directions from NAEP field staff for ensuring that NAEP criteria were applied.

### *District 3*

In 2005, District 3 enrolled 208,945 students in 313 schools, of whom 59,483 students were identified as ELLs. Inclusion decisions in this district were based primarily on the amount of time the student had received English language instruction. ELLs who were native-Spanish speakers received most academic instruction in Spanish at Kindergarten, with a gradual shift to approximately half English and Spanish instruction by Grade 5 (some students continued Spanish instruction in some subjects through middle school). Six of the eight schools reported application of criteria similar to “amount of instruction in English.” Five of the six schools said they relied on the criterion “amount of instruction in English” for both inclusion and accommodations decisions. ELLs in these bilingual programs typically received the accommodation of bilingual test booklets. At one school (Dickens), the principal asked the school counselor to include *all* bilingual Grade 4 students because “in Grade 5 everything is in English,” and he believed the students might benefit from practice in taking a test in English. When NAEP field staff told the Dickens staff that NAEP would only be given in English, they

had to decide “how the child would do,” i.e., whether or not the child could meaningfully participate in the NAEP.

On the whole, there was variability regarding which criteria the District 3 decision makers emphasized during the inclusion decision-making process. Due to the complexity of the decision-making process and the variety of information collected through the LEP Questionnaire, no decision makers were able to tell the research team that they relied exclusively on the two main criteria outlined in the NAEP guidelines (amount of instruction in English and academic performance). Instead, decision makers told us that they relied on a variety of criteria to determine if it would be appropriate for individual ELLs to be included and/or accommodated in NAEP. In other words, decision makers said they were aware of and attempting to base their inclusion decisions on ELL-responsive criteria, e.g., language-related, time-related, or academic-related. Additionally, because the LEP Questionnaire was based on teacher input, opinion-related criterion was used by all decision makers who used the completed LEP Questionnaire as part of their decision-making process.

The criterion related to “amount of instruction in English” was also supplemented by data on the student’s English language proficiency and/or the student’s academic performance. For example, Pyncheon decision makers relied on academic performance and time in the US, while Bronte decision makers relied on academic performance in combination with amount of instruction in English for inclusion decisions regarding students listed on the administrative schedule. However, decision makers at Pyncheon and Bronte told us that the NAEP field staff had made the inclusion decision. At Pyncheon, students who were new to the country this year were excluded from the NAEP. If the student’s ESL level was at 2, 3, or 4, and grades were below C, the student was excluded. At Bronte, ELLs were excluded on the administration schedule based on level of instruction. At yet another school (Fletcher) the decision maker used student test scores. Only one decision maker in District 3 (Chapman) told us she had relied on state policy – but even so, applied ELL-responsive criteria (academic performance and time in the U.S.) that reflected their interpretation of state policy. The Chapman respondent indicated that in making NAEP inclusion decisions the school had attempted to “stick with consistency of how we exempt” ELLs and special education students in state assessments. “We assumed that if they could take the TAKS and the Stanford 10, they would take the NAEP.”

Within this district, the assignment of accommodations was more likely to be determined on the basis of criteria related to English language acquisition (English language proficiency, time in U.S., time in LEP program) rather than other criteria such as whether or not the student had an IEP. The availability of the bilingual booklet influenced the accommodations decisions made at Bronte and Bishop. For example, at Bronte, ELLs were given bilingual booklets for the mathematics NAEP. “They [NAEP] did have materials in Spanish so we did accommodate children that were selected.” The Bishop respondent indicated that the availability of a bilingual test booklet in Math made the decision to include certain students easier. However, the lack of bilingual booklets impacted some ELLs at two schools. Because fewer than needed bilingual booklets were available at Hardy and Pyncheon, some ELLs were excluded. The Hardy respondent told us “Some students were excluded because she did not have enough bilingual materials. They [NAEP field staff] are the ones who chose who to test.” But the major perspective at Hardy was that if students were included in NAEP, they did not need to receive accommodations.

There was a relationship between the inclusion and accommodation decision for seven of the eight schools sampled (Dickens, Bronte, Bishop, Pyncheon, Spenser, Hardy, and Chapman). For example, the Bishop decision maker indicated that the availability of a bilingual test booklet

in Math made the decision to include certain students easier. In fact, because of the availability of the bilingual booklet, one school (Chapman), included all students and offered the bilingual booklet to those who were included. In contrast, the lack of bilingual booklets impacted ELLs at two schools (Hardy and Pynchon), resulting in the exclusion of some ELLs. The Dickens decision maker, who had made her inclusion decision based on ELL students' placement in the school's bilingual program, relied on English language accommodations. The Spenser decision maker based the accommodations decision on the use of accommodations in the classroom. The criteria for assignment of accommodations at one other school (Fletcher) were unclear.

As in District 1 and District 2, District 3 decision makers reported a number of procedural factors that influenced both the inclusion and accommodation decisions for ELLs. These factors reveal the complexity and multiplicity of procedures used among all of the schools sampled, even those following ELL-responsive practices. In District 3, the additional factors that framed the context for both the inclusion and accommodation decisions for ELLs were (1) the involvement of the school administrator in making inclusion and accommodation decisions, (2) the use of NAEP tools, and (3) the role of NAEP field staff.

In six of the eight District 3 schools sampled in this study, school-based decisions were completed primarily by a single individual; however at the two remaining schools (Spenser and Bronte), decisions were completed by a Language Proficiency Acquisition Committee (LPAC) composed of three or more people. The schools with a single decision maker said they relied on their school administrators or their test coordinators, while in those schools that used LPACS, decision makers were drawn from a variety of job roles, including school administrator, test coordinators, language acquisition-related staff, and classroom teachers.

NAEP tools and NAEP field staff played an influential role in the inclusion decisions made in District 3. While it appears that decision makers at all eight of the schools sampled in District 3 relied on NAEP documents, especially the LEP Questionnaire, when making their decision, the manner in which the documents were used varied in relation to the involvement of NAEP field staff in their decisions. For example, at the two schools in which NAEP criteria were applied with minimal guidance from the NAEP field staff, decision makers reporting relying on all three NAEP tools. At Bishop and Chapman, decision makers were less reliant on the NAEP field staff for inclusion and accommodations decisions. This lack of reliance on the NAEP field staff was possible because decision makers at these two schools used the LEP Questionnaire in conjunction with the decision tree, and the NAEP criteria document for inclusion and accommodation decisions.

Decision makers at four of the eight schools (Hardy, Bishop, Spenser, and Chapman) made the ELL inclusion and accommodation decision prior to the pre-assessment visit from NAEP field staff. At Bishop, NAEP field staff confirmed the decision that had already been made, but at Hardy, Spenser and Chapman, NAEP field staff adjusted the decision just before the test was administered. At Spenser, the decision was altered because there were not enough assessment booklets. "When they [NAEP field staff] came to give the test and I gave them the list, I knew there were only thirty students to be tested, so I called them down. However, when I started calling names, they didn't have answer documents [for all of the students on the list I had given them]. We lost a lot of time. I had to keep the kids who weren't being tested in a different room to help with the kids who were not going to be tested." But at one school (Chapman), the inclusion decision was adjusted because too many ELLs had been excluded. "When the ladies [NAEP field staff personnel] came, they asked the school decision makers to review the students selected to participate in the mathematics NAEP and include more LEP students." At the third school (Hardy), the Assistant Principal, reported that he thought all ELLs would be tested in their



native languages for both the Reading and Mathematics NAEP. After finding out this was not possible for the Reading NAEP, the NAEP field staff representative asked the Hardy Assistant Principal to consult with the Grade 4 ESL teacher (who consulted with other Grade 4 teachers) to determine which students might meaningfully take the test in English. The teachers could not identify any students that fit this description. In the end, the NAEP field staff representative made the final decision to include and accommodate students.

At four schools in District 3 (Dickens, Bronte, Pynchon, and Fletcher), the inclusion and accommodation decision was made during the NAEP field staff pre-assessment visit. The Dickens respondent, a school counselor, reported, “I passed out the questionnaire and let NAEP use my office as ‘ground zero.’ I was out of the picture. They did it all.” The Bronte decision maker reported, “The people that came out with NAEP told me who they were going to test. I really had no input on that decision. They were the ones who made the decision [based on the LEP Questionnaires that had been completed by the teachers].” At Pynchon, the respondent, a test coordinator reported, “The NAEP field staff representative did most of work and would consult with [me] regarding the ESL students.

#### *District 4*

In 2005, District 4 enrolled 741,367 students in 760 schools, of whom 315,467 of students were identified as ELLs (National Center for Education Statistics, 2005). ELLs in this district were transitioned to the general education classroom after one year of language support, except for schools that had applied for a waiver to continue offering bilingual instruction.

Decision makers who had used the completed LEP Questionnaire to gather information on the students (i.e., Hesiod, Smithers, and Tully), reported they had used English language proficiency as one of the criteria for including ELLs. For example, the Hesiod decision maker reported, “Teachers know each student’s language level because this information is printed in roll books. The respondent speculated that teachers’ decisions were probably fairly easy to make because of their knowledge of the students’ English language development (ELD) levels.” Even though at Homer teachers filled out LEP Questionnaires, the decision maker explained that the LEP Questionnaires had no real role in deciding which students were included or excluded because decisions were based solely on ELD levels. Once the Title I coordinator and a bilingual coordinator received the administration schedule from NAEP, they automatically eliminated (excluded) students with English Language Levels of 1 or 2.

The pressure to include all ELLs in testing seems to have played a role at Tacitus and Livy. In fact, it appears that the Tacitus respondent thought the initial list of pre-selected students for NAEP was the final list of students to be selected. The decision maker at Livy followed the STAR [California’s Standardized Testing and Reporting program] guidelines for the ELL inclusion decision. He said, “It’s a little more practical. It’s a little harder for us to have different criteria (Olson & Goldstein).”

Two schools assigned accommodations to ELLs based on English language proficiency (Ovid and Tully). Five schools assigned accommodations only to students who had an IEP (e.g., Livy, Josephus, Tacitus, Homer, and Hesiod). The Livy decision maker observed that the school worked very hard to include everybody and that, “really, only special education students were offered accommodations” on NAEP. A Josephus student at an intermediate level of English proficiency was provided the accommodation of testing in small group for the mathematics assessment because that student’s IEP had indicated that this accommodation should be used. It was unclear as to how accommodations were assigned at four of the eight schools (Tacitus, Homer, Hesiod, and Smithers). At three of the four schools, students were tested without

accommodations. Decision makers emphasized the influential role of the accommodation guidelines used for the state assessment. The Tully respondent indicated that no ELLs were offered accommodations, “Because they [the ELLs at his school] are sufficiently proficient [in English], they do not need accommodations. They have scaffolding in place.” It was unclear what criteria decision makers at Smithers relied on when assigning accommodations.

At only one of the eight schools sampled in District 4 (Ovid) there was a relationship between the decision to include and the decision to accommodate ELLs. The decision makers, a bilingual coordinator and an ELD coordinator (two language acquisition specialists), reported they had worked with teachers to try “to include as many students as possible” from the list [administrative schedule] NAEP had provided. They chose to include one student, but to offer him additional time, even though he was at a “high intermediate or advanced” level of English language proficiency because the additional time might help improve his academic performance.

As in the other sample districts, District 4 decision makers reported a number of procedural factors that influenced *both* inclusion and accommodation decisions for ELLs. These additional factors included; (1) the involvement of the school administrator in making inclusion and accommodation decisions; (2) the role of NAEP field staff; and (3) the use of NAEP tools.

Decision makers at the majority of the eight schools sampled were test coordinators and language acquisition specialists, not school administrators, and tended to work individually or in pairs. NAEP field staff appears to have played an influential role at some of the schools in District 4, but not others. The respondents at four of the seven schools sampled in District 4 (Tully, Tacitus, Josephus, and Smithers) reported that school personnel had a minimal role in decision-making and adhered to the student lists provided by NAEP. For example, the Tacitus respondents explained that they had met with the respondents prior to testing and went over the packet of NAEP materials. “Testing went smoothly” since “everything was precoded,” and that they “went with whatever NAEP [field staff] recommended.” The Smithers decision makers explained that NAEP had provided a list of pre-selected students and “tested all.” Furthermore, NAEP field staff appears to have confirmed the decisions made by Homer and Livy – even though these decision makers reported minimal reliance on either NAEP criteria or NAEP materials. Homer decision makers did not use the completed LEP questionnaires, but relied instead on ELD levels (specifically the CELDT). However, decision makers at Livy made their decisions prior to the NAEP pre-assessment visit, relying on the state criteria rather than NAEP guidelines. Ovid decision makers told us that NAEP had provided some guidance, but it was clear as to whether the decision had been made prior or during the NAEP pre-assessment visit. The Hesiod decision maker made the inclusion decision without NAEP field staff involvement, but did not spend too much time worrying over the decision since “there isn’t any [accountability]...A list came to me and then I used the teacher questionnaire. We did not consult with anyone to include/exclude.”

While the LEP Questionnaire was used more than any other documents or tools to make the decisions, even then it was not used that often in District 4 in ELL inclusion and accommodation decision-making. Only four of the seven schools that completed the LEP Questionnaire actually used it in decision-making (Tully, Smithers, Josephus, and Hesiod). As a result, at four of the five schools, decision makers who had used the LEP Questionnaire to gather information on the students (i.e., Hesiod, Smithers, and Tully), also told us they had used English language proficiency as one of the criterion inclusion decisions for ELLs. For example, the Hesiod decision maker reported, “Teachers know each student’s language level because this information is printed in roll books. The respondent speculated that teachers’ decisions were probably fairly easy to make because of their knowledge of the students’ ELD levels.” No

District 4 decision makers reported applying the NAEP decision tree and the criteria document. The eighth school (Ovid) did not indicate the extent to which decision makers relied on the LEP Questionnaire, decision tree, or Criteria for Inclusion.

### **Cross-Case Analysis of Decision-Making Practices in the Four Districts**

The district findings presented in the previous section indicate that NAEP guidelines were subject to multiple interpretations within all four districts. Furthermore, no decision makers were able to tell the research team they relied exclusively on the two main criteria outlined in the NAEP guidelines (amount of instruction in English and academic performance).

This section analyzes inclusion and accommodation decision-making practices found across the four TUDA districts and the possible relationship among inclusion and accommodation decision-making practices through the use of a case study research strategy (see the Methods for a description of the process used). Researchers analyzed data for three rival research propositions regarding the criteria decision-makers applied for decisions about inclusion and accommodations of ELLs on the NAEP:

- A. ELL-responsive criteria
- B. Students with disabilities (SD)–responsive criteria
- C. Treatment of ELLs similarly to native English-speaking students

We next describe findings related to these research propositions for the four sample districts. We also examine the patterns of relationships across accommodation and inclusion decision-making.

### *Comparison of Inclusion Decision-Making Practices Across the Four Districts*

An analysis of decision-making practices for inclusion of ELLs on the NAEP indicated that decision-making aligned with two of the three research propositions. Data for sample schools indicated that a majority matched Proposition A (ELL-responsive criteria). A small number of schools matched Proposition C (treatment of ELLs similarly to native English speakers). Schools that matched Proposition C tended to apply a blanket inclusion decision for all of their ELLs. None of the sample schools relied on SD-responsive criteria (Proposition B) to make inclusion decisions.

#### **Research Propositions**

Decision makers relied on one of the following rival interpretations of state assessment policy for inclusion and/or accommodation decisions:

- A. ELL-responsive criteria
- B. Students with disabilities (SD)–responsive criteria
- C. Treatment of ELLs similarly to native English-speaking students

At 26 of the 29 schools, decision makers told us that they had applied ELL-responsive inclusion criteria for ELL participation in NAEP. These criteria included *language-related*, *academic-related*, *time-related*, and *opinion-related* criteria that considered ELLs' linguistic and socio-cultural needs in relation to assessment. In District 1, decision makers at five of the six schools sampled relied on criteria such as the student's English language proficiency, academic performance and amount of instruction in English to exclude ELLs. In District 2, five of seven decision makers used ELL-responsive criteria when making inclusion decisions. Within District 3, inclusion decisions were most often impacted by the amount of time the student had received English language instruction. Six of the eight District 3 schools reported use of the criteria similar to "time in LEP program," but all eight of eight District 3 schools reported the use of one or more ELL-responsive criteria during inclusion decisions. In District 4, at four of the five schools, decision makers who had used the LEP Questionnaire to gather information on the students (i.e., Hesiod, Smithers, and Tully), also told us they had used English language proficiency as one of the criterion for making inclusion decisions for ELLs.

No decision makers included ELLs based on whether the ELL had an IEP (i.e., based their inclusion decision on the SD-responsive framework). If the ELL did have an IEP that student's inclusion was treated as a decision relating to the student's disability, not their classification as an ELL. However, it should be pointed out that, because there were not enough testing personnel to monitor students during the test at one school (Tiant) in District 1, the NAEP field staff representative re-examined the list of students and found five that should have been excluded based on the students' IEPs.

Decision makers at three schools in the four districts told us they had included all ELLs since that was the requirement. In District 1, at only one school (Tiant) did decision makers choose to include all ELLs since the state assessment policy said that "everyone must be tested." In District 4, the pressure to include all ELLs in testing was an important factor that influenced staff in two schools (Tacitus and Livy). The Tacitus decision maker thought the initial list of pre-selected students for NAEP was the final list of students to be selected and did not attempt to exclude any ELLs. The decision maker at one other school (Livy) followed the STAR [California's Standardized Testing and Reporting program] guidelines when making the inclusion decision. He said, "It's a little more practical [to use the state criteria rather than the NAEP criteria – which he perceived to be different]. It's a little harder for us to have different criteria [for each large-scale assessment]."

Table 8. Cross-Case Analysis of Inclusion Decision-Making Practices for ELLs

District	School	A. ELL-responsive criteria	B. SD-responsive criteria	C. Equal treatment with native English speaking students
District 1	1. Fisk	x		
	2. Rice	x		
	3. Tiant			x
	4. Yazstremski	x		
	5. Boomer	x		
	6. Lynn	x		
	<b>Subtotal</b>	<b>5</b>	<b>0</b>	<b>1</b>
District 2	7. Sandburg	x		
	8. Middleton	x		
	9. Pronto	x		
	10. Day	x		
	11. Raleigh	x		
	12. Bacon	x		
	13. Drake	x		
	<b>Subtotal</b>	<b>8</b>	<b>0</b>	<b>0</b>
District 3	14. Dickens	x		
	15. Hardy	x		
	16. Bronte	x		
	17. Bishop	x		
	18. Pynchon	x		
	19. Spenser	x		
	20. Chapman	x		
	21. Fletcher	x		
	<b>Subtotal</b>	<b>8</b>	<b>0</b>	<b>0</b>
District 4	22. Tully	x		
	23. Tacitus			x
	24. Livy			x
	25. Homer	x		
	26. Smithers	x		
	27. Ovid	x		
	28. Josephus	x		
	29. Hesiod	x		
	<b>Subtotal</b>	<b>6</b>	<b>0</b>	<b>2</b>
	<b>Total</b>	<b>26</b>	<b>0</b>	<b>3</b>

*Comparison of Accommodation Decision-Making Practices Across the Four Districts*

An analysis of accommodation data from decision makers at 29 sample schools in four districts indicated a different pattern from the inclusion analysis. As shown in Table 9, two of the three research propositions were predominant in our data. Nearly half of the sample schools corresponded to Proposition A (ELL-responsive criteria). Data for another large group of schools matched Proposition B, indicating that decision makers relied on criteria designed for students with disabilities when making decisions about accommodating ELLs on the NAEP. A small number of sample schools provided no accommodations to ELLs. This behavior matches Proposition C, suggesting that decision-makers thought ELLs should be treated the same as native English-speaking students.

At 13 of the 29 schools, decision makers told us that they had applied ELL-responsive accommodations criteria. In District 1, decision makers at two Grade 8 schools (Yazstremski and Lynn) assigned accommodations based on ELL-responsive criteria. At three of the seven schools sampled in District 2 (Bacon, Raleigh, and Day), accommodations decisions were made based on the students' English language proficiency. Seven of eight schools in District 3 applied a range of ELL-responsive criteria when determining whether to accommodate ELLs. Two schools in District 4 assigned accommodations to ELLs based on English language proficiency (Ovid and Tully).

Decision makers at 12 of the 29 schools used an SD-responsive framework when making ELL accommodation decisions, specifically by allowing accommodations only to ELLs designated with an IEP. For example, in District 1, three schools (Fisk, Tiant, and Boomer) only provided accommodations to ELLs with IEPs. In District 2, decision makers at three schools (Middleton, Pronto, and Sandburg) explained that, generally, only special education students were allowed to use accommodations on large-scale tests (e.g., per state guidelines).

Within District 4, the assignment of accommodations at most schools appeared to be based on ELL-responsive criteria related to English language acquisition (English language proficiency, time in U.S., and time in LEP program). Five District 4 schools, however, followed Proposition B, assigning accommodations only to students who had an IEP (e.g., Livy, Josephus, Tacitus, Homer, and Hesiod). The Livy decision maker observed that, "really, only special education students were offered accommodations" on NAEP. One Josephus student at an intermediate level of English proficiency was provided the accommodation of being tested in a small group during the NAEP mathematics test because that student's IEP had indicated that this accommodation should be used.

