



SECTION 6: ENGLISH TESTING OPTIONS

6.1 INTRODUCTION

The law requires applicants to have an understanding of the English language, demonstrated by the ability to read, write, and speak words in ordinary usage. In this section, we discuss options for the level of English proficiency required to meet the dictates of the law. The remainder of this section is organized into the following subsections:

- Considerations
- Proficiency Level: Options
- Stakeholder Input

6.2 CONSIDERATIONS

6.2.1 What does the law require. We have shaped this discussion of English testing options around the language provided in the law. With regard to English, the law requires that applicants demonstrate:

"an understanding of the English language, including an ability to read, write, and speak words in ordinary usage in the English language..."

The law further states that:

"the requirements of this paragraph ... shall be met if the applicant can read or write simple words and phrases to the end that a reasonable test of his literacy shall be made...."

The law requires proficiency in reading, writing, and speaking. INS must determine proficiency levels for each of these skills. Options for each are provided in Section 6.3.

6.2.2. What does the naturalization process require. It is possible that the English proficiency necessary to negotiate the naturalization process is higher than the proficiency requirements of the law. Consultation group members, congressional staff, adjudicators, and CBO representatives all raised the issue of the disparity between the law's requirement to understand language in "ordinary usage" and the language required to complete the N-400, answer questions at interview, answer questions during history and government testing, and take the oath.

6.2.3 What do existing resources provide. In the remainder of this section, we discuss several proficiency guidelines that INS can use to inform its decisions regarding the appropriate level of English proficiency on which to test naturalization applicants. Once a specific proficiency level is selected, any of these existing guidelines can serve as a good starting point for INS's test developer. However, because the included guidelines were



like what?

not developed specifically for citizenship testing some modifications may be required. Additionally, most proficiency guidelines proceed through a spectrum of proficiency levels, the majority of which are too high to be consistent with the requirements of the law. Although we have included the full spectrum of proficiency guidelines in many cases, ESL experts have recommended that INS choose a proficiency level at the lower end of the spectrum for the purpose of testing naturalization applicants.

6.3 PROFICIENCY LEVEL: OPTIONS

The English Consultation Group recommended several existing proficiency guidelines. Although there are many different English proficiency guidelines for reading, writing, and speaking, the ones we have included in this section (1) are credible, (2) were created using a valid process, and (3) were specifically recommended by the English, ESL, and literacy experts we convened in our consultation group. They include:

- English-as-a-Second-Language, Model Standards for Adult Education Programs; 2
- Canadian Language Benchmark (CLB); 3
- Foreign Service Institute (FSI) / Interagency Language Roundtable (ILR) Proficiency Guidelines; 1
- Performance-Based Curricula and Outcomes—The Mainstream English Language Training Project (MELT) (*Speaking only*);
- Adult Rating of Oral English (AROE) (*Speaking only*); and 7
- Measure Definitions for the National Reporting System for Adult Education. 4

Is this #5?

Confusing order

On the following pages, we compare these guidelines with two outcomes in mind:

- INS can use the guidelines described to help determine at what level naturalization applicants should read, write, and speak English, based upon the requirements in the law; and
- Once INS has determined at what proficiency level to test naturalization applicants, it can use one or more of the proficiency guidelines as a starting point for its test developer.

Although each set of guidelines uses different scales to measure proficiency, we have divided the levels of proficiency into three levels: low, medium, and high. We believe this will assist INS in making a decision regarding the reading, writing, and speaking proficiency level at which it will test naturalization applicants. We have also indicated on each matrix the proficiency levels suggested by the Consultation Group and the public, based on their understanding of the law. The shaded areas on each matrix represent the average range of proficiency suggested by the consultation group and the public, respectively. Once INS has made a decision regarding general proficiency levels, it may choose a more specific level, or levels, at which to test based upon the standard it chooses to use.

what suggested by court strategy advisors?

think consult - gov lower than public?