At three schools across the sample districts, decision makers told us they had accommodated no ELLs in accordance with their interpretation of state or local policy. In District 1 the Rice respondent explained that she did not assign any accommodations to ELLs because she was following the NAEP requirements listed in the LEP Questionnaire to consult state guidelines, which she interpreted as not allowing "Spanish help." In District 4, the Tully respondent indicated that no ELLs were offered accommodations, explaining, "Because [the ELLs at his school] are sufficiently proficient [in English], they do not need accommodations. They have scaffolding in place."

Table 9. Cross-Case Analysis of Accommodation Decision-Making for ELLs

District	School	A. Applied criteria intended for ELLs	B. Applied criteria intended for SDs	C. Applied criteria intended for native English speaking students	Unclear/ Do not know
District 1	1. Fisk		x		
	2. Rice			x	
	3. Tiant		x		
	4. Yazstremski	x			
	5. Boomer		x		
	6. Lynn	x			
	<b>Subtotal</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>0</b>
District 2	7. Sandburg		x		
	8. Middleton		x		
	9. Pronto		x		
	10. Day	x			
	11. Raleigh	x			
	12. Bacon	x			
	13. Drake			x	
	<b>Subtotal</b>	<b>3</b>	<b>3</b>	<b>1</b>	<b>0</b>
District 3	14. Dickens	x			
	15. Hardy	x			
	16. Bronte	x			
	17. Bishop	x			
	18. Pynchon	x			
	19. Spenser	x			
	20. Chapman	x			
	21. Fletcher				x
	<b>Subtotal</b>	<b>7</b>	<b>0</b>	<b>0</b>	<b>1</b>
District 4	22. Tully			x	
	23. Tacitus		x		
	24. Livy		x		
	25. Homer		x		
	26. Smithers				x
	27. Ovid	x			
	28. Josephus		x		
	29. Hesiod		x		
	<b>Subtotal</b>	<b>1</b>	<b>5</b>	<b>1</b>	<b>1</b>
	<b>Total</b>	<b>13</b>	<b>12</b>	<b>3</b>	<b>2</b>

*Cross-Case Analysis of the Relationship between Inclusion Decision-Making and Accommodation Decision-Making*

To compare the relationship between ELL inclusion and accommodation decisions for the four districts, the research team examined patterns of inclusion and accommodation decisions in relation to the three research propositions. The analysis indicated that in a majority of schools in three of the four districts, there was no relationship between accommodations and inclusion decisions by decision makers. In only one district, District 3, did a majority of decision makers report that their inclusion decisions were influenced by the availability of accommodations.

In District 4, data indicated a relationship between the decision to include and the decision to accommodate at only one school (Ovid). The decision makers at this school, a bilingual coordinator and an ELD coordinator (two language acquisition specialists), worked with teachers to try “to include as many students as possible” from the administrative schedule provided by NAEP field staff. They chose to include one student, but to offer him additional time, even though he was at a “high intermediate or advanced” level of English language proficiency, with the rationale that the additional time might help improve his academic performance. Decision makers at two of six District 1 schools (Yazstremski and Lynn) reported that they had tied the inclusion decision to the provision of accommodations, while there was no evidence of this relationship at the other four schools. In District 2, only two of the seven schools (Day and Bacon) connected the accommodation decision to the inclusion decision, but ELLs in the remaining schools were provided no accommodations, regardless of decision makers’ stated rationale to include. In contrast, decision makers in six of the eight District 3 schools described a relationship between the decision to include and the decision to accommodate – primarily, they said, because of the availability of the bilingual booklet, an accommodation that seemed appropriate for students enrolled in bilingual programs.

Figure 3 illustrates the variable patterns in ELL inclusion and accommodation decision-making across the four districts. This concept map graphically represents inclusion and accommodation decisions for ELLs in sample schools in relationship to the three research propositions. As indicated in this graphic, three of the four patterns (I, II, and III) originate with the application of ELL-responsive criteria for inclusion decisions. For accommodations decisions, these patterns become more variable, diverging into each of the 3 research propositions. The predominant patterns (I and II, followed by 25 of the 29 schools) are indicated by bold arrows, while weaker patterns (III and IV, followed by only 4 of the 29 schools) are indicated by unbolded arrows.



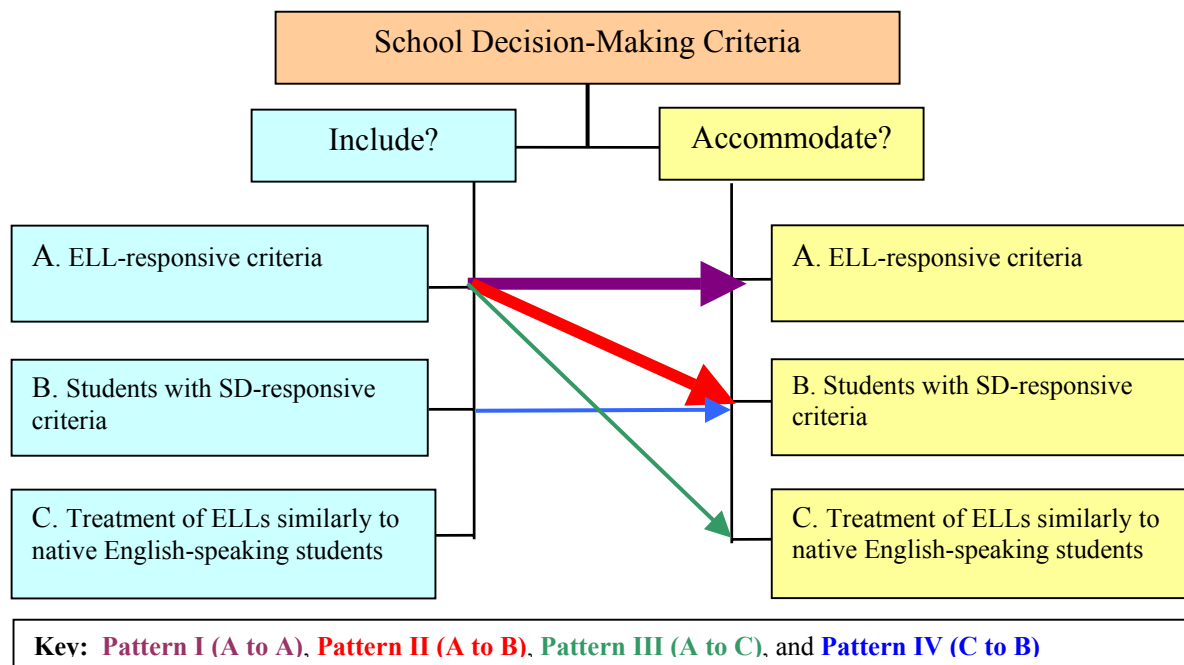


Figure 3. Patterns of Decision-Making in Relation to Three Research Propositions.

In summary, data indicated the following four patterns of inclusion and accommodation decision-making for ELLs in NAEP:

- Pattern I: applied ELL-responsive criteria to make both inclusion and accommodation decisions
- Pattern II: applied ELL-responsive criteria for inclusion decisions, but applied SD-responsive criteria (e.g., student had an IEP) for accommodations decisions
- Pattern III: applied ELL-responsive criteria for inclusion decisions, but offered no accommodations
- Pattern IV: included all ELLs, but applied SD-responsive criteria (e.g., student had an IEP) for accommodations decisions

Table 10 illustrates the decision-making patterns for inclusion and accommodation of ELLs. Pattern I decision-making applied ELL-responsive criteria consistently across inclusion and accommodations decisions. The majority (6 of 8 schools) of sampled decision makers in District 3 followed Pattern I decision-making practices. One of the eight schools applied ELL-responsive criteria for inclusion but offered no accommodations to ELLs (Pattern III). Another one of the eight schools applied ELL-responsive criteria for inclusion, but data were unclear regarding accommodation decision-making.

Decision makers in District 2 were evenly split between Pattern I and Pattern II decision-making practices. While eight of eight District 2 schools sampled used ELL-responsive criteria when including ELLs, they were just as likely to follow Pattern I decision-making practices by relying on criteria for students with disabilities and assigning accommodations only to students

with IEPs (3 of 7 schools) as they were to follow Pattern II decision-making practices by assigning accommodations using ELL-responsive criteria (3 of 7 schools).

Decision makers in District 1 and District 4 were evenly spread among the three patterns when making accommodation decisions -- use of ELL-responsive criteria for inclusion and accommodation decisions for ELLs (Pattern I), use of ELL-responsive criteria for inclusion but assigning accommodations to students with IEPs (Pattern II), and use of ELL-responsive criteria for inclusion decisions, but assigning no accommodations (Pattern III). As such, there was no clear pattern of inclusion and accommodation decision-making in District 4 or District 1 – a sign that multiple interpretations, rather than a district-wide interpretation of NAEP guidelines – was being implemented. Five of six District 1 schools sampled applied ELL-responsive criteria when including or excluding ELLs. Two of the five schools relied on criteria for students with disabilities and assigned accommodations only to students with IEPs; Another two of the five schools assigned accommodations using ELL-responsive criteria and one of the five assigned no accommodations. One school in District 1 included all ELLs, but only offered accommodations to students with IEPs. Six of seven District 4 schools sampled applied ELL-responsive criteria when including/excluding ELLs. These schools were slightly more likely to follow criteria for students with disabilities and assign accommodations only to students with IEPs (3 of 7 schools), rather than assigning accommodations using ELL-responsive criteria (1 of 7 schools) or assigning no accommodations (1 of 7 schools). Two schools in Los Angeles included all ELLs, but only offered accommodations to students with IEPs. In one of the eight schools, decision makers applied ELL-responsive criteria when making inclusion decisions. However, it was unclear from interviews which type of accommodation decision-making practices were applied.

Table 10. Patterns and Variations of Inclusion and Accommodation Decision-Making

District	School	Pattern I. Included and accommodated ELLs based on ELLs-responsive criteria	Pattern II. Applied ELL- responsive criteria to make inclusion decisions and only accommodated students with IEPs	Pattern III. Applied ELL- responsive criteria to make inclusion decisions, but offered no accommodations	Pattern IV Included all ELLs and accommodated only students with IEPs	Unclear/ Do not know
District 1	1. Fisk		x			
	2. Rice			x		
	3. Tiant				x	
	4. Yazstremski	x				
	5. Boomer		x			
	6. Lynn	x				
	<b>Subtotal</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>0</b>
District 2	7. Sandburg		x			
	8. Middleton		x			
	9. Pronto		x			
	10. Day	x				
	11. Raleigh	x				
	12. Bacon	x				
	13. Drake			x		
	<b>Subtotal</b>	<b>3</b>	<b>3</b>	<b>1</b>	<b>0</b>	<b>0</b>
District 3	14. Dickens	x				
	15. Hardy	x				
	16. Bronte	x				
	17. Bishop	x				
	18. Pynchon	x				
	19. Spenser	x				
	20. Chapman	x				
	21. Fletcher					x
	<b>Subtotal</b>	<b>7</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>
District 4	22. Tully			x		
	23. Tacitus				x	
	24. Livy				x	
	25. Homer		x			
	26. Smithers					x
	27. Ovid	x				
	28. Josephus		x			
	29. Hesiod		x			
	<b>Subtotal</b>	<b>1</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>1</b>
	<b>Total</b>	<b>13</b>	<b>8</b>	<b>3</b>	<b>2</b>	<b>2</b>

In conclusion, the cross-case analysis indicated that schools followed four decision-making patterns for inclusion and accommodation of ELLs on the NAEP. A discussion of the implications of these four patterns follows.

### Discussion

The wide variability of interpretation of the criteria for inclusion and accommodations decisions suggests the limitations of school decision-making practices regarding ELL assessment. This variability suggests that decision makers relied on differing conceptual frameworks for selecting the factors that might impact ELLs' abilities to participate in NAEP. This finding also suggests a need for clearer guidance from NAEP in order to assure that school decision makers understand the issues underlying high stakes testing for ELLs and make inclusion and accommodations decisions that are both consistent and valid for ELLs.

For inclusion decisions, a large majority of sample schools said they relied on an ELL-responsive (e.g., *language-related*, *academic-related*, *time-related*, and *opinion-related*) approach to decision-making. Nevertheless, within this approach, decision makers applied a wide range of specific criteria. Thus, while we did find an overall dominant pattern of decision-making for inclusion, the selection of which criteria to apply seemed to result from multiple interpretations of what is meant by "ELL-responsive" criteria. This variability may have been a manifestation of decision makers' attempts to account for the widely varying needs of ELLs. English language learners, by nature, enter school with varying linguistic, academic and socio-cultural backgrounds and may have differing responses to their instructional experiences. This results in differential rates of development of both surface and academic English proficiency (Cummins, 1984, 2000), which in turn influence individual students' readiness to take a high stakes test in English.

For accommodations decisions, we found that decision makers from nearly half of the sample schools interpreted state policies as requiring a disabilities framework, consistent with Rivera et al.'s (2006) review of state policies. Data from the current study suggest two possible misperceptions: either that ELLs should be treated the same as students with disabilities (consistent with findings from Rivera et al.), or that accommodations are intended only for students with disabilities. Further, the NAEP guideline that ELLs receive similar accommodations to those routinely applied in the classroom may have inadvertently reinforced the disabilities framework, since schools' primary recording mechanism for providing accommodations is the Individual Education Plan (IEP). The IEP, however, is designed for students with disabilities, in accordance with the requirements of federal IDEA legislation (U.S. Department of Education, 2004), and is not intended to meet the needs of English language learners. Decision makers relying on the IEP for accommodations decisions would tend to omit ELLs who do not have disabilities from the list of students who may receive accommodations, even if accommodations could have facilitated the meaningful participation of these students.

Another factor that appeared to influence the disabilities framework for accommodations decisions arose from the differing state accommodations guidelines. Specifically, Questions 12 and 13 of the LEP Questionnaire state, "Each eligible ELL is then assigned accommodations that (a) the student has used on state assessment and (b) is approved for use on NAEP."<sup>10</sup> Assessment policies for ELLs vary widely across states. Some state inclusion and accommodation policies fail to distinguish between criteria and accommodations that are appropriate for ELLs and those

<sup>10</sup> Since 2005, NAEP officials have twice revised the ELL inclusion and accommodation decision-making procedures used by school-based decision makers.

criteria and accommodations that are appropriate for students with disabilities. Policies in these states tend to cluster guidelines for these two groups under the single heading of “special needs students,” “at risk students,” or “special populations” (Rivera et al., 2006, p. 45). This guideline may also tend to inadvertently reinforce the practice of treating ELLs similarly to students with disabilities. Findings from this study thus suggest the pervasive reliance on a disabilities framework for ELLs, a framework that tends to be problematic for purposes of NAEP because it can result in inappropriate decisions for including and accommodating ELLs in the assessment.

In addition to differing state policies, differing interpretations of the same policy for accommodations decisions were found. In at least one district (District 2), decision makers who said they had relied on state accommodations guidelines (in accordance with the NAEP decision tree) were equally likely to rely on a disabilities framework as an ELL-responsive framework. This finding suggests that decision makers within a single school district interpreted both NAEP guidelines and state assessment policy in differing ways.

Overall patterns across inclusion and accommodations decisions highlight the wide variability of decision-making processes. As mentioned above, decision makers in District 2 were evenly split between two patterns of decision-making. Interpretation of these two sets of guidelines was even more variable in Districts 1 and 4, in which three patterns of accommodations decision-making were evident. Only one of the four sample districts (District 3) demonstrated a consistent application of ELL-responsive criteria across both inclusion and accommodations decisions. This district had recently aligned its instructional programs for ELLs with both state policy and with a well-developed district-wide program design, factors that may have helped influence decision makers’ understanding of ELL assessment issues as well as the consistency of their decisions regarding ELL participation in NAEP.

In addition to findings regarding patterns of inclusion and accommodations decision-making, data indicated a number of procedural factors that influenced *both* the inclusion and accommodation decisions for ELLs. These factors reveal the complexity of factors that affected decision-making procedures within differing school ecologies, even for decision makers who followed ELL-responsive practices. The following three procedural factors framed *both* the inclusion and accommodation decisions for ELLs: (1) the involvement of the school administrator in making inclusion and accommodation decisions, (2) the role of NAEP field staff, and (3) the use of NAEP tools.

In sum, findings from this study reflect the complex interaction of decision-making criteria and processes with the school culture, district and state policies, and the pressures of daily existence in schools. These findings support the conclusion that (1) reliance on state assessment policy or classroom practice for NAEP inclusion and accommodations criteria may result in inconsistent and inappropriate decisions; (2) that school decision makers need a greater understanding of the needs of ELLs in relation to high stakes testing, in particular regarding the distinction between the needs of ELLs and those of students with disabilities; and (3) that stronger guidance may be needed from NAEP in order to assure the appropriate application of decision-making criteria.

### **Recommendations**

Findings from this study suggest a number of ways for NAEP to support improved decision-making for including and accommodating ELLs. This support can be provided at three moments: prior to the pre-assessment visit, during the pre-assessment visit, and during test administration.

1. Train NAEP staff to recognize common areas of misinterpretation of NAEP guidelines.
2. Provide explicit guidance to school decision makers regarding principles of fair and appropriate testing for ELLs. Make explicit the distinction between the needs of ELLs and those of students with disabilities for participation in high stakes testing.
3. Include a summary of important principles and common areas of misinterpretation in the letter sent to schools along with the new NAEP guidelines prior to the assessment visit.
4. Collaborate with school decision makers during the pre-assessment visit to review their decision-making process. Explicitly guide decision-makers to align their decisions with principles of fair and appropriate assessment of ELLs.
5. Prior to the day of NAEP administration, ensure that there are enough test administration materials for students who will receive accommodations and enough personnel to administer the test.

### *Future Research*

This study has also uncovered areas that would benefit from further research. Future research on this topic might address the following questions:

1. In what ways did concepts or language expressed in school districts' 2005 state assessment policies influence the on-the-ground practices described in the current study?
2. To what extent do the patterns of inclusion and accommodation decision-making identified in this study occur within a statistically generalizeable sample of TUDA districts and schools within those districts? (A large-scale survey of accommodation practices might use questions around the five dimensions and operationalized factors found in Appendix C.)
3. Which and how many accommodations were offered to ELLs other than bilingual booklets? Are bilingual booklets the main accommodation offered to ELLs in NAEP? Did the type of language support program make a difference in the provision of accommodations to ELLs?
4. In what ways did central district staff understanding and leadership regarding ELL-responsive practices and policies differ and how did these differences affect their decisions?

### **Conclusion**

A decade ago, school personnel reported that they were less likely to include ELLs in NAEP due to concerns that students might not be able to participate meaningfully due to their low levels of English language proficiency and because appropriate accommodations were not available (Olson & Goldstein, 1997). Findings from the current study suggest that decision makers' concerns about the meaningful participation of ELLs have not yet been adequately addressed. Although we found an increase in rates of inclusiveness for ELLs in our sample, the low rate of accommodations in some districts and schools and the selection of inappropriate accommodations in other districts and schools is cause for continuing concern. Withholding appropriate accommodations from students who need them not only reduces ELLs' opportunity to participate meaningfully in the assessment but also introduces new sources of error that jeopardize the validity of conclusions regarding the magnitude of achievement gaps. Until ELLs are provided consistent opportunities to participate meaningfully in large-scale assessments, valid and reliable assessment of ELLs is likely to remain one of unrealized promise.

To be effective, accommodations must address the unique needs of the students for whom they are provided. For ELLs, this means assistance in overcoming the linguistic and socio-cultural barriers that prevent them from demonstrating the academic knowledge and skills being tested. Without adequate accommodations, ELL achievement scores cannot accurately reflect what students know and can do (LaCelle-Peterson & Rivera, 1994). By knowing more about how school personnel make decisions about including and accommodating ELLs, NAEP field staff will be better able to provide the guidance necessary to increase the appropriate participation of ELLs in NAEP.

## References

- Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency to academic proficiency among bilingual students. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 2-19). Clevedon, England: Multilingual Matters.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Tonawanda, NY: Multilingual Matters Ltd.
- Haertel, E. (2003). Including students with disabilities and English language learners in NAEP: Effects of differential inclusion rates on accuracy and interpretation of findings. Washington, DC: National Assessment Governing Board. Retrieved September 17, 2006, from <http://www.nagb.org/pubs/conferences/haertle.doc>.
- LaCelle-Peterson, M. & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Education Review*, 64(1), 55-75.
- Maxwell, J. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.
- Mazzeo, J., Carlson, J., Voelkl, K., & Lutkus, A. (2000). Increasing the participation of special needs students in NAEP: A report on 1996 NAEP research activities. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved October 23, 2006, from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000473>.
- Miles, M. & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook (2nd ed.)*. Thousand Oaks, CA: Sage Publications.
- National Center for Education Statistics. (2005). The Common Core of Data: Select 2003 school enrollment data from Boston Public Schools, Chicago Public Schools, Houston Independent School District, and Los Angeles Unified School District. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved October 8, 2005, from <http://nces.ed.gov/ccd/>.
- National Center for Education Statistics. (2006a). NAEP inclusion policy: Inclusion of special needs students. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved September 13, 2006, from <http://nces.ed.gov/naep3/about/inclusion.asp>.
- National Center for Education Statistics. (2006b). NAEP Trial Urban District Assessment. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved September 15, 2006, from <http://nces.ed.gov/nationsreportcard/about/#TUDA>.
- National Center for Education Statistics. (2006c). Percentages of students in urban districts identified, excluded, and assessed in reading and mathematics. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved September 25, 2006, from [http://nces.ed.gov/nationsreportcard/nrc/tuda\\_reading\\_mathematics\\_2005/t0088.asp](http://nces.ed.gov/nationsreportcard/nrc/tuda_reading_mathematics_2005/t0088.asp).



- National Center for Education Statistics. (2006d). Trial Urban District Assessment: Reading 2005. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved October 19, 2006, from <http://nces.ed.gov/nationsreportcard/pdf/dst2005/2006455r.pdf>.
- National Center for Education Statistics. (2006e). Trial Urban District Assessment: Mathematics 2005. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved October 19, 2006, from <http://nces.ed.gov/nationsreportcard/pdf/dst2005/2006457r.pdf>.
- National Center for Education Statistics. (2006f). The NAEP glossary of terms. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved October 21, 2006, from <http://nces.ed.gov/nationsreportcard/glossary.asp>.
- National Center for Education Statistics. (2006g). Reading students assessed by primary type of accommodation, Grades 4 and 8: 1998-2005. Washington, DC: U.S. Department of Education Institute of Education Sciences. Retrieved November 8, 2006, from <http://nces.ed.gov/nationsreportcard/reading/2005readingacctype.asp>.
- Olson, J. & Goldstein, A. (1997). The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress. Washington, DC: U.S. Department of Education National Center for Education Statistics. Retrieved June 4, 2005, from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=97482>.
- Rivera, C., Collum, E., Shafer Willner, L., & Sia Jr., J. K. (2006). An analysis of state assessment policies addressing the accommodation of English language learners. In C. Rivera & E. Collum (Eds.), *A national review of state assessment policy and practice for English language learners*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- U.S. Department of Education. (2004). Individuals with disabilities education improvement act of 2004. *Federal Register Volume 69(249)*. Washington, DC: Author. Retrieved May 11, 2005, from <http://www.ed.gov/legislation/FedRegister/proprule/2004-4/122904a.html>.
- Westat. (2004). NAEP 2005 Assessment Administrator Manual. Bethesda, MD: Westat.
- Yin, R. (2003). *Case study research: Design and methods (3rd edition)*. Thousand Oaks, CA: Sage.

## **Appendix A: Interview Protocol for NAEP School Decision Makers**

School:

Date:

Time:

District:

### **Part 1 - Introductory Questions**

1. Can you tell me a little bit about your position and responsibilities at [school name]?
  - b. Title/position:
  - c. Years in current position:
  - d. Responsibilities, past and current, especially with regard to LEP students:
2. How did you become responsible for deciding whether limited English proficient students in [school name] were included in NAEP?
3. Have you been involved with NAEP before this year?

If YES ask, What were your previous responsibilities for NAEP?

### **Part 2 - Individual Student Decisions**

#### **INCLUDED**

Let's start with STUDENT NUMBER XX. I don't have the names in front of me, so I want to confirm the birth date, found in column "D." STUDENT NUMBER XX's birth date is MO/YR. Is that correct? This student was identified as an LEP student who was included in the NAEP.

If yes, continue. If no, confirm another piece of information to ensure both the respondent and interviewer are discussing the same student.

1. What is the booklet ID number for the student in question?
2. How would you characterize the student's English proficiency?
- 3a. What is your role in the decision-making process for including LEP students in NAEP?

3b. How did you decide that this student should be identified as a limited English proficient student for purposes of NAEP? Probe Respondents about use of documents such as student records, who they talked to, etc.

4. Also, I see on the administration schedule that this student was included in NAEP with/without an accommodation. Is this correct?

5a. Who made the decision to include this student in the NAEP assessment with/without an accommodation, you or another person?

5b. (if another person) Why was this person chosen to make this decision?

5c. Can you give me the name and contact information of this person?

[Note: Interviewer should thank respondent for time and then contact person named here and conduct the interview with that person.]

6. Can you tell me how you decided to include this student in NAEP with/without an accommodation? Try to walk me through this decision, step-by-step.

Use the following probes if not answered in above question:

7a. Did you consult any documents when making this decision (e.g., document with student's LEP designation, other student records)?

7b. If yes, which documents?

7c. If yes, what information did you use from them?

8a. Did you consult with any other school staff when making this decision?

8b. If yes, who did you consult with?

8c. If yes, what information did they provide?

Identify LEP QUESTIONNAIRE

9a. Take a look at the LEP questionnaire. How much, if at all, did you use the LEP questionnaire when making this decision?

If answered a little, some, or a lot:

9b. Can you tell me how you used the LEP questionnaire?

10a. Who completed the LEP questionnaire for this student?

10b. If another person, please specify name and contact information:

10c. On what basis was this person/were you assigned responsibility for completing the questionnaire for this student?

11. Let's take a look at Question 14 in the LEP questionnaire. Can you tell me how this person/you answered Question 14 for this student? Probe what information did the respondent use to answer question 14? Why they choose their answer instead of other response options?

Identify the DECISION TREE on page 3 of the NAEP Inclusion booklet.

12a. How much, if at all, did you also use the decision tree?

If answered a little, some, or a lot:

12b. How did you use the decision tree?

Identify the NAEP Criteria for Inclusion on page 2 of the NAEP Inclusion booklet.

13a. How much, if at all, did you use the document listing the NAEP criteria when making this decision?

If answered a little, some, or a lot:

13b. Can you tell me how you used the document listing the NAEP criteria?

14. Did you use any other sources to make decisions that I have not mentioned?

[List sources]

15a. Let's return to the question about the kind of accommodation used by the student. You decided that the student should or should not be provided with an accommodation.

15b. Was the student allowed to use accommodations other than the one listed on the student administration schedule? If so, can you tell me what these were?

16. How did you decide that this student should be allowed to use this/these accommodation(s)?

17a. How hard was it to decide whether and how to include this student in NAEP?

17b. What about the process was hard/easy

If hard or very hard, ask:

17c. What could make the process of deciding whether and how to include this student easier?

## NOTES

### EXCLUDED

Let's start with STUDENT NUMBER XX. I don't have the names in front of me, so I want to confirm the birth date, found in column "D." STUDENT NUMBER XX's birth date is MO/YR. Is that correct? This student was identified as an LEP student who was excluded from the NAEP.

If yes, continue. If no, confirm another piece of information to ensure both the respondent and interviewer are discussing the same student.

1. What is the booklet ID number for the student in question?

2. How would you characterize the student's English proficiency?

3a. What is your role in the decision-making process for excluding LEP students?

3b. How did you decide that this student should be identified as a limited English proficient student for purposes of NAEP? Probe R about use of documents such as student records, who they talked to, etc.

**\*\*NO QUESTION 4**

5a. Who made the decision to exclude this student in the NAEP assessment, you or another person?

5b. (If another person) Why was this person chosen to make this decision?

5c. Can you give me the name and contact information of this person?

[Note: Interviewer should thank respondent for time and then contact person named here and conduct the interview with that person.]

6. Can you tell me how you decided to exclude this student from NAEP? Try to walk me through this decision, step-by-step.

Use the following probes if not answered in above question:

7a. Did you consult any documents when making this decision (e.g., document with student's LEP designation, other student records)?

7b. If yes, which documents?

7c. If yes, what information did you use from these documents?

8a. Did you consult with any other school staff when making this decision?

8b. If yes, who did you consult with?

## LEP QUESTIONNAIRE

9a. Take a look at the LEP questionnaire. How much, if at all, did you use the LEP questionnaire when making this decision?

If answered a little, some, or a lot:

9b. Can you tell me how you used the LEP questionnaire?

10a. Who completed the LEP questionnaire for this student?

10b. If another person, please specify name and contact information:

10c. On what basis was this person/were you assigned responsibility for completing the questionnaire for this student?

11. Let's take a look at Question 14 in the LEP questionnaire. Can you tell me how this person/you answered Question 14 for this student? Probe what information did the respondent use to answer question 14? Why they choose their answer instead of other response options?

Review of the DECISION TREE on page 3 of the NAEP Inclusion booklet.

12a. How much, if at all, did you also use the decision tree?

If answered a little, some, or a lot:

12b. How did you use the decision tree?

Identify the NAEP Criteria for Inclusion on page 2 of the NAEP Inclusion booklet.

13a. How much, if at all, did you use the document listing the NAEP criteria when making this decision?

If answered a little, some, or a lot:

13b. Can you tell me how you used the document listing the NAEP criteria?

14. Did you use any other sources to make decisions that I have not mentioned?

[List sources]

*N/A [13(14I)]*

NO 15a, 15b, 16

17a. How hard was it to decide whether and how to exclude this student in NAEP?

17b. What about the process was hard/easy

If hard or very hard, ask:

17c. What could make the process of deciding whether and how to exclude this student easier?

## NOTES

**Part 3 - Concluding Questions**

1a. Did you make decisions about whether or not to include LEP students in NAEP prior to or during Westat's visit to your school?

1b. What guidance, if any, did Westat provide in making this decision?

If not answered in Part 1, ask:

2a. Have you been involved with other assessments at your school, for example the state assessments?

2b. If YES, ask Which one(s)?

If respondent has responsibility for state or district assessments, ask:

2c. What is your role in the decision-making process for this/these other assessment(s)?

3a. Did the EXCLUDED student we discussed earlier participate in state assessment? If so, how?

3b. How did the INCLUDED student we discussed earlier participate in the regular state academic assessment in [reading/math]?

3c. If other or accommodations, ask respondent to specify:

4. How does the process of deciding whether or not to include students in NAEP differ from deciding whether or not to include students in your state or district assessments? How is it the same?

5. Is there anything else you would like to add that we haven't talked about today?



## Appendix B: Analysis Codes

### Respondent

- Building administrator
- Program Coordinator/Specialist
- Teacher
- Language acquisition expertise
- No language acquisition expertise
- Respondent not decision maker
- More than one respondent at school
- Respondent job description

### Final Decision Makers

- Teacher
- Administrator
- Testing Coordinator and teachers
- Administrator and teachers
- School-based Committee
- Field Staff Representative only
- Field Staff Representative and teachers
- Field Staff Representative and test coordinators and teachers
- Field Staff Representative and administrators

### Time of Decision

- Prior to Field Staff Representative visit
- During Field Staff Representative visit
- After Field Staff Representative visit

### Extent Field Staff Representative Participated in Decision

- Field Staff Representative decision
- Field Staff Representative guidance
- Field Staff Representative confirmation
- No Field Staff Representative input

### Criteria for ID of Student as LEP

- Time in country
- Time in school program
- English language proficiency
- Academic test score
- Home language survey
- Parent waiver
- Other

### Inclusion/Exclusion Decision Process Used

- Description of inclusion/exclusion process

- Criteria
  - language-related: English language proficiency
  - academic-related: Student grades, classroom performance
  - time-related: Time in US
  - Time-related: Time in LEP program
  - opinion-related: I asked so and so and she said....
  - I don't know

#### Number of Inclusion/Exclusion Decision Actions Used Above

- None
- One
- Two
- Three
- Four
- Five or more

#### Inclusion/Exclusion of Students

- All LEPs included in reading
- All LEPs included in mathematics
- All LEPs excluded from reading
- All LEPs excluded from mathematics
- Beginners excluded from mathematics
- Beginners excluded from reading
- English language takers excluded from test

#### Inclusion/Exclusion Sources of Information

- NAEP document
  - NAEP LEP Questionnaire
  - NAEP Decision tree
  - NAEP Criteria for Inclusion
  - NAEP assessment administrator manual
- Non-NAEP document
  - State policy document
  - District policy document
  - School policy document
- Non-NAEP test results
  - National test results
  - State test results
  - District test results
  - School test results
- Verbal Source
  - Consulting of other school staff
  - Guidance offered by Westat
- Use of other sources

#### How NAEP Documents Used

- Guided included/excluded decision
- Confirmed inclusion/exclusion decision
- Guided the accommodations decision
- Confirmed the accommodations decision
- Not used with decision, but teachers completed
- Not used at all

#### How NAEP and State Guidelines Compare

- Same
- State all students included; NAEP excluded some\
- State all students included
- State guidelines don't allow accommodations
- State allow accommodations
- State guidelines do not use teacher questionnaire
- No say in who tested
- Not answered by respondent

#### How NAEP and School-Based Criteria Compare

- Same
- School-based required all included; NAEP excludes some
- School-based no accommodations
- School-based no teacher questionnaire
- Not answered by respondent

#### Inclusion/Exclusion Primary Guidance From (choose one)

- NAEP guidelines directions
- State policy directions
- District guidelines directions
- School criteria directions
- Other source of information
- I don't know

#### Accommodations Decision Process Used

- Description of accommodation decision-making process
- Criteria
  - language-related: English language proficiency
  - academic-related: Student grades, classroom performance
  - time-related: Time in US
  - time-related: Time speaking English
  - opinion-related: I asked so and so and she said....
  - IEP
  - I don't know
- Other Use of Accommodations
  - Accommodations Criteria Used on other tests

- Accommodations not used on other tests
- Accommodations regularly used in classroom

#### Number of Accommodations Decision Actions Used

- None
- One
- Two
- Three
- Four
- Five or more

#### Priority Given to (choose one)

- NAEP policy directions
- State policy directions
- District policy directions
- School policy directions
- Verbal source of information
- Other source of information
- I don't know

#### Accommodations Sources of Information

- NAEP LEP Questionnaire
- NAEP Decision tree
- NAEP Criteria for Inclusion
- State guidelines document
- State English language proficiency
- District guidelines document
- School-based criteria
- Consulting of other school staff
- Guidance offered by Westat
- Use of other sources
- I don't know

#### Primary Guidance From

- NAEP guidelines directions
- State guidelines directions
- District guidelines direction
- School criteria directions
- Verbal source of information
- Other source of information
- I don't know

#### Why Accommodation Chosen

- Allowed by NAEP

- Allowed on national test
- Allowed on state test
- Allowed by district
- Used in classroom
- Did not offer what we wanted
- Found in student's IEP
- Other

#### Type of Accommodation

- Direct Linguistic Support
  - Test questions read aloud/English audiotape
  - Directions reading aloud/English audiotape
  - Bilingual booklet
  - Bilingual word list glossary
  - Bilingual dictionary no definitions
- Indirect Linguistic Support
  - Small group
  - One-on-one testing
  - Extra time
  - Preferential seating

#### Accommodations Framework Referenced

- Define according to disabilities framework
- Define according to linguistic support

## Appendix C: District Overviews and School Profiles for the Four TUDA Districts

As described in Stage Two of the methods, Appendix C contains district and school arranged according to five dimensions of decision-making practices that emerged from the data: (1) school personnel involved, (2) involvement of NAEP field staff, (3) NAEP tools used, (4) inclusion criteria used, and (5) accommodations criteria used. The codification of operationalized factors within each of the five dimensions allowed the research team to explore relationships between factors within each of the five dimensions to see if any patterns could be found. For example, as part of the exploration of relationships within the data, the following questions were explored for the five dimensions:

- Dimension 1: School personnel involved. Did the decision makers from schools that used certain types of criteria have a particular job role or expertise (e.g., language acquisition specialist, test coordinator, or school administrator) or have a particular number of people in school decision-making unit (e.g., one person, two people, three or more people)?
- Dimension 2: Involvement of NAEP field staff. Did the timing of the decision impact whether NAEP tools were used or the extent to which NAEP field staff were involved in the decision-making process?
- Dimension 3: NAEP tools used. Were NAEP documents or documents from other sources (e.g., state documents) used as the primary guidance for the decision?
- Dimension 4: Inclusion criteria used. Were the NAEP criteria (e.g., amount of instruction in English and academic performance) the only criteria used by decision makers?
- Dimension 5: Accommodation criteria used. Were criteria commonly used with students with disabilities used by decision makers?

The exploration of the relationships among the operationalized factors helped create a more individualized profile of the different facets associated with each districts' decision-making processes.

### District 1

The research team conducted interviews at six schools in District 1: three elementary schools, two middle schools, and one school that housed Grades 4 and 8. A seventh school was targeted to be interviewed. Efforts to include this school or to find a replacement school were unsuccessful. Table E1 provides a list of the schools, grade levels, and student enrollment for each district participating in the study.

Table E1. Enrollment per Grade Level(s) in District 1 Schools Participating in the Study

School	Grade	Student Enrollment for that Grade
Fisk	4	105
Rice	4	36
Tiant	4	118
Yazstremski	4 + 8	72 + 89
Boomer	8	230
Lynn	8	194

Source: National Center for Education Statistics (2005)

Dimension 1: School personnel involved. When comparing the data in Table E2 for “number of decision makers” and “role/expertise to contribute to decision,” we can see that, in

five of six schools in District 1, school administrators were the key decision makers who participated in making the ELL inclusion and accommodation decision. At two schools, the school administrator (e.g., assistant principal) worked with a committee of teachers and specialists; at two other schools, the school administrator worked in tandem with a testing coordinator or language acquisition specialist; and at one school, the school administrator made the decisions alone. At the two schools in which school specialists participated in the decision-making process, the specialists worked in pairs or as part of a committee. These units always made the inclusion and accommodation decision as a team.

As the NAEP guideline suggest, “knowledgeable school staff, such as ESL teachers” be involved in decision-making, school coordinators, the operationalized factor of “Role/Job expertise” was charted. At only two of the six schools held a job that required a background in language acquisition issues

Table E2. Number of Individuals in the Decision-Making Unit in District 1 and Their Role/Expertise

School Name	Grade	Number in Decision-Making Unit			Role/Job Expertise			
		One individual	Two individuals	Committee	Classroom instruction	Language acquisition	Test coordination	School administration
Fisk	4			x	x			x
Rice	4		x			x	x	
Tiant	4	x						x
Yazstremski	4 + 8		x				x	x
Boomer	8		x			x		x
Lynn	8			x	x			x
<b>Summary</b>		<b>1</b>	<b>3</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>5</b>

Dimension 2: Involvement of NAEP field staff. Some decision makers in District 1 reported that they had made the decision to include and accommodate ELLs well before the NAEP pre-assessment visit. As shown in Table E3 when comparing the data for “timing of decision” and “involvement of NAEP field staff,” at two of the six schools (Tiant and Boomer), the decision to include all students had been made by the school staff well before the NAEP pre-assessment visit. However, the day of the test, at one school (Tiant) NAEP field staff failed to arrive with the promised number of personnel needed to monitor students and so according to, the school coordinator, the NAEP field staff representative re-examined the list of students and found five that should have been excluded based on the students’ IEPs.

Table E3. Involvement of NAEP Field Staff-District 1

School name	Grade	Timing of Decision					Involvement of NAEP Field Staff				
		Prior to NAEP field staff visit	During NAEP field staff visit	After NAEP field staff visit	Just before test administered	Unclear/ Do not know	No role	Confirmation of prior decision	Gave strong guidance	Adjusted prior decision already just before test	Unclear/ Do not know
Fisk	4		x					x			
Rice	4			x			x				
Tiant	4	x			x		x			x	
Boomer	4 + 8	x					x				
Yazstremski	8			x			x				
Lynn	8		x					x		x	
<b>Summary</b>		<b>2</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>4</b>	<b>2</b>	<b>0</b>	<b>2</b>	<b>0</b>

At the second school (Boomer) where the decision was made prior to the NAEP field staff visit, the school coordinators told us that the NAEP field staff had played no role in the ELL inclusion and accommodation decisions that had been made.

At the other four schools (Fisk, Yazstremski, Rice, and Lynn) the ELL inclusion and accommodation decision was made during or after the NAEP pre-assessment visit. At two schools (Rice and Yazstremski), where the decision to include and accommodate ELLs was made after the NAEP field visit, decision makers told us that the NAEP field staff played little or no role at all. At two schools (Fisk and Lynn) the decision to include and accommodate ELLs was made during the NAEP field visit and the NAEP field staff simply confirmed the decisions made by the school-based decision makers. At one school (Rice) respondents made it clear they did not need for help making inclusion decisions. At yet another school (Lynn), the school coordinators reported that “They [NAEP staff] did not come with enough people. They wanted everybody in the cafeteria for testing because they didn’t have enough NAEP monitors to administer the test using our schedule.”

Dimension 3: NAEP tools used. As shown in Table E4 when comparing the data for “central document(s) used with decision-making” and “NAEP tool used,” we can see that, of the three tools used by NAEP to inform decision-making, the LEP Questionnaire was the only tool used in all District 1 schools. At each school the LEP Questionnaire was completed by the teacher most familiar with the selected students. However, at three of the six schools (Boomer, Tiant, and Fisk), the LEP Questionnaire was not used in the inclusion decision-making process. At one school (Tiant), the decision to include was based on the state assessment policy that “everyone must be tested.” At another school (Boomer) the decision was based on students’ scores and state and district criteria. At another school (Fisk) decision makers completed the LEP Questionnaire but did not use it to make the inclusion decision.