Judith: public is less informed

any cuts on that 36

A COMPARISON OF EXISTING ENGLISH PROFIC

Low Proficiency

CA ESL
CA ESL Model
Standards)

ESL Beginning Literacy Unable to read English	ESL Beginning (Low) Recognizes letters and numbers	ESL Beginning (High) Gets limited meaning from print with successive rereading and checking
---	--	---

CLB
Lian Language
nchmarks)

Stage 1, Benchmark 1 Initial competence at reading simple texts	Stage 1, Ben Developing competence
---	--

Note: Selected competencies for Stage 1 include simple instructional formatted,

NRS
tional Reporting
System)

Beginning Literacy - No or minimal skills such as the ability to read own name or simple isolated words - Little or no comprehension of how print corresponds to spoken language	Beginning ESL Reads numbers and letters but has limited understanding of print only through frequent rereading
--	---

ACTFL
merican Council of
achers of Foreign
Languages)

Novice (Low) Occasionally identifies isolated words and/or major phrases when strongly supported by context	Novice (Mid) - Identifies an increasing number of highly contextualized words and/or phrases - Material understood rarely exceeds a single phrase at a time and rereading may be required
--	---

FSI/ILR
gn Service Institute
ar / Language
Recordtable)

Novice—Prefunctional Proficiency Reads signs and menu items	- Reads material with - May get only the mai - Reads simple texts a - things: basic direction
--	--

Consultation Group Recommendation

ESL Intermediate (Low)

- Reads simplified material on familiar subjects and some authentic material dealing with everyday matters

Benchmark 2

at reading simple texts

Stage 1, Benchmark 3

- Adequate competence at reading simple texts

unformatted, and informational texts

Low Intermediate ESL

- Reads simple material on familiar subjects and comprehends simple and compound sentences
- Difficulty with authentic materials

Novice (High)

- Interprets written language in areas of practical need
- Reads learned messages, phrases, or expressions
- Occasionally derives meaning from higher level material where context is supportive

Intermediate—Survival Proficiency

common words and simple grammatical constructions. Ideas and will probably have to reread a lot of material about predictable subjects: short descriptions of persons, places, and things and descriptions of what a place or person looks like.

A COMPARISON OF EXISTING ENGLISH PROFICIENCY SCALE

Medium Proficiency

CA ESL

Public Focus Groups Recommendation	
ESL Intermediate (Low) (Continued)	ESL Intermediate (High) - Reads materials on familiar subjects - Reads authentic materials with limited success

CLB

Stage 1, Benchmark 4 Fluent competence at reading simple texts	Stage 2, Benchmark 5 Initial competence at reading texts of medium complexity	Stage 2, Benchmark 6 Developing competence at reading texts of medium complexity
---	--	---

NRS

Note: Selected competencies for Stage 2 include instructional formatted uniform	
Low Intermediate ESL (Continued)	High Intermediate ESL - Reads simple texts on familiar subjects but has difficulty with more complex - Interprets actions required in written directions - Reviews a paragraph and identifies errors

ACTFL

Intermediate (Low) - Understands the main ideas of most speech in a standard dialect - May not be able to comprehend complex extended discourse - Emerging awareness of culturally implied meanings but may fail to grasp sociocultural nuances	Intermediate (High) - Reads consistently with understanding of basic and social needs - Examples may include: local news, travel guides, and popular fiction for a wide audience
--	--

FSI/ILR

Advanced—Limited Working Proficiency - Understands at least the main ideas and some of the details in clearly presented prose with familiar sentence patterns and basic vocabulary - Abstract topics and complicated grammatical structures may be beyond the reader - May be able to read health or agricultural manuals, simple story, local newspapers, oral literature, and short stories	
--	--

CIENCY GUIDE LINES (Reading)

y

	ESL Advanced (Low) • Reads authentic materials on everyday subjects and nontechnical prose • Reads technical materials with difficulty
--	---

Benchmark 6 ce at reading texts of	Stage 2, Benchmark 7 • Adequate competence at reading texts of medium complexity
---	--

atted, and analytical	
aterials	

Intermediate (Mid) an increased understanding simple connected texts dealing with a variety of ds short straightforward descriptions of persons, places and things written	

vocabulary	

A COMPARISON OF EXISTING ENGLISH

High Proficiency

CA ESL

Advanced (High) • Reads authentic materials on familiar subjects and nontechnical prose

CLB

Stage 2, Benchmark 8 • Fluent competence at reading texts of medium complexity	Stage 3, Benchmark 9 • Initial competence at reading complex texts	Stage 3, Benchmark 10 • Developing competence at reading complex texts
--	--	--

Note: Selected competencies for Stage 3 include complex instructional, formatted

NRS

Advanced ESL • Reads simple narratives on familiar subjects • Makes some minimal inferences about familiar texts and compares and contrasts information from such texts, but not consistently	• Reads authentic materials • Interprets narratives on
--	---