At the other three schools (Yazstremski, Lynn, and Rice), the LEP Questionnaire was used to shape the final inclusion and accommodation decision. At Yazstremski, the LEP Questionnaire was used “a lot” to inform inclusion and accommodation decisions. At Yazstremski, the LEP Questionnaire was used “a lot” to inform inclusion and accommodation decisions. The Yazstremski respondent found the questionnaire quite useful, as he or she was responsible for reviewing a “vast number of students and could not interview each and every



one!” The Yazstremski respondent explained that she had relied on “teacher opinion exclusively” and had used the questionnaire to ascertain teachers’ judgments regarding particular students. At Lynn, the LEP Questionnaire was used “a lot” to inform inclusion and accommodation decisions. The Lynn respondent indicated that he or she relied on the judgment of the particular student’s classroom teacher. Respondents at Lynn used the decision tree some of the time and felt that the decision tree provided an “easy graphic” and a “clear graphic reference” for decision-making. The respondent also used the NAEP inclusion criteria document “a lot.” The Rice decision makers used NAEP guidelines “to confirm” the decision that had already been made based on decision makers’ knowledge of students. Fisk school coordinators gave more weight to “materials in the school and teacher knowledge” than to NAEP guidelines. Furthermore, the respondent at Fisk indicated that the LEP Questionnaire was used “a little,” but that decisions were based on “who the kid was” and on “materials in school and teacher knowledge.”

Table E4. NAEP Tools Used-District 1

School Name	Grade	Central Document(s) Used With Decision-Making				NAEP Tool Used		
		NAEP tools	State tools	School or district test scores	Staff opinion	NAEP LEP Questionnaire	NAEP Decision tree	NAEP criteria document
Fisk	4		x		x	x		
Rice	4				x	x	x	x
Tiant	4		x					
Yazstremski	4 + 8	x			x	x		
Boomer	8			x				x
Lynn	8	x			x	x	x	x
<b>Summary</b>		<b>2</b>	<b>2</b>	<b>1</b>	<b>4</b>	<b>4</b>	<b>2</b>	<b>3</b>

Dimension 4: Inclusion criteria used. As shown in Table E5, at four of the six schools students’ English language proficiency and academic level played a role in some inclusion decisions – especially those involving lower level ELLs. Four of the six schools referred to English language proficiency (Fisk, Rice, Yazstremski, and Lynn). The Boomer respondent explained, “We looked at the NAEP criteria and they are very broad. Anyone who has been here for less than 3 years should be excluded. Well, we know our kids better and some who have been here a shorter time should take the test, not be excluded.” The Boomer respondent explained, “We used our own criteria.” Nonetheless, the primary criteria used at Boomer were state and district mechanisms used for classifying ELLs: test results—MELA-O (given in fall and spring), MEPA (brand new language proficiency assessment), and descriptors provided by the Office of Language Services. Academic performance was mentioned by school coordinators at two schools (Rice and Boomer). Among the District 1 decision makers surveyed, only two of the six mentioned using Time in English language instruction as a criteria (Fisk and Rice). Four of the six schools referred to the criterion of “time in the U.S.” (Rice, Yazstremski, Boomer, and Lynn). In at least one instance at Fisk, it was judged that taking the reading test would be “far too frustrating” for newly arrived ELLs. The respondent indicated that in applying this general rule,

the decision makers based their decisions on “who the kid was.” (It should be noted that Fisk decision makers told the research team that they included all ELLs except for those “with no English language ability whatsoever.”)

However, there was one commonality among three of the six schools. Decision makers at Fisk, Tiant, and Boomer said they relied on state inclusion criteria for ELLs when making the ELL inclusion decision for NAEP. The Fisk respondent explained, “NAEP criteria are the same as those used for MCAS testing anyway, so of course we did [follow NAEP guidelines].” At Tiant, the decision maker also on state assessment policy, but had a differing interpretation that “everyone must be tested.”

Table E5. Inclusion Criteria Used-District 1

School Name	Grade	English Language Proficiency	Academic Performance	Time in US	Amount of Instruction in English	Teacher Opinion	Unclear/ Do not know
Fisk	4	x			x	x	
Rice	4	x	x	x	x		
Tiant	4						
Yazstremski	4 + 8	x		x			
Boomer	8		x	x			
Lynn	8	x		x			
<b>Summary</b>		<b>4</b>	<b>2</b>	<b>4</b>	<b>2</b>	<b>1</b>	<b>0</b>

Dimension 5: Accommodations criteria used. Within District 1, there were two clear distinctions among the way decision makers assigned accommodations to ELLs. As shown in Table E6, one group of decision makers (two out of the six - Yazstremski and Lynn) used the student’s English language proficiency as the primary criterion for providing accommodations. In District 1, ELP was used for the assignment of accommodations in two of three six schools, but of whom had 8<sup>th</sup> graders. Not surprisingly these same two of six District 1 schools (Yazstremski and Lynn) tied the inclusion decision to the provision of accommodations. The respondent at Yazstremski indicated that the decision to let a student take the NAEP with the read aloud accommodation was based on teacher recommendation and the MELA-O listening score. The Yazstremski respondent also observed that the LEP Questionnaire was useful in this decision because “it was clear, so it was easy to define accommodations. Yazstremski respondents indicated that usually [in state testing] a students’ IEP was an important reference point for deciding which students were included with or without accommodations. A common observation was that only students with IEPs were excluded from testing.

Also shown in Table E6, another group of decision makers (four out of the six –Fisk, Tiant, Rice, and Yazstremski) explained that accommodations were only offered to students with IEPs. Only students with IEPs were excluded from testing. The Rice respondent did not give any accommodations to ELLs because she was following the NAEP requirements listed in the LEP Questionnaire (which asked decision makers to consult state accommodations requirements). She explained that, according to MCAS requirements, ELLs do not “get any Spanish help” (i.e., native language accommodations are not available for students taking MCAS). Similarly, at Boomer, the respondent observed that availability of accommodations varied among state and district assessments and that NAEP provides broader accommodations. The state says only IEP;

NAEP was not as narrow in who might receive an accommodation.

Table E6. Accommodations Criteria Used-District 1

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Offered no accommodations	IEP	Unclear/ Do not know
Fisk	4						x	
Rice	4						x	
Tiant	4						x	
Yazstremski	4 + 8	x						
Boomer	8					x	x	
Lynn	8	x						
<b>Summary</b>		<b>2</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>3</b>	<b>0</b>

### *Fisk*

Dimension 1: School personnel involved. The respondent, an assistant principal, explained that he or she “sat around a table” with classroom teachers, resource teachers, and the NAEP field staff representative to make a “group decision” regarding individual students. Although the principal does not seem to have been involved in discussions about individual students, she set the agenda by indicating that all ELLs were to be considered eligible to participate in NAEP unless the student’s IEP indicated otherwise or the student had “no English speaking or reading ability whatsoever.”

Dimension 2: Involvement of NAEP field staff. The respondent indicated that the decision to use accommodations was based on a group decision in which NAEP field staff took part.

Dimension 3: NAEP tools used. Fisk decision makers gave more weight to “materials in the school and teacher knowledge” than to NAEP guidelines. The respondent at Fisk indicated that the LEP Questionnaire was used “a little,” but that decisions were based on “who the kid was” and on “materials in school and teacher knowledge.” In to at least one instance, the respondent indicated that the student’s level of English language proficiency, which was listed in the questionnaire, was used to make inclusion decisions. So it seems likely that the questionnaire served as a vehicle to convey the teacher’s assessment of the student’s English language proficiency.

Dimension 4: Inclusion criteria used. Decision makers were influenced by state guidelines; however, exceptions were made based on individual student considerations. The Fisk principal set the primary criterion for inclusion: All ELLs were to be considered eligible to participate in the NAEP unless the student’s IEP indicated otherwise or the student had “no English speaking or reading ability whatsoever.” According to the Fisk respondent, “NAEP criteria are the same as those used for MCAS testing anyway, so of course we did [follow NAEP guidelines].”

Yet the respondent’s answers suggest that Fisk decision makers actually gave more weight to “materials in school and teacher knowledge” than to NAEP guidelines and often based decisions on “who the kid was.” Indeed, it appears that NAEP documents were not used in

decision-making. In at least one instance, it was judged that taking the reading test would be “far too frustrating” for newly arrived ELLs. The respondent indicated that in applying this general rule, the decision makers based their decisions on “who the kid was.” For example, a student excluded from the reading NAEP was described as having “no proficiency.” He was excluded because he had been at the school only since December 2003, “was not literate in either language,” and therefore would have found the reading test “far too frustrating.”

Dimension 5: Accommodations criteria used. The Fisk respondent reported that only students who had an IEP (and were classified as special education) were allowed to have accommodations. The respondent provided an example of a special education student who had received the accommodation of taking the Math NAEP in a small group because this accommodation was explicitly required by the student’s IEP.

### *Rice*

Dimension 1: School personnel involved. The respondents, a literacy testing coordinator and a sheltered English instruction teacher, described the decision-making process as “cut and dried.” The sheltered English instruction teacher, who had worked with all the ELLs sampled by NAEP, and the literacy testing coordinator worked closely together. Both respondents worked with teachers to gather information and make decisions.

Dimension 2: Involvement of NAEP field staff. Although the NAEP field staff representative was available to provide guidance, the respondents did not feel that they needed any help to make inclusion decisions. The decision was made after the NAEP field staff pre-assessment visit.

Dimension 3: NAEP tools used. The Rice decision makers used NAEP guidelines “to confirm” the decision that had already been made based on decision makers’ knowledge of students. Hence, at Rice, the LEP Questionnaire, decision tree, and criteria were used “some.” The respondent (a sheltered English instruction teacher) “filled out the questionnaire because that was what I was supposed to do, but it was really a formality because I know my students well at this time of year.” Respondents at Rice felt that decision tree provided an “easy graphic” and a “clear graphic reference” for decision-making, but only referred to it some of the time.

Dimension 4: Inclusion criteria used. The actual decision to include or exclude LEP students was based on English language proficiency, time in the US, and academic level. The Rice respondent had worked with all the students listed on the NAEP administration schedule and so had a basic understanding of their readiness to participate in NAEP. Students’ English proficiency was gauged by test performance on large-scale tests (MELA-O, LAS, MCAS, and/or SAT9) and formative assessments. As it turned out, most of the ELLs selected by NAEP field staff had been reading below grade level and so were not included.

The sheltered English instruction teacher stressed that the reason most of their students were excluded was because Rice is only in its second year of sheltered English instruction after having been a bilingual school. Because of this, the students hadn’t been in sheltered English instruction long enough to have the English skills needed to take the test. The teacher noted that almost all his or her students were reading at least two levels below grade level in English. Only three or four of his or her students were reading at the appropriate grade level and “were probably good enough to be included.”

Dimension 5: Accommodations criteria used. Using the NAEP requirements listed in the LEP Questionnaire (to consult state accommodations requirements), the Rice respondent also

indicated that according to MCAS requirements, ELLs do not “get any Spanish help” (i.e., native language accommodations are not available for students taking MCAS).

### *Tiant*

Dimension 1: School personnel involved. The respondent, a Director of Instruction, also worked with teachers to make inclusion/accommodation decisions, which was presented to NAEP field staff on the day of the test.

Dimension 2: Involvement of NAEP field staff. The NAEP field staff representative seems not to have played a role in the initial decision. However, because the NAEP field staff did not arrive with the promised number of personnel on the day of the test, there were not enough testing personnel to monitor students during the test, the NAEP field staff representative decided to exclude five students on the day of the test based on the students’ IEPs.

Dimension 3: NAEP tools used. At Tiant, the respondent indicated that the LEP Questionnaire “went straight to the teacher,” who had the best knowledge of the student’s ability. However, decisions were based, not on this tool, but on the state assessment policy that “everyone must be tested.” Neither the NAEP decision tree nor the NAEP inclusion criteria document were used at all in making decisions to include and accommodate ELLs, but the respondent did not elaborate on why these tools were not used.

Dimension 4: Inclusion criteria used. An initial decision was made to include all students in the NAEP (following the policy for state assessment that “everyone must be tested”). The respondent noted, “We think everybody should be tested. We did this a couple of years ago and we said that we wanted to do it with all the 4<sup>th</sup> grade students.” “Anything that came in we gave to the teacher because she has the most accurate knowledge of how the student would manage a test such as this. The teacher looked over the NAEP criteria, weighed them against other information [MELA-O, MEPA, MCAS, citywide testing “We do testing, testing, testing, testing.”] and decided that a student would take the test. We included ALL the students. Every 4<sup>th</sup> grader was supposed to be included. This is a state mandate.” The only criteria explicitly identified by the respondent as being used in making inclusion and accommodation decisions for ELLs were scores from MELA-O and MEPA, which are used to measure, respectively, students’ oral and written proficiency in English.

However, because the NAEP field staff did not arrive with the promised number of personnel, there were not enough testing personnel to monitor students during the test, the NAEP field staff representative decided to exclude five students on the day of the test based on the students’ IEPs. The respondent explained that “this school is used to testing everyone” and was disappointed not to be able to test all students simply because there were not enough people to monitor the students during the test.

Dimension 5: Accommodations criteria used. Tiant staff had planned to test every student listed on the NAEP administration schedules. The respondent indicated that “accommodations [for state test] come from special education. The teacher knows all accommodations, and I go over these with the teachers.” (A read-aloud accommodation for the Math NAEP was listed in the NAEP administration schedule but not mentioned during the interview.)

However, the Tiant respondent reported that the NAEP field staff representatives were unprepared to accommodate LEP students. “They [NAEP field staff] were supposed to come with 10 or more testers and they didn’t. We were in a predicament. I was expecting the people here, at least a good 10 of them –we had students with accommodation, all the students, and we

only had 5 NAEP testers. He [NAEP field staff representative] said ‘I can’t accommodate them. I’ll eliminate them.’ The hard thing is that we as a school decided to test everyone rather than a random number. With eliminations, we had to find space for students who would not be tested rather than testing in their rooms. The people who are coming to proctor the test should have a general idea of who needs what when. They were not prepared when they came. It was a problem.

### *Yazstremski*

Dimension 1: School personnel involved. The respondent, a student support coordinator and counselor, indicated that he or she “worked with teachers and the LEP Questionnaire.” The respondent was new at the school so he also consulted with a more experienced staff member (an “academy director”—the school is divided into three academies). In this case, the teacher made the decision and the respondent did the necessary paperwork.

Dimension 2: Involvement of NAEP field staff. Although the NAEP field staff representative visited the school to discuss the inclusion and accommodation process, the respondent did not base decisions on the input of the NAEP field staff representative.

Dimension 3: NAEP tools used. At Yazstremski, the LEP Questionnaire was used “a lot” to inform inclusion and accommodation decisions. The Yazstremski respondent found the questionnaire quite useful, as he or she was responsible for reviewing a “vast number of students and could not interview each and every one!” This respondent relied on “teacher opinion exclusively” and used the questionnaire to ascertain teachers’ judgments regarding particular students. Neither the NAEP decision tree nor the NAEP inclusion criteria documents were used in making decisions to include and accommodate ELLs.

Dimension 4: Inclusion criteria used. Teacher opinion was conveyed primarily through the LEP Questionnaire, which teachers completed for each student selected to participate in NAEP. Criteria for decisions at the school were based on the length of time the student had been in the U.S. and English language proficiency level (based at least in part on MELA-O scores). The respondent also considered the student’s “sheltered English instruction level” in making inclusion decisions. For instance, two students who were excluded from the NAEP were discussed. One was a beginning-level English language proficiency student and the other was described as having no proficiency in English; both students had only recently arrived in the U.S.

Dimension 5: Accommodations criteria used. A read aloud accommodation was offered to students at Yazstremski. The respondent at Yazstremski indicated that the decision to let a student take the NAEP with the read aloud accommodation was based on teacher recommendation and the MELA-O listening score. The respondent also observed that the LEP Questionnaire was useful in this decision because “it was clear, so it was easy to define accommodations.

Yazstremski respondents indicated that usually [in state testing] a students’ IEP was an important reference point for deciding which students were included with or without accommodations. A common observation was that only students with IEPs were excluded from testing.

### *Boomer*

Dimension 1: School personnel involved. The respondents, a director of instruction and the language assessment team (LAT) coordinator, worked closely together in the decision-

making process. Students' classroom teachers completed LEP Questionnaires, but do not seem to have been actively involved in decision-making.

Dimension 2: Involvement of NAEP field staff. Decisions were made prior to the visit from the NAEP field staff representative, who seems to have played a minimal role in decision-making.

Dimension 3: NAEP tools used. The Boomer respondent simply stated that, although teachers completed the LEP Questionnaire for their students, questionnaires were not used in the decision-making process. The NAEP decision tree was not used at all in making decisions to include and accommodate ELLs. Respondents did not elaborate on why the tool was not used. The NAEP inclusion criteria document was used "a little," primarily "to confirm" a decision that had already been made based on decision makers' knowledge of students.

Dimension 4: Inclusion criteria used. The respondent simply stated that, although teachers completed LEP Questionnaire for their students, the questionnaire was not used in the decision-making process. "When the list came from NAEP, the teachers were given the questionnaire and filled it out. However, we didn't really use them [questionnaires]. I looked at students' scores and state and district criteria rather than NAEP." They continued, "We looked at the NAEP criteria and they are very broad. Anyone who has been here for less than 3 years should be excluded. Well, we know our kids better and some who have been here a shorter time should take the test, not be excluded." The respondent explained, "we used our own criteria." Nonetheless, the primary criteria used were state and district mechanisms used for classifying ELLs: test results—MELA-O (given in fall and spring), MEPA (brand new language proficiency assessment), and descriptors provided by the Office of Language Services.

In Boomer, all ELLs selected to participate in the NAEP who were judged to be at Intermediate, Early Intermediate, or Transitional levels of English language proficiency participated in the test. Although decisions were based on time in program/U.S. and English language proficiency level, these decisions were also framed by teacher recommendation about the student's readiness to take the test. Students take SRI (Scholastic Reading Inventory, a district test) that measures reading comprehension. They have to meet benchmarks in order to move through levels in SEI and out of that program. There are specific district requirements for each level.

The Boomer respondent indicated that a student was excluded from NAEP because taking the test "would have been completely overwhelming" for the student. "It would be great if all students could take this. However, those with no proficiency should be excluded." The Boomer respondent noted that "this student and 2 others—only 3 were excluded—came to us in the summer of 2004 with no English facility whatsoever at the very beginnings of English learning."

Dimension 5: Accommodations criteria used. Boomer respondents indicated that a students' IEP was an important reference point for deciding which students were included with or without accommodations. Only students with IEPs were excluded from testing. The respondent commented that availability of accommodations varied among state and district assessments and that NAEP provides broader accommodations. The state says only IEP; NAEP was not as narrow in who might receive an accommodation.

*Lynn*

Dimension 1: School personnel involved. The respondent, a director of instruction, indicated that decisions were made during the visit from the NAEP field staff representative.

Dimension 2: Involvement of NAEP field staff. In making decisions, the respondent worked “in concert with” students’ classroom teachers often deferring to teacher judgment because the teachers “know the students so well.” The respondent reported, “We waited for the criteria to make our decisions and the information came prior to testing. It was mailed out beforehand. The woman in charge was “available for questions” and who had “kept in regular touch.”

“I have one complaint. We set up a whole schedule for testing, so that when the team came in, everything would run smoothly. However, they did not come with enough people. They wanted everybody on the cafeteria for testing because they didn’t have enough NAEP monitors to administer the test using our schedule. This was not advantageous for our students and did not give our students the best accommodations. We are wondering why the classroom teachers are not administering this assessment. It would make it much easier. We could keep kids in their homerooms for one thing. Why do they need NAEP administrators in each room? In our school, the way we had the schedule was better. We had 8 group lists and they came with only 5 administrators. When we considered pull-outs, we had 11 groups and there were not enough NAEP people. If NAEP is administered in a school with a large population, there should be availability of staff to test students in more ideal conditions.”

Dimension 3: NAEP tools used. At Lynn, the questionnaire was used “a lot” to inform inclusion and accommodation decisions. The Lynn respondent indicated that he or she relied on the judgment of the particular student’s classroom teacher. This respondent reviewed the questionnaire with the teacher who completed it “so that there would be no ambiguity.” Respondents at Lynn used the decision tree some of the time and felt that the decision tree provided an “easy graphic” and a “clear graphic reference” for decision-making. The respondent also used the NAEP inclusion criteria document “a lot.”

Dimension 4: Inclusion criteria used. Interestingly, the Lynn respondent commented, “We used the questionnaire primarily but teachers are biased because they know the students so well.” Lynn decision makers also based the criteria for decisions on the length of time the student had been in the U.S. and English language proficiency level. Only special education students were excluded from the NAEP

Dimension 5: Accommodations criteria used. The accommodation offered in Math was ideal for including students that needed help that otherwise would have to have been excluded. The respondent notes, “In math, students had access to glossary which was most helpful.” In making decisions, the respondent relied on the judgment of the particular student’s classroom teacher. The respondent indicated that in making a recommendation for inclusion or accommodation, the teacher considered such factors as classroom performance, length of time in country, and mastery of English, which was in part determined on the basis of “scores on language tests.” The respondent went on to explain the use of accommodations in general based on ELLs’ *linguistic* needs: “To receive an accommodation for MCAS, a student does not necessarily have to be classified as special education.”



## District 2

The research team conducted interviews at seven schools in District 2: three schools for Grade 4, three schools housing Grade 8, and one school enrolling students in Grades 4 and 8. Table E7 lists the seven schools that participated in the GW-CEEE study and school enrollment by grade level.

Table E7. Enrollment for Grade Level(s) in District 2 Schools Participating in Study

School	Grade	Student Enrollment for that Grade
Sandburg	4	174
Middleton	8	110
Pronto	4	203
Raleigh	4 + 8	120 + 136
Bacon	8	402
Drake	8	85
Day	4	55

Source: National Center for Education Statistics (2005)

Dimension 1: School personnel involved. Within District 2, test coordinators were most likely to be involved in the ELL inclusion and accommodation decision. As can be seen by comparing the data for the “number of decision makers” and the “role/expertise to contribute to decision” in Table E8, test coordinators served on decision-making committees, worked in tandem with another decision maker, or made the decision individually. Decision makers with language acquisition background either assisted on committees or worked with another decision maker, but never made the decision individually.

Table E8. School Personnel Involved-District 2

School Name	Grade	Number of Decision Makers			Role/Expertise to Contribute to Decision			
		One individual	Two	Three or more	Classroom instruction	Specialist in language acquisition	Test coordination	School administration
Sandburg	4	x					x	
Day	4		x				x	x
Pronto	4	x					x	
Middleton	8			x	x	x		x
Raleigh	4 + 8			x	x	x	x	x
Bacon	8	x					x	x
Drake	8		x			x	x	
<b>District 2</b>		<b>3</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>6</b>	<b>4</b>

Dimension 2: Involvement of NAEP field staff. As shown in Table E9 when comparing the data for “timing of decision” and “involvement of NAEP field staff,” we can see that in every school in District 2 except Bacon, respondents indicated that decisions regarding which students to include or exclude from the NAEP were made prior to the “visit from NAEP field staff.” NAEP field staff, in general, confirmed the decisions that had already been made when they came to the school for the pre-assessment visit. At Bacon, once the school received the list of students selected for participation in the NAEP, the principal and “NAEP administrator” (presumably, the NAEP field staff representative) made decisions regarding inclusion of students.

Table E9. Dimension 2: Involvement of NAEP Field Staff-District 2

School name	Grade	Timing of decision					Involvement of NAEP field staff				
		Prior to NAEP field staff visit	During NAEP field staff visit	After NAEP field staff visit	Just before test administered	Unclear/ Do not know	No role	Confirmation of prior decision	Gave strong guidance	Adjusted prior decision just before test	Unclear/ Do not know
Sandburg	4	x						x			
Day	4	x					x				
Pronto	4					x					x
Middleton	8	x						x			
Bacon	4 + 8		x						x		
Raleigh	8	x						x			
Drake	8	x						x			
<b>District 2</b>		<b>5</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>4</b>	<b>1</b>	<b>0</b>	<b>1</b>

Dimension 3: NAEP tools used. As shown in Table E10 when comparing the data for “central document(s) used with decision-making” and “NAEP tool used,” we can see that, all three NAEP tools were used at three schools (Sandburg, Raleigh, and Drake) as a central part of their decision-making processes. Two of the three NAEP tools were used at Middleton and Day. It appears that only one school, Pronto, used only one NAEP tools, the LEP Questionnaire. Of the three tools distributed by NAEP field staff to inform decision-making in District 2, the LEP Questionnaire was the only one to have been used in all schools sampled in District 2. At each school the questionnaire was completed by the teachers most familiar with the selected students. At only one school, Bacon it was unclear the extent to which decision makers at Bacon used the completed LEP Questionnaire, Decision Tree, and NAEP Criteria of Inclusion during decision-making.

Table E10. NAEP Tools Used-District 2

School Name	Grade	Central Document(s) Used with Decision-Making				NAEP Tool Used		
		NAEP tools	State tools	School or district test scores	Staff opinion	NAEP LEP Questionnaire	NAEP Decision tree	NAEP criteria document
Sandburg	4	x		x		x	x	x
Day	4	x			x	x		x
Pronto	4	x		x		x		
Middleton	8				x	x		x
Raleigh	4 + 8			x	x	x	x	x
Bacon	8	x			x	unclear	unclear	unclear
Drake	8			x		x	x	x
<b>District 2</b>		<b>4</b>	<b>0</b>	<b>4</b>	<b>4</b>	<b>6</b>	<b>3</b>	<b>5</b>

Dimension 4: Inclusion criteria used. District 2 decision makers did not seem to be influenced by the state inclusion policy (and feel that all ELLs needed to be included). Neither did they report that the list of pre-selected students was the final list of students for the test. The Day respondent summarized, “The NAEP process differs from the state. For the state, student must be in program year 3 or higher to take a [standardized] test in English, but we also give it in Spanish. For NAEP, we were given a choice.

As shown in Table E11, within District 2, a majority of decision makers did use the NAEP criteria of amount of instruction in English and academic performance. However, in addition to the NAEP criteria, decision makers at five of the seven schools sampled also referred to English language proficiency and time in the U.S. Using the criteria of English language proficiency and academic performance, lower level students were excluded from the NAEP at Drake. However, decision makers also considered the impact of the test on the student; hence, a student would be excluded if the decision makers thought that taking the test would be “detrimental to the child.”

Table E11. Inclusion Criteria Used-District 2

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Teacher opinion	Unclear/ Do not know
Sandburg	4	x	x		x		
Day	4	x			x	x	
Pronto	4		x				
Middleton	8	x				x	
Raleigh	4 + 8		x		x	x	
Bacon	8	x					
Drake	8	x	x	x			
<b>District 2 Summary</b>		<b>5</b>	<b>4</b>	<b>1</b>	<b>3</b>	<b>3</b>	<b>0</b>

Dimension 5: Accommodations criteria used. As shown in Table E12, at two of the seven schools sampled in District 2 (Bacon and Day), accommodations decisions were made based on the students' English language proficiency. In three schools (Raleigh, Day, and Bacon) respondents indicated they were able to include student because they were able to provide the students with accommodations. Because of the availability of the bilingual test booklet, no student was excluded at Day from the mathematics NAEP. (Using the LEP Questionnaire, one Day student had been given a bilingual booklet for the mathematics assessment, while other students at Day were excluded from reading test because a dual-language booklet was not available.) However, in at least one instance, the Raleigh respondent felt that a student who was in the U.S. and bilingual program for less than 2 years "could manage the test with the accommodation of extended time." Because NAEP offered this accommodation, the teacher wanted students to attempt the test. The respondent reiterated, however, that, ordinarily, LEP students do not receive any accommodation unless they are also identified as special education.

However, according to decision makers, the state policy on the use of accommodations (only for students with IEPs) exerted an influence on how they accommodated. In discussing decision-making for accommodating students participating in the NAEP, the most common response (at four schools: Raleigh, Middleton, Pronto, and Sandburg) was that, generally, only special education students were allowed to use accommodations on large-scale tests (e.g., per state guidelines). The Raleigh respondent explained that, generally, only special education students (students with an IEP) were given accommodation. However, in at least one instance, the respondent felt that a student who was in the U.S. and bilingual program for less than 2 years "could manage the test with the accommodation of extended time." Because NAEP offered this accommodation, the teacher wanted students to attempt the test. The respondent reiterated, however, that, ordinarily, LEP students do not receive any accommodation unless they are also identified as special education. The decision maker at Drake (who had made her decision prior to the pre-assessment visit from NAEP) commented, "For us, 'accommodations' means Special Ed (or 504). NAEP doesn't differ from the state [criteria] because we use standardized tests and follow the same procedure....Children in the bilingual programs have a tendency to score real well in math. We don't give accommodations on some tests and not on others." Interestingly, the Drake respondent also commented, "No accommodation was offered to students during the NAEP; however if this were a mandatory assessment, I would look at lower-level testing. But in this case, if I thought the student was capable, the student would take the on-level test."

Table E12. Accommodations Criteria Used-District 2

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Offered no accommodations	IEP	Unclear/ Do not know
Sandburg	4						x	
Day	8	x						
Pronto	4						x	
Middleton	4						x	
Raleigh	4 + 8				x		x	
Bacon	8	x						
Drake	8					x		
<b>Summary</b>		<b>2</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>4</b>	<b>0</b>

### *Sandburg*

Dimension 1: School personnel involved. The respondent, a school counselor who is also the testing coordinator, reported receiving “packages” from NAEP field staff. The NAEP field staff provided assistance in “making sure we read guidelines for inclusion and exclusion.”

Dimension 2: Involvement of NAEP field staff. Inclusion and accommodation decisions were made prior to NAEP field staff’s visit to the school.

Dimension 3: NAEP tools used. The respondent at Sandburg used the teacher-completed questionnaire and sought supporting data in making decisions. At Sandburg, the NAEP inclusion criteria document “definitely came into play,” The decision makers at Sandburg used the decision tree “some,” finding it “helpful to have an outline” of the decision-making process.

Dimension 4: Inclusion criteria used. The respondent felt that the guidelines for NAEP were clear, consistent, easy to use, and were identical to those used to made inclusion and exclusion decisions for local assessments. To make decisions, the respondent reviewed categories found on the NAEP list and applied information from files (including ITBS scores, IMAGE level, and years in program) along with teacher recommendation. Overall, the respondent indicated that the two most important criteria for determining whether or not student would participate in NAEP were (1) time in program (if less than 3 years they would not participate) and (2) proficiency (was student proficient enough to participate in standardized testing?). However, decision makers also considered the impact of the test on the student; hence, a student would be excluded if the decision makers thought that taking the test would be “detrimental to the child.”

Dimension 5: Accommodations criteria used. Accommodations could be used if the student has an IEP status.

### *Middleton*

Dimension 1: School personnel involved. The respondent for Middleton was the Assistant Principal who had been at the job 10 years and had served previously as bilingual lead teacher. The respondent conferred with teachers to make inclusion/exclusion decisions for ELLs.

Dimension 2: Involvement of NAEP field staff. Decision-making materials were sent from NAEP field staff, but no one visited school prior to testing. However, the respondent did talk to NAEP field staff representative over the phone. The respondent felt that NAEP field staff representative was very helpful.

Dimension 3: NAEP tools used. The respondents at Middleton reported that decision makers used the LEP Questionnaire “a lot.” The Middleton respondent indicated that although the teachers completed the LEP Questionnaire and he or she completed the school questionnaire, they reviewed the questionnaires together. “Basically, we went over stuff together and decided. We did use the NAEP criteria...Remember, we didn’t choose the students to take the test. However, deciding who to include was not difficult because they [the students] are already categorized. It is not difficult to match up. I understand a national assessment needs to cover all the bases. I did not think this was all that difficult. Actually, the teacher questionnaire was the most difficult part of the NAEP, really.” Decision makers at Middleton did not use the NAEP decision tree at all.

Dimension 4: Inclusion criteria used. The Middleton respondent indicated that although teachers completed the LEP Questionnaire and he or she completed the school questionnaire, they “all worked together with this material.” Decisions at Middleton were based on student scores on English language proficiency tests (which are used to classify students) and teacher observation. The respondent found it relatively easy to make determinations regarding student participation in the NAEP because the students were already categorized; it was just a matter of “matching students’ categories” with teacher input.

Dimension 5: Accommodations criteria used. Accommodations were assigned based on IEP at Middleton, as only special education students were allowed to use accommodations on large-scale tests (e.g., per state/district guidelines).

### *Pronto*

Dimension 1: School personnel involved. A teacher served as respondent to our questions for this study. She explained, “I took over testing after decisions had been made.” The school’s testing coordinator had decided to go back to the classroom before she retired. The testing coordinator was involved in the set-up for NAEP but had left the school by the time the research team began conducting telephone interviews.

Dimension 2: Involvement of NAEP field staff. The respondent did not know the extent to which the NAEP field staff representative had participated in the decision or when the decision had been made in relation to the NAEP field staff representative’s site visit.

Dimension 3: NAEP tools used. At Pronto, the LEP Questionnaire was used to “decide accommodations and to include or exclude after NAEP gave list of students selected.” The decision maker at Pronto did not use the NAEP decision tree or the NAEP inclusion criteria document at all.

Dimension 4: Inclusion criteria used. The ITBS score, which indicates academic level, was the most important criterion used by the Pronto decision maker. However, the LEP Questionnaire completed by students’ teachers was used to help decide which students to include in or exclude from the NAEP and which accommodations should be made available to particular students.

Dimension 5: Accommodations criteria used. If the student was not in special education, he or she participated in the NAEP without accommodation.

### *Day*

Dimension 1: School personnel involved. The respondent, the assistant principal and coordinator for testing, first gathered suggestions from teachers and then made final decision. They [NAEP field staff] chose the administrative codes but we [the two respondents] decided together. The NAEP field staff filled out the questionnaires and provided us with the answers we needed.”

Dimension 2: Involvement of NAEP field staff. Decisions regarding students’ participation in NAEP were made prior to the visit from the NAEP field staff representative. They explained, “We gathered suggestions from teachers, but we made the final decision.”

Dimension 3: NAEP tools used. The Day respondent read the completed LEP questionnaire and used the information about English language proficiency. Decision makers at Day did not use the NAEP decision tree at all. Decision makers at Day used the NAEP inclusion criteria document “some.” At Day, the criteria document was used because it reinforced the usual practice at the school.

Dimension 4: Inclusion criteria used. The Day respondent remarked, “I based the decision on teacher recommendations, program year and what I know about the student. The teachers were assigned the responsibility for completing the [LEP] questionnaire. They used observation, assessment from the classroom and if students were beginning ESL [English language proficiency level]. Whatever the strongest language was, that was what we used.” Because of the availability of the bilingual test booklet, no student was excluded from the Math NAEP.

Students were excluded from the Reading NAEP because a dual-language booklet was not available. The respondent explained, “Our school is different because we are dual language. Most schools don’t give LOGRAMOS after the second year. We teach first and second grade completely in Spanish. This is when students learn reading strategies - so for our kids, taking a test like NAEP in 4<sup>th</sup> grade doesn’t work because their skills are not in English. They should be tested in Spanish. For our students, a Spanish NAEP would be better because it would more accurately test their reading skills. We excluded them because the test was not in Spanish. If the tests were in Spanish, it would have made it much easier to decide.”

The Day respondent summarized, “The NAEP process differs from the state. For the state, student must be in program year 3 or higher to take a [standardized] test in English, but we also give it in Spanish. For NAEP, we were given a choice.

Dimension 5: Accommodations criteria used. Because of the availability of the bilingual test booklet, no student was excluded from the Mathematics NAEP. (Using the LEP Questionnaire, one student had been given a bilingual booklet for the mathematics assessment.) Students were excluded from reading test because a dual-language booklet was not available.

### *Raleigh*

Dimension 1: School personnel involved. As NAEP coordinator, the respondent (a school counselor) worked with principal (a former bilingual coordinator) and teacher recommendation to include and accommodate students selected for participation in the NAEP. The respondent explained that the decision makers arrived at their determination “as a group” consisting of the

respondent, the school principal, the NAEP field staff representative, and, in some cases, classroom teachers.

Dimension 2: Involvement of NAEP field staff. Initially, NAEP field staff provided Raleigh with a list of students selected to participate in NAEP. The respondent attended an initial NAEP information meeting. By the time of the next visit from the NAEP field staff representative, the respondent and the principal had worked together to establish parameters regarding which students to include or exclude. The NAEP field staff representative confirmed the decision, but did not directly guide it.

Dimension 3: NAEP tools used. Decision makers at Raleigh used the LEP Questionnaire, the NAEP decision tree, and the NAEP inclusion criteria document “a little,” one respondent indicating that the decision tree was made available to teachers, the other simply that it provided a “guide.” The Raleigh respondent indicated that NAEP documents were used along with other sources of criteria including ITBS criteria and teacher judgment.

Dimension 4: Inclusion criteria used. The respondent and principal decided to use the same criteria for ELLs inclusion decisions as that used for ITBS, which is administered in the spring. “We made a general rule about who should or should not be included. We looked at the LEP questionnaire for any discrepancies and we looked at teacher opinion as expressed on the LEP form.” The criteria is as follows: If a student is in at least the third year of the bilingual program, he or she takes the ITBS; if in Year 1–2, the student takes Logramos. If the student is not progressing in English after Year 3, it is recommended that he or she take 2 more years in the bilingual program. If the student is still in the program in the fifth year, he or she may take a special bilingual test in English called the IMAGE (Illinois Measure of Annual Growth in English). Hence, as students enrolled in the bilingual program for “at least 3 years” were eligible to take ITBS, they were also considered eligible to participate in NAEP. In some cases, the two primary criteria, ITBS data and years in program, were overruled either by NAEP criteria or by teacher judgment. For instance, the respondent explained that “A couple of kids that we would have tested based on their ITBS levels were not tested based on NAEP criteria. In two cases, teachers said not to test students. In these cases, we deferred to teacher judgment.”

Dimension 5: Accommodations criteria used. The Raleigh respondent explained that, generally, only special education students (students with an IEP) were given accommodation. However, in at least one instance, the respondent felt that a student who was in the U.S. and bilingual program for less than 2 years “could manage the test with the accommodation of extended time.” Because NAEP offered this accommodation, the teacher wanted students to attempt the test. The respondent reiterated, however, that, ordinarily, LEP students do not receive any accommodation unless they are also identified as special education.

### *Bacon*

Dimension 1: School personnel involved. The respondent was case manager for special education students but, at the time of the interview, was overseeing the entire school.

Dimension 2: Involvement of NAEP field staff. This individual indicated that once the school received the list of students selected for participation in the NAEP, the principal and “NAEP administrator” (presumably, the NAEP field staff representative) made decisions regarding inclusion of students.

Dimension 3: NAEP tools used. At Bacon, LEP Questionnaires were completed (by their students’ teachers), yet the extent to which the LEP Questionnaire and the NAEP documents were used in making decisions was unclear.



Dimension 4: Inclusion criteria used. Once the school received the list of students selected for participation in the NAEP, the principal and “NAEP administrator” made decisions regarding inclusion of students based on LEP status. Students who would otherwise have been excluded from the math NAEP were included because of the availability of the bilingual test booklet during testing at Bacon. One student who used the bilingual booklet was judged as having “average English proficiency.” The respondent remarked that this accommodation gave the student a better chance to succeed on the NAEP.

Dimension 5: Accommodations criteria used. Students who would otherwise have been excluded from the Mathematics NAEP were included because of the availability of the bilingual test booklet during testing at Bacon. One student who used the bilingual booklet was judged as having “average English proficiency.” The respondent remarked that this accommodation gave the student a better chance to succeed on the NAEP.

### *Drake*

Dimension 1: School personnel involved. The Drake testing coordinator worked with the school’s bilingual coordinator. One respondent, a school counselor, indicated that he or she worked with the second respondent, the school’s bilingual coordinator, to decide which students were to be included or excluded from the NAEP. The counselor chose this arrangement because the bilingual coordinator knew students’ strengths and weaknesses. When asked what guidance, if any, NAEP field staff provided in making the final decision, the Drake respondent remarked, “they were available for help and we did sit with the NAEP facilitator who helped us decide if we had any questions. However, I know my students and they don’t!”

Dimension 2: Involvement of NAEP field staff. Decisions to include or exclude were made on a child-by-child basis prior to the visit from NAEP field staff. Respondents met face-to-face with the NAEP field staff representative, who helped them resolve questions.

Dimension 3: NAEP tools used. Decision makers at Drake used the LEP Questionnaire and NAEP decision tree “a little,” one respondent indicating that it was made available to teachers, the other simply that it provided a “guide.” At Drake, the criteria document was used “a little.” The respondent at Drake reported that the criteria document was used “as a guideline to follow,” but that the most salient factors in decision-making were the academic progress of the individual student.

Dimension 4: Inclusion criteria used. “When I first received that NAEP list, I reviewed the categories with the teachers and participated in their decision-making. The teachers provided test scores. The teachers used the NAEP questionnaire and answered it, but they were getting information from the school program and the school and curriculum supporting information as well as from state assessments (IMAGE). It is only through standardized data that we can measure language proficiency.” Decisions to include or exclude were made on a child-by-child basis based on (1) English language proficiency test, (2) student’s academic progress, which was determined by scores from state and local tests, and (3) whether or not the student “was ready and strong enough to assess.”

The respondent remarked, “We really looked at the individual child's progress and used documentation rather than NAEP information.” In some cases, if the child had been in the program for only one year but was really proficient, then that child was included. In other cases, if the child had been in the program for 3 years but did not meet these other criteria, he or she was excluded. The decision to exclude was made because “to have this child participate in this

assessment would have caused the student even more frustration,” and the respondent did not want to increase the student’s stress. A second excluded student was beginning-intermediate with proficiency in reading and math comprehension. The fact that this was the student’s second year in country and that he or she was considered to be very timid was also a factor in the decision to exclude.