ACTFL

Intermediate (High) • Reads consistently with full understanding simple connected texts dealing with basic personal and social needs • Gets some main ideas from texts at the next higher level featuring description and narration • Has some difficulty with the cohesive factors in discourse • Comprehension is less consistent than at the advanced level • May have to read material several times for understanding	• Reads • Gets • Com • Incl • busi
--	--

FSI/ILR

No Equivalent
Proficiency Levels

ESL PROFICIENCY GUIDELINES (Reading)

Proficiency

	Program Exit • Responds to questions on forms and applications and writes simple discourse of several paragraphs
--	--

Mark 10 Competence at reading	Stage 3, Benchmark 11 • Adequate competence at reading complex texts	Stage 3, Benchmark 12 • Fluent competence at reading complex texts
--	--	--

...l, unformatted, and analytical texts.

Proficient ESL ...als on everyday subjects ...familiar topics and gains meaning from unfamiliar topics using complex language	
--	--

Advanced ...ds longer prose of several paragraphs in length with familiar sentence patterns ...s the main ideas and facts and misses some details ...prehension derives from situational and subject-matter knowledge and increasing control of the language ...ides descriptions and narrations such as simple short stories, news items, bibliographical information, personal and ...ness correspondence, and simple technical material	
---	--

A COMPARISON OF EXISTING ENG

Low

CA ESL

CA ESL Model Standards)

ESL Beginning Literacy Unable to write any English other than own name and address	ESL Beginning (Low) May be able to write own name and address	ESL Beginning (High) Copies words and phrases previously learned
---	--	---

CLB

Canadian Language Benchmarks)

Stage 1, Benchmark 1 Initial competence in writing simple texts
--

Note: Selected competencies for Stage 1 include

NRS

National Reporting System)

Beginning Literacy - No or minimal writing skills - May write letters or numbers and copy simple words - No or incomplete recognition of the alphabet - May have difficulty using a writing instrument - Little or no comprehension of how print corresponds to spoken language	Beginning (Low) - Prints numbers, letters, and sight words and - May write simple sentences or phrases (e.g., and letters - Narrative writing is disorganized and unclear
--	--

ACTFL

American Council of Teachers of Foreign Languages)

Novice (Low) - Forms some letters in an alphabetic system - Copies and produces the basic strokes	Novice (Mid) - Copies or transcribes familiar words or phrases and reproduces some from memory - No practical communicative writing skills	Novice (High) - Writes simple fixed expressions - Supplies information - Writes names, numbers, simple lists
---	--	---

ISI/LR

Interagency Language Roundtable)

Novice - Prefunctional Proficiency Mostly practicing the writing system rather than using it for functional purposes	Intermediate - Survival Proficiency - Meets limited practical writing needs by recombining structures into simple sentences - Unlikely to make a lot of spelling and grammatical
---	--

ESL PROFICIENCY GUIDELINES (Writing)

Proficiency

		Consultation Group Recommendation
Beginning (High) Copies and writes sentences based on materials		ESL Intermediate (Low) Writes short messages and notes within the scope of their limited language experience
Stage 1, Benchmark 2 Developing competence in writing simple texts	Stage 1, Benchmark 3 Adequate competence in writing simple texts	Stage 1, Benchmark 4 Fluent competence in writing simple texts
copying information, filling out forms, describing personal situations, and expressing simple ideas		
Beginning ESL Uses lists of familiar words and phrases (address and phone number) and simple messages. Frequent errors		Low Intermediate ESL Writes simple notes and messages on familiar situations using simple and compound sentences. Often lacks complete clarity and focus. Sentence structure is simple, but shows some control of basic structure and mechanics
Novice (High) Expressions and limited memorized material in simple forms and documents (names, dates, own nationality, other simple autobiographical information) and some short phrases and		Intermediate (Low) Meets limited practical writing needs. Writes short messages, postcards, and simple notes. Creates limited statements/questions. Makes frequent errors, but writing can be understood
Proficiency Expanding vocabulary and Mistakes		

A COMPARISON OF EXISTING ENGLISH

Medium Proficiency

Public Focus Groups Recommendation

ESL Intermediate (High)

- Can perform basic writing tasks in similar contexts

Stage 2, Benchmark 5

- Initial competence in writing text of medium complexity

Stage 2, Benchmark 6

- Developing competence in writing texts of medium complexity

Stage 2, Benchmark 7

- Adequate competence in writing texts of medium capacity

Note: Selected competencies for Stage 2 include reproducing information, filling out forms, conveying formal messages, and expressing opinions.