Dimension 5: Accommodations criteria used. “For us, ‘accommodations’ means Special Ed (or 504). NAEP doesn’t differ from the state [criteria] because we use standardized tests and follow the same procedure.....Children in the bilingual programs have a tendency to score real well in math. We don’t give accommodations on some tests and not on others.” Interestingly, the Drake respondent also commented, “No accommodation was offered to students during the NAEP; however if this were a mandatory assessment, I would look at lower-level testing. But in this case, if I thought the student was capable, the student would take the on-level test.”

### District 3

The research team conducted interviews at eight schools in District 3: four schools for Grade 4, and four schools for Grade 8. Table 17 lists the eight schools that participated in the GW-CEEE study and the enrollment of these schools by grade level.

Table E13. Enrollment for Grade Level(s) in District 3 Schools Participating in Study

School	Grade	Student Enrollment for that Grade
Dickens	4	131
Hardy	4	109
Bronte	4	101
Bishop	4	162
Pynchon	8	430
Spenser	8	376
Chapman	8	401
Fletcher	8	428

Source: National Center for Education Statistics (2005)

Dimension 1: School personnel involved. When comparing the data in Table E14 for “number of decision makers” and “role/expertise to contribute to decision, we can see that, in six of the eight District 3 schools sampled in this study, school-based decisions were mainly completed by a single individual; however at the two other schools (Spenser and Bronte), decisions were completed by three or more people on a Language Proficiency Acquisition Committee (LPAC). The schools with only one of their staff members involved in the NAEP inclusion and accommodation decision-making process relied on their school administrators or their test coordinators, while in those schools that used three or more decision makers, decision makers were drawn from a variety of job roles, school administrator, test coordinators, language acquisition-related staff, and classroom teachers.

Table E14. School Personnel Involved-District 3

School Name	Grade	Number of Decision Makers			Role/Expertise to Contribute to Decision			
		One individual	Two	Three or more	Classroom instruction	Specialist in language acquisition	Test coordination	School administration
Dickens	4	x						x
Hardy	4	x				x		x
Bronte	4	x					x	
Bishop	4	x					x	
Pynchon	8	x					x	
Spenser	8			x	x	x	x	
Chapman	8			x	x	x	x	
Fletcher	8	x					x	
<b>Summary</b>		<b>6</b>	<b>0</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>6</b>	<b>2</b>

Dimension 2: Involvement of NAEP field staff. As shown in Table E15 when comparing the data for “timing of decision” and “involvement of NAEP field staff,” we can see that within District 3, decision makers at fourth of the eight schools (Hardy, Bishop, Spenser, and Chapman) made the ELL inclusion and accommodation decision prior to the pre-assessment visit from NAEP field staff. At Bishop, NAEP field staff confirmed the decision that had already been made, but at Spenser and Chapman, NAEP field staff adjusted the decision. At Spenser, the decision was adjusted because there were not enough assessment booklets. “When they [NAEP field staff] came to give the test and I gave them the list, I knew there were only thirty students to be tested, so I called them down. However, when I started calling names, they didn’t have answer documents [for all of the students on the list I had given them]. We lost a lot of time. I had to keep the kids who weren’t being tested in a different room to help with the kids who were not going to be tested.” But at Chapman, the decision was adjusted because too many ELLs had been excluded. “When the ladies [NAEP field staff personnel] came, they asked the school decision makers to review the students selected to participate in the mathematics NAEP and include more LEP students.” At the fourth school (Hardy), the Assistant Principal told us that he thought all ELLs would be tested in their native languages for both the Reading and Mathematics NAEP. After finding out this was not the case for the Reading NAEP, the NAEP field staff representative asked the Hardy Assistant Principal to consult with the Grade 4 ESL teacher (who consulted with other Grade 4 teachers) to determine which students might meaningfully take the test in English. The teachers could not identify any students that fit this description. In the end, the NAEP field staff representative made final inclusion and accommodation decisions for Hardy.

At four schools in District 3 (Dickens, Bronte, Pynchon, and Fletcher), the inclusion and accommodation decision was made during the NAEP field staff pre-assessment visit. The Dickens respondent, a school counselor, reported, “I passed out the questionnaire and let NAEP use my office as ‘ground zero.’ I was out of the picture. They did it all.” The Bronte decision

maker reported, “The people that came out with NAEP told me who they were going to test. I really had no input on that decision. They were the ones who made the decision [based on the LEP Questionnaires that had been completed by the teachers].” At Pynchon, the respondent, a test coordinator reported, “The NAEP field staff representative did most of work and would consult with the respondent regarding the ESL students.

Table E15. Involvement of NAEP Field Staff-District 3

School Name	Grade	Timing of Decision					Involvement of NAEP Field Staff				
		Prior to NAEP field staff visit	During NAEP field staff visit	After NAEP field staff visit	Just before test administered	Unclear/ Do not know	No role	Confirmation of prior decision	Gave strong guidance	Adjusted prior decision just before test	Unclear/ Do not know
Dickens	4		x						x		
Hardy	4	x			x					x	
Bronte	4		x						x		
Bishop	4	x						x			
Pynchon	8		x						x		
Spenser	8	x			x		x			x	
Chapman	8	x			x		x			x	
Fletcher	8		x						x		
<b>District 3 Summary</b>		<b>4</b>	<b>4</b>	<b>0</b>	<b>3</b>	<b>0</b>	<b>2</b>	<b>1</b>	<b>4</b>	<b>3</b>	<b>0</b>

Dimension 3: NAEP tools used. As shown in Table E15 and Table E16, when comparing the data for “central document(s) used with decision-making” and “NAEP tool used,” we can see that, while it appears that decision makers at all eight of the schools sampled in District 3 relied on NAEP documents, especially the LEP Questionnaire, when making their decision, the manner in which the documents were used varied in relation to the involvement of NAEP in their decisions. For example, at the two schools in District 3 in which NAEP criteria were used, but the NAEP field staff played a smaller role, decision makers reporting using all three NAEP documents. At Bishop and Chapman, decision makers were less reliant on the NAEP field staff when making the inclusion and accommodation decision. This lack of reliance on the NAEP field staff was possible because decision makers at these two schools used the LEP Questionnaire in conjunction with the decision tree and NAEP criteria document when making inclusion and accommodation decisions for ELLs.

Decision makers at Dickens and Hardy relied on NAEP to use the completed LEP Questionnaire to make the ELL inclusion and accommodation decision. Similarly, the decision maker at Pynchon reported that she relied on the NAEP field representative to do “most of work,” but was consulted regarding the ESL students. During these discussions with the NAEP field staff representative, the Pynchon decision maker relied on the LEP Questionnaire but did not use the decision tree nor NAEP Criteria for Inclusion.

In contrast, decision makers at Fletcher and Bronte said they relied more on the NAEP decision tree and the NAEP Criteria for Inclusion than the LEP Questionnaire – mainly because they relied on the NAEP field staff representatives to make the actual decision for them. (Although it was unclear the extent to which Fletcher decision makers used the completed LEP

Questionnaire during decision-making. The Fletcher respondent worked with the NAEP field staff representative and based decisions on the NAEP decision tree and the NAEP inclusion criteria document. It was unclear whether the LEP Questionnaire was used during decision-making.) Somewhat similar to the Fletcher and Bronte decision makers, Spenser decision makers indicated not remembering using the LEP Questionnaire very much “but the teachers did fill it out.” Prior to the NAEP pre-assessment visit, Spenser decision makers relied on their committee discussions around the NAEP Criteria for Inclusion, but did not use the LEP Questionnaire.

Table E16. NAEP Tools Used-District 3

School Name	Grade	Central Document(s) Used With Decision Making				NAEP Tool Used		
		NAEP tools	State tools	School or district test scores	Staff opinion	NAEP LEP Questionnaire	NAEP Decision tree	NAEP criteria document
Dickens	4	x				x		
Hardy	4	x				x		
Bronte	4	x				x	x	x
Bishop	4	x		x	x	x		x
Pynchon	8	x				x		
Spenser	8			x				x
Chapman	8	x				x	x	x
Fletcher	8	x				unclear	x	x
<b>District 3 Summary</b>		<b>7</b>	<b>0</b>	<b>2</b>	<b>1</b>	<b>6</b>	<b>3</b>	<b>5</b>

Dimension 4: Inclusion criteria used. Within District 3, inclusion decisions were most often impacted by the amount of time the student had received English language instruction. As detailed in Table E17, six of the eight schools reported use of criteria similar to “time in LEP program.” Yet, these decisions were often impacted by the student’s placement in the school’s bilingual program in opposite ways. The Hardy decision maker strove to exclude more ELLs based on the amount of instruction they had received in English, while the Dickens principal had asked the school counselor (the Dickens decision maker) to include all bilingual Grade 4 because “in Grade 5 everything is in English.” When the Dickens staff were told the test was only in English, they had to decide “how the child would do,” i.e., whether or not the child could meaningfully participate in the NAEP.

There was variability regarding which criteria decision makers emphasized during the inclusion decision-making process. Application of criteria related to “amount of instruction in English” was also supplemented by data on the student’s English language proficiency and/or the student’s academic performance. For example, Pynchon decision makers relied on academic performance, time in the US, while Bronte decision makers relied on academic performance in combination with amount of instruction in English to determine inclusion of students listed on the administrative schedule. However, it should be noted that the Pynchon and Bronte told us that the NAEP field staff had made the inclusion decision. At Pynchon, students who were new to country this year were excluded from the NAEP. If ESL level (2, 3, or 4) and grades [below C] indicated that the student was not progressing, the student was excluded. At Bronte, ELLs

were excluded on the administration schedule due to level of instruction and class lists. At yet another school, Fletcher, the decision maker used student test scores

Only one decision maker in District 3 (at Chapman) told us they had relied on state policy to include all students. Chapman decision makers told us that the school had attempted to “stick with consistency of how we exempt” ELLs and special education students. “We assumed that if they could take the TAKS and the Stanford 10, they would take the NAEP.” Hence, as a general rule, any Chapman student exempt from regular testing would not be included in NAEP. These Chapman students were most likely to be at intermediate or advanced English language proficiency levels.

Table E17. Inclusion Criteria Used-District 3

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Teacher opinion	Unclear/ Do not know
Dickens	4				x		
Hardy	4	x			x		
Bronte	4		x		x		
Bishop	4	x	x				
Pynchon	8		x	x	x		
Spenser	8	x	x		x		
Chapman	8		x	x			
Fletcher	8	x	x		x		
<b>District 3 Summary</b>		<b>4</b>	<b>6</b>	<b>2</b>	<b>6</b>	<b>0</b>	<b>0</b>

Dimension 5: Accommodations criteria used. As shown in Table E 18, within District 3, the assignment of accommodations was more likely to be determined by criteria related to English language acquisition (English language proficiency, time in U.S., time in LEP program) rather than non-ELL-responsive criteria such as whether or not the student had an IEP. The availability of the bilingual booklet influenced the accommodations decisions made at Bronte and Bishop. For example, at Bronte, ELLs were given bilingual booklets for the mathematics NAEP. “They [NAEP] did have materials in Spanish so we did accommodate children that were selected.” The Bishop respondent indicated that the availability of a bilingual test booklet in Math made the decision to include certain students easier.

However, the lack of bilingual booklets impacted ELLs at two schools. Because fewer than needed bilingual booklets were available at Hardy and Pynchon, some ELLs were excluded. The Hardy respondent told us “Some students were excluded because she did not have enough bilingual materials. They [NAEP field staff] are the ones who chose who to test.” But the major influence at Hardy was the view that if students are included, they do not need to receive accommodations. Initially, the Hardy Assistant Principal explained, “We have no LEP Special Ed students in 4<sup>th</sup> grade. Accommodations were decided by NAEP [NAEP field staff representative].” When asked further, the Assistant Principal replied: “All 4<sup>th</sup> graders get tested. If a student is in a regular classroom, that student should not have been excluded. [Our] 4<sup>th</sup> grade students’ reading skills are in Spanish. Some students were excluded because she did not have enough bilingual materials. They [NAEP field staff] are the ones who chose who to test.”

The Dickens decision maker, who had made her inclusion decision based on the ELL's placement in the school's bilingual program, relied on English language accommodations. The Spenser decision maker based the accommodations decision based on its use in the classroom. She explained, "If teacher had been using accommodation, then the committee would have allowed the student to use that accommodation on the NAEP if it were available. "We received information [for the NAEP] in January. We were supposed to be doing certain accommodations, especially in math. But the teachers weren't doing these accommodations. The only accommodation they were doing was translation. On the day of the test, if our teachers had been able to use this technique they could have helped students better. I know that this coming year there were eight accommodations they could be using so they are going to be using these accommodations."

Pynchon decision makers offered accommodations based on English language proficiency and academic performance, but they also offered accommodations to ELLs who had special education needs. To highlight the difference between the NAEP and state policy on accommodations, the Bronte respondent explained that "Whereas NAEP offers accommodations, the state exempts students only through special education or parent refusal."

It was unclear which accommodations criteria were used by Chapman and Fletcher decision makers. Chapman decision makers did offer bilingual booklets to some ELLs, but did not say which criteria were involved in the accommodations decision. An "advanced LEP" student at Fletcher was included in the NAEP with accommodation in Math, but the decision maker did not give clear information on the accommodations criteria that had been used.

Table E18. Accommodations Criteria Used-District 3

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Offered no accommodations	IEP	Unclear/ Do not know
Dickens	4		x					
Hardy	4					x		
Bronte	4				x			
Bishop	4	x			x			
Pynchon	8	x	x					
Spenser	8		x		x			
Chapman	8							x
Fletcher	8							x
<b>Summary</b>		<b>3</b>	<b>3</b>	<b>0</b>	<b>3</b>	<b>1</b>	<b>0</b>	<b>2</b>

### *Dickens*

Dimension 1: School personnel involved. The respondent, a school counselor, said that decisions were made by NAEP field staff. "I passed out the questionnaire and let NAEP use my office as 'ground zero.' I was out of the picture. They did it all."

Dimension 2: Involvement of NAEP field staff. The respondent, a school counselor, reported that decisions were made by NAEP field staff during their visit to the school. Teachers returned completed the LEP Questionnaires and returned them to NAEP field staff personnel, who made inclusion decisions based on teachers' responses.

Dimension 3: NAEP tools used. Teachers filled out the questionnaire and submitted it to the NAEP field staff who used the questionnaire to include or exclude students. However, the respondent also felt that the questionnaire was too lengthy and that the decision makers had the appropriate information about the students without completing the questionnaire. Respondents at Dickens reported that neither the decision tree nor the NAEP inclusion criteria document provided by NAEP were used.

Dimension 4: Inclusion criteria used. Decision makers at the school appear to have included ELLs in the NAEP based on the criteria "time in LEP program." The principal had asked that all bilingual Grade 4 students be tested because "in Grade 5 everything is in English." The respondent, a school counselor explained, "I test all bilingual fourth graders. If they test high enough on the state test (TAKS) then they can exit the program because in fifth grade, everybody goes to English [school is PreK—5]. The more we can exit the program the better." When the Dickens staff were told the test was only in English, they had to decide "how the child would do," i.e., whether or not the child could meaningfully participate in the NAEP. The respondent distributed LEP Questionnaires to teachers. Teachers returned completed LEP Questionnaires to NAEP field staff personnel, who made inclusion decisions based on teachers' responses. In the end, all the LEP students in the school were included in the math NAEP.

Dimension 5: Accommodations criteria used. In considering which accommodation a student should use during the mathematics NAEP, the teacher and NAEP (i.e., NAEP field staff representative) decided reading instructions aloud would help students based on classroom observation and language proficiency (especially oral and listening proficiency).

*Hardy*

Dimension 1: School personnel involved. The respondent, the Assistant Principal, thought all ELLs would be tested in their native languages for both the Reading and Mathematics NAEP. After finding out this was not the case for the Reading NAEP, the NAEP field staff representative asked the respondent to consult with the Grade 4 ESL teacher (who consulted with other Grade 4 teachers) to determine which students might meaningfully take the test in English. The teachers could not identify any students that fit this description. In the end, the NAEP field staff representative made final inclusion and accommodation decisions.

Dimension 2: Involvement of NAEP field staff. "[NAEP field staff] called and said they were coming out to meet with me and that never happened until 2 days before the test... We had to do it in a hurry. It was not easy to decide. Some of the things I should have had done before she [NAEP field staff] got here. I was not given a lot of detail about testing, no. I should have had more information earlier." As mentioned above, in the end, the NAEP field staff representative made final inclusion and accommodation decisions.

The Assistant Principal noted, "I don't know how much flexibility you have with testing dates, but she [NAEP field staff representative] had that test [NAEP] set up for the same week that we were taking a field test with TAKS [state test] and the 4<sup>th</sup> grade was involved. I don't know how far in advance you're going to look for this and if we're going to do this next year or not, but if so, please check on a calendar for our testing."



Dimension 3: NAEP tools used. The Hardy Assistant Principal reported being told by NAEP that if all students were going to participate in the NAEP, then questionnaires should be filled out only for those students who would be taking the test in English. She explained, “When we got the questionnaire, the ESL teacher started using it and had lots [of students] to do the test. Then we were told that if everyone was going to take the test in the native language, it was not necessary to fill out the questionnaire... ‘just fill out the ones in the native language.’” The respondent at Hardy did not know how the decision tree or NAEP inclusion criteria document were used.

Dimension 4: Inclusion criteria used. Decision makers thought ELLs would be included in NAEP according to the language of academic instruction. At the beginning of the school year decision makers at Hardy had used English language proficiency to assign ELLs to particular LEP/bilingual programs. “Students who do not score high enough on the SAT10 (40% in reading and language) are automatically classified as LEP and that is the program they go into, unless parents waiver them out.” Students in the bilingual program are taught primarily in Spanish through Grade 4. They receive ESL for only an hour per day. The school-based decision makers assumed there would be able to assign bilingual booklets (since they thought all ELLs would be tested in their native language). No students were excluded from the Math test because of the availability of Spanish language test materials: “This made a huge difference in testing in our program.”

But during the Reading NAEP, there was some confusion regarding the use of bilingual booklets for that assessment. It appears that the respondent never fully understood the NAEP criteria that ELLs could not use a bilingual booklet for the Reading NAEP. The Assistant Principal explained, “My understanding was, in the beginning, that all of the LEP students would be tested, but then I found out there were not enough test booklets submitted so [NAEP field team] had to redo the test list. There was a complete breakdown of the list. I was not really clear about what was going on. When I was first contacted, all my LEP students were to be tested in their native language. Then [NAEP field staff] came through and said there were no materials. They said, ‘That’s not really how it is. I need your bilingual teachers to let me know which one of their students can take the test in English because most of my books are in English and most of my testers are English testers.’” So the respondent consulted with bilingual teachers to determine which students could meaningfully take the test in English.

The Assistant Principal continued, “My school has a bilingual grant. Our kids are kept on the bilingual program for 4 years. They are exited in the 5<sup>th</sup> year. We have no ESL transition here. Hopefully by 4<sup>th</sup> grade, the students are proficient enough to test in English in 5<sup>th</sup> grade. They receive all major instruction in their native language and then have one hour of ESL a day. The students in this program are not proficient enough to that the test. The bilingual teachers said it wouldn’t be fair. That’s when we had a breakdown with the testing....Fourth grade students who exit are a small percentage. My percentages of not-exited are high because we don’t have ESL transitional.” The Assistant Principal relied heavily on a Grade 4 ESL teacher for information about students. The Grade 4 teacher met with the other Grade 4 teachers to decide which students had sufficient English skills to take the test. The teachers could not identify any students that fit this description. “She met with the other 4<sup>th</sup> grade teachers to find someone strong enough in English to test but they couldn’t. All the kids tested were tested as exceptions....That is what the teachers thought was happening.” In the end, the NAEP field staff representative made the inclusion decisions.

Dimension 5: Accommodations criteria used. Initially, when asked about accommodations, the Assistant Principal explained, “We have no LEP Special Ed students in 4<sup>th</sup> grade. Accommodations were decided by NAEP [NAEP field staff representative].” When asked further, the Assistant Principal replied: “All 4<sup>th</sup> graders get tested. If a student is in a regular classroom, that student should not have been excluded. [Our] 4<sup>th</sup> grade students’ reading skills are in Spanish. Some students were excluded because she did not have enough bilingual materials. They [NAEP field staff] are the ones who chose who to test.” The Assistant Principal also indicated that the school wanted to provide a spelling accommodation but this accommodation was not available through NAEP.

### *Bronte*

Dimension 1: School personnel involved. The respondent, a Reading First coach, reported that NAEP sent a list of students selected for participation in NAEP (presumably in the NAEP administration schedules). “This is my second year with NAEP as the testing coordinator. I played no role. They [NAEP field staff representative] asked me for class lists when they did the testing. They used them to decide who would take the test.”

Dimension 2: Involvement of NAEP field staff. “The people that came out with NAEP told me who they were going to test. I really had no input on that decision. They were the ones who made the decision [based on the LEP Questionnaires that had been completed by the teachers].”