Low Intermediate ESL

(Continued)

High Intermediate ESL

- Can write simple paragraphs on familiar topics
- Can write messages and notes related to basic needs and short paragraphs
- Basic spelling/punctuation with some errors

Intermediate (Low)

(Continued)

Intermediate (High)

- Meets a number of practical writing needs and content involves personal preferences, daily routine, personal experience
- Writing tends to be a loose collection of sentences with little conscious organization
- Can generally be understood by natives

Advanced-Limited Working Proficiency

- Meets most practical writing needs in work or study situations
- Uses a fairly good range of vocabulary and structures
- Still makes some grammatical errors and words will not be idiomatic

CA ESL

CLB

NRS

ACTFL

ILR

PROFICIENCY GUIDELINES (Writing)

Proficiency

	ESL Advanced (Low) Can produce routine correspondence and paragraphs about previously discussed topics	
Benchmark 7 Can write texts of	Stage 2, Benchmark 8 Fluent competence in writing texts of medium complexity	Stage 3, Benchmark 9 Initial competence in writing complex texts
g complex ideas.		Note: Selected competencies for Stage 3 include writing instructional, formatted and unformatted texts; summarizing, analyzing, evaluating, and persuading.
Can write about daily activities	Advanced ESL - Can write simple narrative descriptions and short essays on familiar topics - Can handle basic writing mechanics with few errors but makes errors with complex structures	
Intermediate (Mid) Writes short, simple letters about routine everyday events, and other topics grounded in concrete experiences or sentence fragments on a given topic and lacks		Intermediate (High) - Meets a number of practical writing needs and limited social demands - Takes notes in detail on familiar topics - Writes simple letters, brief synopses, and paraphrases - Emerging ability to describe and narrate in paragraphs - Writing is generally comprehensible

A COMPARISON OF EXISTING ENGL

High P

CA ESL

ESL Advanced (High) • Can produce descriptions, essays, and summaries	Program Exit • Can respond to questions on forms and write simple discourse of several parag
--	---

CLB

Stage 3, Benchmark 10 • Developing competence in writing complex texts	Stage 3, Benchmark 11 • Adequate competence in writing complex texts	• Flu - co
---	---	---

NRS

Advanced ESL (Continued)	Proficient ESL • Writes multiparagraph essays • Writes well formed sentences and appropriate mechanics • Makes few errors
--	--

ACTFL

Intermediate (High) (Continued)	Advanced • Writes routine social correspondence and joins sentences in simple discourse • Takes notes and writes cohesive summaries, resumes, and narratives of a factual nature • May still make errors in punctuation and spelling • Good control of syntax, but makes frequent errors in producing complex sentences • Writing may resemble literal translations from the native language, but a sense of organization is emerging • Writing is understandable
---	---

FSI/ILR

No Equivalent
Proficiency Levels

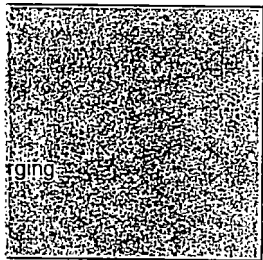
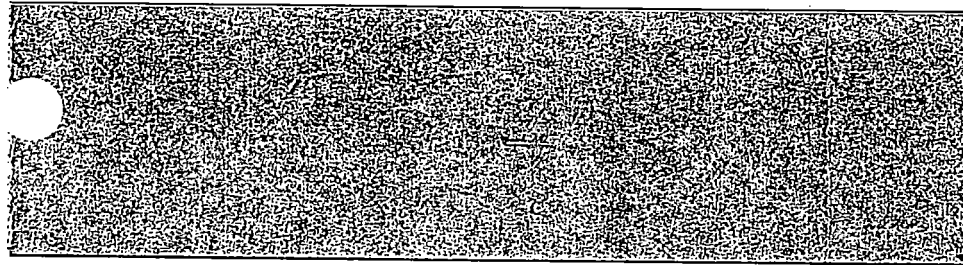
ISH PROFICIENCY GUIDE . NES (Writing)

Proficiency

applications and
graphs

Stage 3, Benchmark 12

ent competence in writing
mplex texts



NATURALIZATION TESTING: CONTENT OPTIONS



Below, we provide a brief description and summary of the development methodology for each of the guidelines described in the English proficiency matrices. Many of the guidelines apply to all three areas of the English language: reading, writing, and speaking. Two apply only to speaking and we have marked these guidelines with an asterisks:

- English-as-a-Second-Language, Model Standards for Adult Education Programs.** In 1988, an Adult Education Advisory Committee was established in California and developed a report entitled *Adult Education for the 21st Century: Strategic Plan to Meet California's Long-term Adult Education Needs*. The recommendations in this strategic plan provided standards and performance measures for programs to follow while developing adult education programs, as well as standards by which to measure their success. From this plan, the *English-as-a-Second-Language (ESL) Model Standards for Adult Education Programs* were created. The model standards were developed by a field-based ESL curriculum committee, which served in an advisory capacity to the Adult Education Unit of the California Department of Education. All standards set forth in *The ESL Model Standards for Adult Education Programs* were validated through a field review process. *The ESL Model Standards for Adult Education Programs* describe general standards, define language proficiency levels, and identify possible tests for measuring language proficiency. The purpose of these standards are to ensure that adults participating in ESL programs acquire language and cultural proficiencies that will enable them to fulfill personal, vocational, academic, and civic goals.²³
- Canadian Language Benchmark (CLB).** The CLB is a task-based guideline that uses a four staged process with 12 benchmarks. The CLB provides oral/aural, reading and writing standards. The benchmarks provide a description of the ability to use the English language in specifically defined performance and situation conditions. The standards are hierarchical, building upon skills acquired when the individual met the last benchmark. Proficiency descriptions, and current theories of language (both linguistics and psychological) were considered during the formulation of the benchmarks. In addition, learners were identified to test content validity.²⁴
- Foreign Service Institute (FSI) / Interagency Language Roundtable (ILR) Proficiency Guidelines.** In 1952, the Secretary of State called upon Foreign Service Institute (FSI) to create criteria that could be used to identify the foreign language proficiency of U.S. Government employees. The Civil Service Commission established a committee comprised of representatives of government agencies concerned with foreign languages. This committee was tasked with developing definitions for key levels of proficiency and creating an inventory of employees with foreign language proficiency in the various agencies. These definitions were field tested and then revised. In 1982, the American Council of the Teaching of Foreign

²³ *English-as-a-Second-Language: Model Standards for Adult Education Programs*. California Department of Education, Sacramento, 1992.

²⁴ *Canadian Language Benchmarks: English as a Second Language for adults English as a second language for literacy learners*. Working Document, 1996.



Languages (ACTFL) published a modified version of the Government's Proficiency Scale to meet the needs of secondary schools and colleges. The revised guidelines were distributed for comment to foreign language teaching professionals and field-tested to determine if the new guidelines were appropriate. In addition, language-specific guidelines were developed and distributed for comment.²⁵

- *** Performance-Based Curricula and Outcomes—The Mainstream English Language Training Project (MELT) (*Speaking only*).** The Federal Office of Refugee Resettlement's (ORR) MELT initiative established standards for ORR-funded adult refugee English Language Training (ELT) programs. These standards were established using a phased approach:
 - Research;
 - Field testing;
 - Materials development; and
 - Dissemination and implementation of English language training standards.

ORR funded seven national ELT projects, and 19 refugee ELT programs participated as demonstration sites with 1,375 refugees participating in the field-testing. The outcomes of these projects were a competency-based core curriculum guide, a system for defining Student Performance Levels (SPLs), and a valid and reliable instrument for assessing students' English proficiency, the Basic English Skills Test (BEST). Although the ELT project initially developed SPLs for reading and writing proficiency as well as oral proficiency, these guidelines are now outdated. For this reason, we have included only the SPLs for speaking in this document.²⁶

- *** Adult Rating of Oral English (AROE) (*Speaking only*).** The Adult Rating of Oral English was developed by Development Associates, a Virginia-based education consulting firm that works with bilingual programs and limited English proficient students at the adult and elementary levels. Development Associates recognized the need to create a simple tool for rating oral English proficiency of limited English proficient adults entering training programs, specifically vocational training programs.

Development Associates conducted research to determine the value of the AROE. They conducted reliability assessments in relations to existing measurement scales and cross-validated it with other oral assessment tools. In addition, Development Associates conducted inter-rater reliability assessments and found moderately high reliability even among raters trained in ESL and raters not trained in ESL.²⁷

²⁵ ACTFL Proficiency Guidelines. American Council for Teachers of Foreign Languages, 1989.

²⁶ Grognet, Allene G. *Performance-Based Curricula and Outcomes: The Mainstream English Language Training Project (MELT) Updated for the 1990s and Beyond*. Center for Applied Linguistics, 1997.