Dimension 3: NAEP tools used. Respondents at Bronte reported that the LEP Questionnaire was used “some.” The Bronte respondent simply indicated that the LEP Questionnaire was given to teachers “who teach in Spanish” to fill out. The Bronte respondent indicated that the decision tree and NAEP inclusion criteria document were used as a reference.

Dimension 4: Inclusion criteria used. The respondent indicated that the NAEP field staff representative made inclusion and accommodation decisions based on the class lists that the respondent provided. In completing the questionnaire, teachers considered the “total assessment package” as well as their experience of the students’ written and oral performance in class. Students were excluded on the administration schedule due to level of instruction. One student, who technically had not yet qualified for Special Ed but he could not read, was excluded because “he has no proficiency.... We did use some of NAEP criteria [for this student], but at this point in the year, it was not hard to decide.”

Dimension 5: Accommodations criteria used. For the mathematics NAEP, ELLs were given bilingual booklets. “They [NAEP] did have materials in Spanish so we did accommodate children that were selected.” The decision to make this accommodation available to the student was based on the level of the student’s instruction (whether in English or Spanish). The SASI, the Texas test that identifies LEP students, was used to determine which ELLs would be allowed to use bilingual booklets for NAEP: If student received language instruction in English then they received an English booklet; if instruction was given in Spanish, they received a bilingual booklet.

The Bronte respondent told that, whereas NAEP offers accommodations, the state exempts students only through special education or parent refusal. “To qualify for Special Ed is the only way to get an accommodation [on the state assessment]. If a student is to be exempted, the state must assess that student. IEP is only mechanism for granting exemption or allowing accommodation.”

### *Bishop*

Dimension 1: School personnel involved. The respondent, lead science teacher and science coordinator, was appointed NAEP coordinator at the school by the consultant-principal because, as an enrichment teacher, the respondent taught all students and had access to students' language development portfolios, which are updated yearly.

Dimension 2: Involvement of NAEP field staff. Inclusion decisions were made prior to the NAEP field staff visit. In consultation with teachers, the respondent gathered appropriate data and then met twice with the NAEP field staff representative. The final meeting was to "make adjustments," presumably to the final decisions regarding inclusion and accommodation of LEP students.

Dimension 3: NAEP tools used. At Bishop the LEP Questionnaire was used "as a guide in conjunction with other knowledge/information." Respondents at Bishop reported that the decision tree provided by NAEP was not used at all. At Bishop, the NAEP inclusion criteria document was "compared and contrasted with information on hand."

Dimension 4: Inclusion criteria used. A science enrichment teacher coordinated the decision-making process at the school. In making decisions, the respondent had access to students' language development portfolios, which are updated yearly, and hence could document students' language progression. She asked classroom teachers to fill out LEP Questionnaires for selected students and to make a recommendation regarding inclusion and accommodation. Teachers based their recommendation on classroom observation and Stanford 10 scores. The respondent considered teacher input, HISD questionnaire, and respondent's knowledge of students along with the LEP Questionnaire and the NAEP criteria. Teacher judgment was also instrumental in making inclusion decisions. For instance, it was decided that an LEP student described as a beginning level of English proficiency would experience too much stress if he or she participated in the NAEP. The respondent observed that "Some days she's better than others. We didn't want to stress her out with this test. She is borderline, but we decided that the test would be unnecessarily frustrating."

Dimension 5: Accommodations criteria used. The respondent indicated that the availability of a bilingual test booklet in Math made the decision to include certain students easier. At points there was minor disagreement about whether to allow students to use accommodations. In one instance, a teacher recommended that a student be accommodated with a bilingual test booklet (for Math). Because the respondent considered the student to be bright and to have "pretty good" proficiency in English, the respondent was not certain why the classroom teacher decided to allow the student to use the accommodation but surmised that it was to make the student "more comfortable" with the assessment.

### *Pyncheon*

Dimension 1: School personnel involved. The NAEP field staff representative did most of work and would consult with the respondent regarding the ESL students. The respondent stated that "we [the respondent – a test coordinator – and NAEP field staff representative] selected students" and then gave the LEP Questionnaires to the students' ESL teachers, who then filled them out. The reading teachers completed LEP Questionnaire for students selected for Reading NAEP.

Dimension 2: Involvement of NAEP field staff. The respondent explained that inclusion decisions were made two weeks before testing with the NAEP field staff representative. The

NAEP field staff representative did most of work and would consult with the respondent regarding the ESL students.

Dimension 3: NAEP tools used. Respondents at Pynchon reported that the LEP Questionnaire was used “some.” The respondent stated that “we [the respondent and NAEP field staff representative] selected students” and then gave the LEP Questionnaires to the students’ ESL teachers, who then filled them out. Reading teachers completed LEP Questionnaire for students selected for Reading NAEP. Respondents at Pynchon reported that the decision tree and NAEP inclusion criteria document provided by NAEP were not used.

Dimension 4: Inclusion criteria used. The teachers completed the questionnaire, but the respondent did not examine it, although the NAEP Coordinator did. The respondent indicated that students who were new to country this year were excluded from the NAEP. If ESL level (2, 3, or 4) and grades [below C] indicated that the student was not progressing, the student was excluded. Teachers expressed surprise when, after reviewing the questionnaire, the NAEP field staff representative decided to exclude certain students.

Dimension 5: Accommodations criteria used. Because there was no native language accommodation available for the Reading NAEP—no bilingual booklet or dictionary—one LEP student discussed in the interview “couldn’t possibly take the test.” For the Reading NAEP, everything was in English. For the Mathematics NAEP, if the student made grades B or C, he or she could participate in NAEP but would need accommodation. To make this determination, the NAEP field staff representative and the respondent considered ESL levels and grades. Because of English proficiency level, it was determined that one group of students would have extended time. These students were tested separately from the larger groups of NAEP participants. The NAEP coordinator made the decision and also looked at those students with accommodations (if these students were classified as special education students).

### *Spenser*

Dimension 1: School personnel involved. The respondent, the bilingual counselor, was in charge of testing and is a member of the Language Proficiency Acquisition Committee (LPAC) which decides the program level and schedule for each ELL at the school. “I had already made the decision by the time [NAEP field staff representative] had come.” She explained, “I sat down with the [NAEP] representative who came here and then I also talked to some of the teachers. She [the NAEP field staff representative] told me to check with them [the teachers]. They knew who should and shouldn’t be tested.”

Dimension 2: Involvement of NAEP field staff. NAEP field staff personnel helped in decision-making, but the respondent had already made the decisions working with classroom teachers. When the NAEP field staff representative left, the respondent checked with the teacher about which students to include or exclude.

But, some changes needed to be made to the original decision when, on the day of the test, Spenser had to exclude more students because there were not answer booklets for some of the students on the list. “When they [NAEP field staff] came to give the test and I gave them the list, I knew there were only thirty students to be tested, so I called them down. However, when I started calling names, they didn’t have answer documents [for all of the students on the list I had given them]. We lost a lot of time. I had to keep the kids who weren’t being tested in a different room to help with the kids who were not going to be tested. They did not have answer documents for all the kids on that list. I thought they were going to us blank documents for those students,

but they didn't. Without the documents, there were so many children who were not going to be tested that, at the last minute, we had to find a place for all these kids."

Dimension 3: NAEP tools used. The Spenser respondent indicated not remembering using the LEP Questionnaire very much "but the teachers did fill it out." At Spenser, the NAEP inclusion criteria document was used by teachers and "definitely helped" them make decisions. Respondents at Spenser reported that the decision tree provided by NAEP was not used at all.

Dimension 4: Inclusion criteria used. As part of LPAC, the respondent worked with teacher and student information to make decisions. They sat down with each teacher and looked at each child individually. "We look at entry date into the country and information from the classroom teacher. There are also test scores that mark progress. This year the students took TAKS in math; for reading they took RPTE.... Students also give five samples of writing. An observation protocol is also considered: It is the teachers who decide if the student is intermediate, beginning. All of this helped us to decide whether the child is progressing."

Dimension 5: Accommodations criteria used. If teacher had been using accommodation, then the committee would have allowed the student to use that accommodation on the NAEP if it were available. "We received information [for the NAEP] in January. We were supposed to be doing certain accommodations, especially in math. But the teachers weren't doing these accommodations. The only accommodation they were doing was translation. On the day of the test, if our teachers had been able to use this technique they could have helped students better. I know that this coming year there were eight accommodations they could be using so they are going to be using these accommodations."

Translation of directions was the only accommodation used for math on the NAEP. The respondent felt that more students might have been accommodated if the school had had more time to determine which accommodations were available for NAEP. "We didn't really have time to sort all this out as testing came in January. We might have offered more student accommodations with bilingual information if we'd had more time. Accommodations were easier in math because there are glossaries, pictures, etc. to help students."

### *Chapman*

Dimension 1: School personnel involved. The respondent, a counselor and Dean of ESL, "relied heavily on teacher input." The respondent and classroom teachers were all on the LPAC, although in making inclusion decisions for NAEP they were not acting in an official LPAC capacity. They "relied heavily on teacher input."

Dimension 2: Involvement of NAEP field staff. The respondent indicated that inclusion and accommodation decisions were made prior to the NAEP field staff visit. The respondent got some guidance but "knew there was some kind of dilemma": The school ended up excluding so many students that there were not enough students to constitute a valid sample. "When the ladies [NAEP field staff personnel] came, they asked the school decision makers to review the students selected to participate in the mathematics NAEP and include more LEP students."

Dimension 3: NAEP tools used. At Chapman, the decision tree was "heavily used" since it listed the grade levels. The NAEP inclusion criteria document was also "very helpful." However, the LEP Questionnaire was not used that much since they were relying on the other two documents the most.

Dimension 4: Inclusion criteria used. The Chapman respondent indicated that in making decisions for inclusion in the NAEP the school had attempted to "stick with consistency of how we exempt" ELLs and special education students. "We assumed that if they could take the TAKS

and the Stanford 10, they would take the NAEP.” Hence, as a general rule, any student exempt from regular testing would not be included in NAEP. These students were most likely to be at intermediate or advanced English language proficiency levels. Two primary criteria govern inclusion in regular testing: time in U.S. and scores on the RPTE. LEP students who had not been in the U.S. for 3 or more years and who did not score high enough on the RPTE would be excluded. The respondent reported that most students included in the NAEP were also included in regular testing.

The Chapman respondent also noted, “The list of students to take the test [NAEP sent us] was very outdated. A lot of students they showed as LEP had already exited the program. It seemed that the list was outdated because most of our new arrivals were not on the list – so most of the LEP data was very inaccurate.”

The respondent indicated that, because they were asked by NAEP field staff personnel to review the list of selected students for inclusion in the Math NAEP, decision makers looked more closely at academic level rather than simply considering the student’s time in the country for inclusion decisions for this test.

Dimension 5: Accommodations criteria used. The respondent also reported that a group of students at the school took the mathematics NAEP using the Spanish booklet but did not provide any details about criteria used in making the accommodations decision.

### *Fletcher*

Dimension 1: School personnel involved. The school Counselor and the NAEP field staff representative were the primary decision makers. The respondent was a counselor responsible for checking the progress of students, especially those being monitored. “It was just me. I didn’t consult other school staff.”

Dimension 2: Involvement of NAEP field staff. Inclusion decisions were made during the NAEP field staff visit to the school. Teachers participated insofar as they completed the LEP Questionnaires; however, the respondent reported that the questionnaires were not used in making inclusion and accommodation decisions. “The classroom teachers answered the questions. I did not look at them.” Nonetheless, in making decisions, the respondent worked with the NAEP field staff representative and based decisions on the NAEP decision tree and the NAEP inclusion criteria document. The Fletcher respondent commented, “We made the decisions before the testing. We [respondent and NAEP field staff representative] worked in tandem. She gave me good guidance.” [Additional calendar note: “We did the writing part of TAKS one day before the NAEP. It was chaos. The state gives us a calendar with dates when these could be administered.”]

Dimension 3: NAEP tools used. In making decisions, the respondent worked with the NAEP field staff representative and based decisions on the NAEP decision tree and the NAEP inclusion criteria document. It was unclear whether the LEP Questionnaire was used during decision-making.

Dimension 4: Inclusion criteria used. The respondent worked with the NAEP field staff representative to make inclusion and accommodation decisions informed by the NAEP decision tree and the NAEP inclusion criteria document. In addition to test scores, the respondent felt that it was sometimes necessary to consult such sources as records from the last school attended or the home language survey. In relating the decision-making process for students included in the NAEP, the respondent referenced the fact that the students discussed were advanced LEP

students who had exited the ESL program and were now in the monitored program. One student excluded from Reading NAEP was judged “not sufficiently proficient to take the test.” The student was in the ESL program and had been “in and out of the U.S. and therefore in and out of the program.” Furthermore, the student had not yet met the exit criteria and “barely knows how to read English.” In the end, “lack on English proficiency drove classroom teacher decision” regarding this student’s participation in the test.

Dimension 5: Accommodations criteria used. The respondent discussed an “advanced LEP” student included in the NAEP with accommodation in Math, but did not give clear information on the accommodations criteria that had been used.

### District 4

The research team was able to conduct interviews at eight schools in District 4: four elementary schools and four middle schools. Table E19 provides the total enrollment in Grades 4 and 8 of each LA school.

Table E19. Enrollment for Grade Level(s) in District 4 Schools Participating in Study

School	Grade	Student Enrollment for that Grade
Tully	4	174
Tacitus	4	159
Livy	4	207
Homer	4	335
Smithers	8	727
Ovid	8	702
Josephus	8	952
Hesiod	8	613

Source: National Center for Education Statistics (2005)

Dimension 1: School personnel involved. When comparing the data in Table E20 for “number of decision makers” and “role/expertise to contribute to decision, we can see that, within District 4, school-based decision-making tended to be completed by an individual or pair, but not by three or more individuals from the school. The decision makers were not school administrators, but testing coordinators and language acquisition specialist.

Table E20. School-Based Decision Makers-District 4

School Name	Grade	Number of Decision Makers			Role/Expertise to Contribute to Decision			
		One individual	Two	Three or more	Classroom instruction	Specialist in language acquisition	Test coordination	School administration
Tully	4	x						x
Tacitus	4		x			x	x	
Livy	4	x					x	
Homer	4		x			x	x	
Smithers	8	x					x	
Ovid	8		x x*			x x*		
Josephus	8		x			x	x	
Hesiod	8	x					x	
<b>Summary</b>		<b>4</b>	<b>5</b>	<b>0</b>	<b>0</b>	<b>5</b>	<b>6</b>	<b>1</b>

\*The respondents at Ovid were a bilingual coordinator and an ELD coordinator

Dimension 2: Involvement of NAEP field staff. Respondents at Tully, Tacitus, Josephus, and Smithers reported that school personnel had a minimal role in decision-making and adhered to the student lists provided by NAEP. For example, the Tacitus respondents explained that they had met with the respondents prior to testing and went over the packet of NAEP materials. “Testing went smoothly” since “everything was “precoded,” and that they “went with whatever NAEP [field staff] recommended.” The Smithers decision makers explained that NAEP had provided a list of pre-selected students and “tested all.”

The Hesiod decision maker made the inclusion decision, but did not spend too much time worrying over the decision since “there isn’t any [accountability]. It’s a general report of the country. We don’t think about it. We’re not concerned about it.” To make the decision, the Hesiod decision maker used NAEP guidance and the LEP questionnaire. She explained, “A list came to me and then I used the teacher questionnaire. We did not consult with anyone (to include/exclude)...[NAEP] is fairly easy to implement [since] NAEP sends their own proctors.”

As shown in Table E21 when comparing the data for “timing of decision” and “involvement of NAEP field staff,” NAEP field staff appears to have confirmed the decisions made by Homer and Livy – even though these decision makers did not follow NAEP criteria or use NAEP materials. Homer decision makers did not use the completed LEP questionnaires, but relied instead on ELD levels (specifically the CELDT). Decision makers at Livy made their decisions prior to the NAEP pre-assessment visit, using the state criteria rather than NAEP criteria. Ovid decision makers told us that NAEP had provided some guidance, but it was clear as to whether the decision had been made prior or during the NAEP pre-assessment visit.



Table E21. Involvement of NAEP Field Staff-District 4

School name	Grade	Timing of Decision					Involvement of NAEP Field Staff				
		Prior to NAEP field staff visit	During NAEP field staff visit	After NAEP field staff visit	Just before test administered	Unclear/ Do not know	No role	Confirmation of prior decision	Gave strong guidance	Adjusted prior decision just before test	Unclear/ Do not know
Tully	4		x						x		
Tacitus	4		x						x		
Livy	4	x						x			
Homer	4			x				x			
Smithers	8		x						x		
Ovid	8					x			x		
Josephus	8		x						x		
Hesiod	8			x				x			
<b>Summary</b>		<b>1</b>	<b>4</b>	<b>2</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>3</b>	<b>5</b>	<b>0</b>	<b>0</b>

Dimension 3: NAEP tools used. Table E21 shows that, of the three documents distributed by NAEP field staff to inform decision-making, the LEP Questionnaire was the only one to have been used in seven of eight schools sampled for the District 4 case study. At the seven schools the questionnaire was completed by the teachers most familiar with the selected students and then given back to decision makers. However, decision makers at the seven District 4 schools (Tully, Tacitus, Livy, Homer, Smithers, Ovid, Josephus, and Hesiod) completed the LEP Questionnaire. Three of eight schools ignored LEP Questionnaire in their decision-making (Tacitus, Livy, and Homer). Four other schools completed the LEP Questionnaire and gave it to the NAEP field staff representative to use (Tully, Smithers, Josephus, and Hesiod). The eighth school Ovid did not indicate the extent to which the LEP Questionnaire, decision tree, or Criteria for Inclusion for used by decision makers.

Table E22 reveals that, while NAEP documents (in the form of the LEP Questionnaire) were used more than any other document to make the decisions, even then the LEP Questionnaire was not used that much. The NAEP decision tree and the criteria document were not used by District 4 decision makers. Respondents at the eighth school Ovid did not indicate the extent to which the LEP Questionnaire, decision tree, or NAEP Criteria for Inclusion was used by decision makers, but it appears the NAEP tools were used.

Table E22. NAEP Tools Used -District 4

		Central Document(s) Used with Decision-Making				NAEP Tool Used		
		NAEP tools	State tools	School or district test scores	Staff opinion	NAEP LEP Questionnaire	NAEP Decision tree	NAEP criteria document
School Name	Grade							
Tully	4	x				x		
Tacitus	4	x						
Livy	4		x					
Homer	4			x		x		
Smithers	8	x				x		
Ovid	8			x		unclear	unclear	unclear
Josephus	8	x				x	x	x
Hesiod	8	x				x		
<b>Summary</b>		<b>5</b>	<b>1</b>	<b>2</b>	<b>0</b>	<b>4</b>	<b>0</b>	<b>0</b>

Dimension 4: Inclusion criteria used. As mentioned previously in Table E22, in four of the eight schools sampled in District 4, either decision makers or the NAEP field staff representative used the information gathered via the LEP Questionnaire to make the inclusion decision. As a result, Table E23 supplements the data in Table E22, showing that decision makers who had used the LEP Questionnaire to gather information on the students (i.e., Hesiod, Smithers, Tully), told the research team they had used English language proficiency as one of the criterion inclusion decisions for ELLs. For example, the Hesiod decision maker reported, “Teachers know each student’s language level because this information is printed in roll books. The respondent speculated that teachers’ decisions were probably fairly easy to make because of their knowledge of the students’ ELD levels.”

Even though at Homer teachers filled out LEP Questionnaires, the decision maker explained that the LEP Questionnaires had no real role in deciding which students were included or excluded because decisions were based solely on ELD levels. Once the Title I coordinator and a bilingual coordinator received the administration schedule from NAEP, they automatically eliminated (excluded) students with English Language Levels of 1 or 2.

The pressure to include all ELLs in testing seems to have played a role at Tacitus and Livy. In fact, it appears that the Tacitus respondents thought the initial list of pre-selected students for the NAEP was the final list of students to be selected. The Tacitus respondents said they “looked at” the criteria document and “followed” it “but tested all.” Decision makers at Ovid “worked with teachers on the list provided by NAEP “to include as many students as possible.” Ovid decision makers used English language proficiency to inform their decision about student placement, but in most cases went with the initial list sent by NAEP since they thought that was the final list.

The decision maker at Livy followed the STAR [California’s Standardized Testing and Reporting program] guidelines for the ELL inclusion decision. He said, “It’s a little more practical. It’s a little harder for us to have different criteria [for each large-scale assessment].” He explained, “We kind of followed the STAR testing criteria because they include everybody.

Because it is anonymous and because our school would add to the state profile, we didn't want to exclude anybody.” When interviewer observed that the codes in the administration schedule did not reflect full inclusion, the respondent suggested that a student might have been “at the nurse or something” and stressed that making inclusion and accommodation decisions for the NAEP was easy because everyone was included. He explained, “I remember there were some kids that just didn't come in. There were circumstances -- they were absent, late, in the nurse, maybe. We tested them [even if the student's English level was 1]. We were really making it easier for us.”

Table E23. Inclusion Criteria Used-District 4

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Teacher opinion	Unclear/ Do not know
Tully	4	x			x		
Tacitus	4						
Livy	4						
Homer	4	x					
Smithers	8	x			x		
Ovid	8	x		x		x	
Josephus	8						
Hesiod	8	x				x	
<b>Summary</b>		<b>5</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>0</b>

Dimension 5: Accommodations criteria used. In District 4, all accommodations discussed were given to students participating in the Math NAEP. The Ovid decision maker allowed one student with advanced English language proficiency to have extended time because the respondent “knew that the student would work better [have improved academic performance] without time pressure.”

As shown in Table E24, five schools assigned accommodations only to students who had an IEP (e.g., Livy and Josephus). The Livy decision maker observed that the school worked very hard to include everybody and that, “really, only special education students were offered accommodations” on NAEP. A Josephus student at an intermediate level of English proficiency was provided with testing in small group during the NAEP in mathematics because that student's IEP had indicated that this accommodation should be used. The Tacitus decision maker indicated that reading directions in Spanish is the only accommodation offered for state tests, but that students with IEPs were generally included with accommodations. At Homer, no students were discussed who participated in the NAEP with accommodations, but indicated that “only students with special needs/IEP are offered accommodations in state assessment.” At Hesiod, all students were tested without accommodations. The Hesiod decision maker went to explain that “all students are tested in English; only students with IEPs have accommodations in state assessment.” She continued, “The accommodation we can provide is translating directions, but we can't always do that.” “They may also have a word translator list that they use on a regular basis in their classes....The teacher and bilingual coordinator determine [if child may have these accommodations].” It was unclear what criteria decision makers at Smithers used when assigning accommodations.

Tully decision makers offered no accommodations, explaining, “because they [the ELLs at his school] are sufficiently proficient [in English], they do not need accommodations. They have scaffolding in place.”

Table E24. Accommodations Criteria Used-District 4

School Name	Grade	English language proficiency	Academic performance	Time in US	Amount of instruction in English	Offered no accommodations	IEP	Unclear/ Do not know
Tully	4					x		
Tacitus	4						x	
Livy	4						x	
Homer	4						x	
Smithers	8							x
Ovid	8		x					
Josephus	8						x	
Hesiod	8						x	
<b>Summary</b>		<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>5</b>	<b>1</b>

### *Tully*

Dimension 1: School personnel involved. The respondent was an assistant principal who emphasized that she did not make inclusion decisions for NAEP but that the school was simply given a roster of randomly selected students by the NAEP field staff representative. Teachers filled out the LEP Questionnaire, made initial judgments, and then submitted the questionnaire to the assistant principal. The assistance principal then met with the NAEP field staff representative, who made the final decision. The respondent reported that the process made it very easy to include children on the NAEP list. Only students new to the school were excluded.

Dimension 2: Involvement of NAEP field staff. The respondent, an assistant principal, is responsible for the bilingual students at the school (which is 50% of the school). She monitors their testing and follows their progress. She indicated that the NAEP representative was the primary decision maker for inclusion and accommodation and remarked ““We did not do anything out of the ordinary for LEP students. We just administered the exam.” The assistant principal explained, “We didn’t have any choice in the matter. Our NAEP coordinator [NAEP field staff representative] for our school gave us the names. Tully had nothing to do with this. The information came from them [NAEP].”

Dimension 3: NAEP tools used. The LEP Questionnaire was completed by the teachers of the sampled students and either given directly to the NAEP field staff or to the School NAEP Coordinator. “The teachers fill that [questionnaire] out and you turned that in. I don’t have it. They turn it into me as the coordinator and I give it to the ‘lady’ [the NAEP field staff representative]. Based on what the teachers filled out on those questionnaires. The NAEP decision tree and the NAEP inclusion criteria document were not used at all by decision makers.

Dimension 4: Inclusion criteria used. The respondent was an assistant principal who emphasized that he did not make inclusion decisions for NAEP but that the school was simply given a roster of randomly selected students by the NAEP field staff representative. Teachers filled out the LEP Questionnaire, made initial judgments, and then submitted the questionnaire to

the assistant principal. Teachers' initial judgments seemed to have been based on CELDT results and the student's progress vis-à-vis state ELD standards. The assistant principal explained, "Our school district has various assessments to determine a child's progress in language acquisition. There are all kinds of things that we use." The school coordinator then met with the NAEP field staff representative, who made the final decision. The respondent reported that the process made it very easy to include children on the NAEP list. Only students new to the school were excluded.

Dimension 5: Accommodations criteria used. No accommodations were offered to students included in the NAEP. "Instructions in Spanish are literally the only accommodation you could make and these children just don't need that at all." The respondent indicated that Grade 4 students do not need accommodations in Reading because they had been in LA Unified for several years. "Because they are sufficiently proficient [in English], they do not need accommodations. They have scaffolding in place."

### *Tacitus*

Dimension 1: School personnel involved. The respondents, an ELL Coordinator and an instructional/Title I coordinator, indicated that school personnel "did not make any decisions here."

Dimension 2: Involvement of NAEP field staff. The respondent explained that the NAEP field staff representative was the primary decision maker and maintained that "We tested all students who were on the NAEP list unless they were absent on the day of testing." Tacitus respondents explained that they had met with the respondents prior to testing and went over the packet of NAEP materials. "Testing went smoothly" since "everything was "precoded," and that they "went with whatever NAEP [field staff] recommended."

Dimension 3: NAEP tools used. At every school, the LEP Questionnaire was completed by the teachers of the sampled students and either given directly to the NAEP field staff or to the School NAEP Coordinator. The extent to which this document was used by decision makers, however, varied. The Tacitus respondents felt the LEP Questionnaire to be "a lot of work" but also commented that this tool was "a good thing to have to show that the school exercised the rights of the students; in other words, it was good for our documentation." Respondents at Tacitus did not address the use of the decision tree by decision makers. Respondents at Tacitus did not indicate the extent to which the criteria document was used by decision makers. The Tacitus respondent commented that the decision makers "looked at" the criteria document and "followed" it "but tested all."

Dimension 4: Inclusion criteria used. It appears that the Tacitus respondent thought the initial list of pre-selected students for the NAEP was the final list of students to be selected. The Tacitus respondent commented that the decision makers "looked at" the criteria document and "followed" it "but tested all." The respondent explained that the NAEP field staff representative was the primary decision maker and maintained that "We tested all students who were on the NAEP list unless they were absent on the day of testing." The respondents, an ELL coordinator and an instructional/Title I coordinator, indicated that school personnel "did not make any decisions here."

Respondents received a list of randomly selected students and the school tested everybody. "When we got the list, [students] had already been selected. When we get the list, we say every child should be included. It is not up to us to say some can take the test and some cannot." All students on the NAEP list were tested unless absent on the day of testing. The

respondents worked together, making no changes to the list of selected students unless the student had an IEP.

The respondents noted several times that Tacitus made no decision to include or exclude students but “went with the list and looked at teacher answers on the LEP Questionnaire.” Interestingly, the Tacitus respondent commented, “We are broad in our inclusion and feel it is a good thing to have students participate in as much of this kind of assessment as possible. . . . Overall, we went with whatever NAEP recommended. Teachers received those [LEP] questionnaires and filled them out—they are a good thing to have to show that we exercised the rights of the students—good for our documentation. Because we are broad in our inclusion and work hard to have our students participate, our students’ English level is higher—we are a community that works! We encourage full participation.”

The respondent explained that the NAEP field staff representative was the primary decision maker and maintained that “We tested all students who were on the NAEP list unless they were absent on the day of testing.” Apparently, this inclusiveness was possible because all the students on the list were of intermediate or native proficiency. Students new to the program or who have been in the program for less than a year are not tested.

Dimension 5: Accommodations criteria used. The respondent indicated that reading directions in Spanish is the only accommodation offered for state tests and mentioned several times that it was the school’s practice to include all students. On state tests, the only accommodation is that directions to the test can be read in Spanish if needed.” Students with IEPs were generally included with accommodations. “Special Ed students would get accommodations anyway.”

### *Livy*

Dimension 1: School personnel involved. The respondent was a programs coordinator responsible for Title I and English language compliance and testing. He was new to the job and stayed less than 1 year. (He had left his position at this school by the time he was contacted, but was in the process of training someone at the school to take over for him.)

Dimension 2: Involvement of NAEP field staff. The programs coordinator also consulted with classroom teachers, who provided information based on close observation of students in class. Teacher filled out the NAEP Questionnaire, but the questionnaire “was not a big factor” in the decision. “I think we tried to do it before [the NAEP field staff visit]. Yes, we got [guidance from [NAEP field staff personnel]]. They did come in. I had a lot of contact over the phone. They were supportive. They made it easier for me.”

Dimension 3: NAEP tools used. Even though Livy decision makers had had teacher complete the LEP Questionnaire, they did not use it, nor did they use the decision tree or NAEP criteria. The Livy respondent said that they used the state testing guidelines for STAR.

Dimension 4: Inclusion criteria used. The respondent looked at the NAEP test as a good opportunity to give students assessment practice so that they could become more comfortable with testing: “Because from what I understood, we followed the STAR [California’s Standardized Testing and Reporting program] guidelines. It’s a little more practical I thought. It’s a little harder for us to have different criteria [for each large-scale assessment]” This program requires use of the SAT9. “I understood that when students were taking the test, it was supposed to be anonymous, so I would not have much information about the student...I don’t have enough information to tell you specifically [about English language proficiency.] We wanted to include everybody. That was our criteria and we did as much as possible -- unless there was a request not to from the teacher or the parent....We kind of followed the STAR testing criteria because they

include everybody. Because it is anonymous and because our school would add to the state profile, we didn't want to exclude anybody." He continued, "We include students in the regular STAR test, but we give them another test, the FABLE, in their native language as well as CELDT and portfolio that moves students thorough English language development level. Each student should be progressing one English language level a year (requirement from the state). We given them a test (DPI—diagnostic placement inventory—to find out what English language level they should be in middle school."

When interviewer observed that the codes in the administration schedule did not reflect full inclusion, the respondent suggested that a student might have been "at the nurse or something" and stressed that making inclusion and accommodation decisions for the NAEP was easy because everyone was included. He explained, "I remember there were some kids that just didn't come in. There were circumstances -- they were absent, late, in the nurse, maybe. We tested them [even if the student's Eng level was 1]. We were really making it easier for us."

Dimension 5: Accommodations criteria used. The respondent indicated that he used STAR guidelines as a means of determining whether to exclude or include a student with or without an accommodation. The respondent indicated that STAR criteria for LEP students showed which accommodations were available, especially for language arts in Grades 4–6. The respondent explained, "For STAR, LEP students have some accommodations that were available if we decided to provide them. STAR is testing language arts. [NAEP] also provided us with a sort of practice for STAR because it doesn't hold us to any accountability, so we used it for practice." However, no students discussed in the interview were provided with accommodations. The respondent observed that the school worked very hard to include everybody and that, "really, only special education students were offered accommodations" on NAEP. He also noted, "'It was my first year, I was still learning. I was still learning the process of accommodation. In reality, if it was there, I didn't even have time to look it up."

### *Homer*

Dimension 1: School personnel involved. According to the respondents, a Title I coordinator and a bilingual coordinator, NAEP (NAEP field staff) provided the Title I coordinator a list of randomly selected students designated to take the test. The Homer respondents explained that teachers filled out LEP Questionnaires, but that they had no real role in deciding which students were included or excluded because decisions were based solely on ELD levels. Once the Title I coordinator received the list, he or she automatically eliminated (excluded) students with English Language Levels of 1 or 2.

Dimension 2: Involvement of NAEP field staff. Homer decision makers used no NAEP field staff input when making their decision. According to the respondents, a Title I coordinator and a bilingual coordinator, NAEP (NAEP field staff) provided the Title I coordinator a list of randomly selected students designated to take the test. Once the Title I coordinator received the list, he or she automatically eliminated (excluded) students with English Language Levels of 1 or 2.

Dimension 3: NAEP tools used. The Homer respondents explained that teachers filled out LEP Questionnaires, but that they had no real role in deciding which students were included or excluded because decisions were based solely on ELD levels. In effect, the LEP Questionnaire, the NAEP decision tree, and the NAEP inclusion criteria document were not used. "We used CELDT. That is what identifies them as LEP and establishes their level of proficiency....The

teachers had no part in this [completing the LEP Questionnaire] and we were asked only for their ELD levels. Only [student's] LEP score was used."

Dimension 4: Inclusion criteria used. At Homer, decisions were based solely on ELD levels. "I was told by NAEP that certain students would be selected at random. From that list, I looked at their English Language Levels and we eliminated those with Levels 1 and 2. English language levels are established by the CELDT test, which is given at the beginning of every year." English language development levels are determined by student performance on CELDT (California English Language Development Test). The Home Language Survey response is what triggers the CELDT testing."

Dimension 5: Accommodations criteria used. No students were discussed who participated in the NAEP with accommodations. "They [NAEP field staff] came to the school. They had their list of students to be pulled out. I asked them if they needed help. They said no. They did it all themselves. I never got any feedback on the questionnaires." The respondent indicated that only students with special needs/IEP are offered accommodations in state assessment.

### *Smithers*

Dimension 1: School personnel involved. The test coordinator was the main school-based decision maker involved. The respondent was not present when NAEP field staff visited the school to conference with and train teachers and did not consider herself to be a part of the decision to include an ELL, but simply adhered to the list provided by NAEP. Indeed, the respondent reported being responsible only for "getting students to the cafeteria for testing."

Dimension 2: Involvement of NAEP field staff. The respondent, a testing coordinator and counselor at the school, reported that, based on the list of pre-selected students provided by the NAEP field staff representative, teachers completed LEP Questionnaires for those selected students. From there, the NAEP field staff representative used these questionnaires to make decisions regarding the inclusion and accommodation of students.

Dimension 3: NAEP tools used. Respondents at Smithers only referred to the use of the LEP Questionnaire, but not the decision tree nor the NAEP criteria document.

Dimension 4: Inclusion criteria used. The respondent, a test coordinator, said that the NAEP field staff representative used information on the LEP Questionnaires to make decisions regarding the inclusion or exclusion. The teachers completed LEP Questionnaires for the selected students (The respondent mentioned that students were "placed in classes based on CELDT test" which suggests that, indirectly, ELD levels and program placement were criteria in the decision-making process.) The NAEP field staff representative used the LEP Questionnaires to make decisions regarding the inclusion and accommodation of students in the NAEP.

Dimension 5: Accommodations criteria used. No students discussed in the interview were given accommodations on the NAEP. The respondent was unable to answer questions regarding accommodations.

### *Ovid*

Dimension 1: School personnel involved. The respondents, a bilingual coordinator and an ELD coordinator, worked with teachers on the "list" (NAEP administration schedules) to include as many students as possible.



Dimension 2: Involvement of NAEP field staff. The respondents reported that NAEP field staff personnel provided guidance and administered the test.

Dimension 3: NAEP tools used. Respondents at Ovid did not indicate the extent to which the LEP Questionnaire, decision tree, or NAEP inclusion criteria document were used by decision makers.

Dimension 4: Inclusion criteria used. The respondents, a bilingual coordinator and an ELD coordinator, worked with teachers on the list provided by NAEP “to include as many students as possible.” The Ovid respondent indicated that the primary criteria for determining participation in the NAEP were ESL level, time in country, and classroom teacher input. Information relevant to the decision to include or exclude students was drawn from the following assessments: CELDT, LAS Espanol, CAT6, and SABE. Students who were included were at “high intermediate” levels of English proficiency. No students were excluded from the Reading NAEP.

An Ovid student was discussed who had been excluded from the Math NAEP because of the student’s language level and time in country. When asked about the decision to include a student in the Reading NAEP without an accommodation, the Ovid respondent indicated that the student was at a “high intermediate or advanced” level of ELD.

Dimension 5: Accommodations criteria used. When asked about the decision to include a student in the Reading NAEP without an accommodation, the respondents indicated that the student was at a “high intermediate or advanced” level of ELD. The respondents indicated that a student participating in the Math NAEP was offered the extended time accommodation because the respondent “knew that the student would work better without time pressure.”

### *Josephus*

Dimension 1: School personnel involved. A coordinator of testing (who is a year-round employee who not in the classroom), testing coordinator/bilingual clerk, and teachers.

Dimension 2: Involvement of NAEP field staff. The respondent found the process easy because NAEP provided the respondent with a list of pre-selected students. Also, NAEP field staff personnel came 2 to 3 times per week and worked with “the testing crew” to organize decision-making and test administration: “The NAEP crew was excellent here—they did everything. They gave us a punchlist; everything really easy.” NAEP provided the respondent with a list of 160-170 students selected to participate in the test. The respondent observed that this was the largest group ever tested and that, after testing, NAEP took all materials. Hence, the respondent had “no recollection from information available about the students” included or excluded from the test.

Dimension 3: NAEP tools used. The Josephus respondent simply indicated that teachers “filled out LEP Questionnaire, turned into NAEP, we never saw them after that.” According to the respondent, decisions were made by NAEP field staff personnel based on information from the LEP Questionnaires, which were completed by the students’ teachers. In completing the questionnaire, teachers made reference to classroom observations and their knowledge of students’ testing level. The NAEP Decision Tree and Criteria for Inclusion were not used by Josephus staff.

Dimension 4: Inclusion criteria used. The respondent indicated that there were two primary criteria used for determining whether or not and how an LEP student was to be included in the NAEP: ELD level (determined by CELDT and bilingual office) and teacher recommendation (based on observation, testing level, and ability). The LEP students included in

math and reading NAEP had IEPs, so this was a factor in decisions regarding these students as well.

According to the respondent, decisions were made by NAEP field staff personnel based on information from the LEP Questionnaires, which were completed by the students' teachers. In completing the questionnaire, teachers made reference to classroom observations and their knowledge of students' testing level. The Josephus respondent simply indicated that teachers "filled out LEP Questionnaire, turned into NAEP, we never saw them after that." The respondent found the process easy because NAEP provided the respondent with a list of pre-selected students.

Dimension 5: Accommodations criteria used. A Josephus student at an intermediate level of English proficiency was provided with testing in small group during the NAEP in mathematics because that student's IEP had indicated that this accommodation should be used.

### *Hesiod*

Dimension 1: School personnel involved. The respondent was an assistant principal.

Dimension 2: Involvement of NAEP field staff. NAEP field staff "came prior to the testing to answer any questions we might have, to determine testing locations. And then she came a second time to make sure we were on task." The respondent indicated that decisions to include students in the NAEP were made primarily with reference to a "teacher survey," i.e., the LEP Questionnaire, which conveyed the recommendations of selected students' teachers regarding participation in the NAEP. "A list came to me and then I used the teacher questionnaire. We did not consult with anyone (to include/exclude)...[NAEP] is fairly easy to implement [since] NAEP sends their own proctors. We don't have too much information on what to ask and how it happens. LEP Questionnaires were for teachers. They were instructed by NAEP to fill them out...Part of it is that there isn't any [accountability]. It's a general report of the country. We don't think about it. We're not concerned about it."

Dimension 3: NAEP tools used. Decisions regarding the inclusion of students in the NAEP were made based on the LEP Questionnaire completed by teachers. According to the respondent, teachers' decisions "would be easy because teacher [knows the] language level" of the student. The Hesiod respondent indicated that the LEP Questionnaire was used "a lot" to determine who to include. The classroom teacher completed the LEP questionnaire based on classroom observations and language assessments scores. The NAEP decision tree was not used at all by decision makers at Hesiod. The Hesiod respondent was unsure whether or not the teachers received the NAEP criteria. "I don't recall placing it [using it or giving it to the teachers]" she explained.

Dimension 4: Inclusion criteria used. Decisions regarding the inclusion of students in the NAEP were made based on the LEP Questionnaire completed by teachers. According to the respondent, teachers' decisions "would be easy because the teacher [knows] the language level" of the student. Teacher recommendation was based on classroom observation and language assessment scores. Teachers know each student's language level because this information is printed in roll books. The respondent speculated that teachers' decisions were probably fairly easy to make because of their knowledge of the students' ELD levels. The respondent also observed that "all students are tested in English; only students with IEPs have accommodations in state assessment." "My role is to be sure that each student is tested."

The respondent did not know the reason that an LEP student at an advanced level of English language proficiency was excluded from the NAEP. Students excluded from math:

“Proficiency is beginning. Language proficiency is what exclusion is based on, not time in country.” “In California, we cannot offer instruction in home language. We can use it to clarify and assist.”

Dimension 5: Accommodations criteria used. Students were tested without accommodations. “Accommodations were done through a survey given to teachers. The ones who were not excluded were preparing to reevaluate and were therefore at a higher level where accommodations were not necessary.....My guess is that when teacher felt the [student’s] level was too low, they were just excluded.” With regard to a student tested in Reading, the respondent indicated that the decision to test with or without an accommodation was based on the student’s “level,” which may refer to ELD level.

The respondent also observed that “all students are tested in English; only students with IEPs have accommodations in state assessment.” She continued, “The accommodation we can provide is translating directions, but we can’t always do that.” “They may also have a word translator list that they use on a regular basis in their classes....The teacher and bilingual coordinator determine [if child may have these accommodations].”