²⁷ Interview with Paul Hopstock regarding the Adult Rating of Oral English (AROE), Development Associates, Inc., December 29, 1998.



- **Measure Definitions for the National Reporting System for Adult Education.** National Reporting System (NRS) is in the process of establishing a system for measuring and documenting student outcomes resulting from adult education instruction throughout the country. The success of this system is dependent upon establishing a common set of measure definitions as well as a uniform methodology for collecting data. The NRS has developed three types of measures along with definitions and major topic areas and is awaiting pilot testing by states and local programs. It is important to note that the overriding principles during the definition phase were the need to accommodate the diversity of the adult education delivery system and the need for compatibility of the definitions with related adult education and training programs.²⁸

6.4 STAKEHOLDER INPUT

As described in Section 3 “Methodology,” we spoke with a variety of stakeholders while developing this document, including subject matter experts, Congressional staff, CBOs, the public, and field staff. Below, we discuss the input of each set of stakeholders regarding the level of English proficiency new citizens should demonstrate.

6.4.1 Consultation Groups. The consultation group members believed that the language of the law was reasonable in terms of the proficiency required. They suggested that the law required a basic or “low proficiency” level. The consultation group provided the following examples and clarifications regarding what a basic proficiency level is:

- **Speaking.** Based upon the language of the law, the Consultation Group believed that applicants should be able to: (1) verbally provide basic personal information (e.g., information on residence, children, and family), (2) participate in basic daily interactions (e.g., shopping, simple social conversations), (3) use clarification strategies (e.g., asking “can you repeat that?”), and (4) express opinions.
- **Reading.** Based upon the language of the law, the consultation group believed that applicants should be able to read “environmental print” -- print necessary to function on a daily basis and preserve personal safety. Discussion of reading ability also focused on the functions of reading. Consultation group members suggested that the most important functions of reading were to obtain information, answer questions, and receive instructions.
- **Writing.** Because of the different ways in which many people learn language, often focusing first on oral ability, the consultation group believed that the level of English proficiency required in writing should be lower than that required in speaking. Based upon the language of the law, the consultation group believed applicants should be able to write short notes such as telephone messages. The consultation group also thought that much of the language related to history and government may be beyond the proficiency level described in the law (e.g., judiciary, legislature).

²⁸ *Measure Definitions for the National Reporting System for Adult Education: Pilot Test Measures*. National Reporting System (NRS) for Adult Education, July 1998.



6.4.2 Adjudicators. Adjudicators placed the proficiency level as anywhere from 2nd or 3rd grade level to 7th to 9th grade level. Some adjudicators suggested applicants should be proficient enough in English to participate in ordinary life activities such as shopping, banking, and going to the post office. Others suggest applicants should be able to read street signs. Most adjudicators suggested that applicants should have enough English proficiency to respond to questions about their applications, although they readily agreed that these questions often went beyond “ordinary” English. To address this issue they reported that they often rephrased questions for applicants who did not understand.

6.4.3 INS field leadership. INS field leaders were more concerned with the logistics of testing (i.e., when it happens, who conducts it, and where it occurs) than with at what proficiency level applicants are tested. One District Director believed INS should raise the “bar” for applicants by making the new test harder, and an ADDE stated that INS should use open-ended questions to test written English. In order to reduce the effects of memorization, the ADDE also believed the questions should not be available to applicants prior to the test. Other field leaders were more concerned that the test be created through a credible process and that it be consistently applied and well-monitored.

6.4.4 CBOs. When asked for specific guidelines for levels of language proficiency that INS should expect of naturalization applicants, CBO representatives provided the following:

- Phonetic spelling and incorrect grammar should be acceptable as long as sentences are understandable and communicative;
- Applicants should be able to hold everyday conversations about basic needs (enough for a tourist to survive in a foreign land); and
- English reading, writing, and speaking tasks should be related to work and personal life.

6.4.5 The public. During focus groups with the public conducted in Chicago, Houston, and San Francisco, we asked participants to vote on the English proficiency that they believed the law required. The options we presented were based loosely upon the California ESL Model Standards.²⁹ For **speaking**, we presented five options:

- Communicating with native language and gestures only;
- Communicating using a few English words and phrases;
- Communicating survival needs using learned phrases and sentences;
- Participating in basic conversations like those that occur in routine social situations; and
- Participating in face-to-face conversations beyond the basic level.

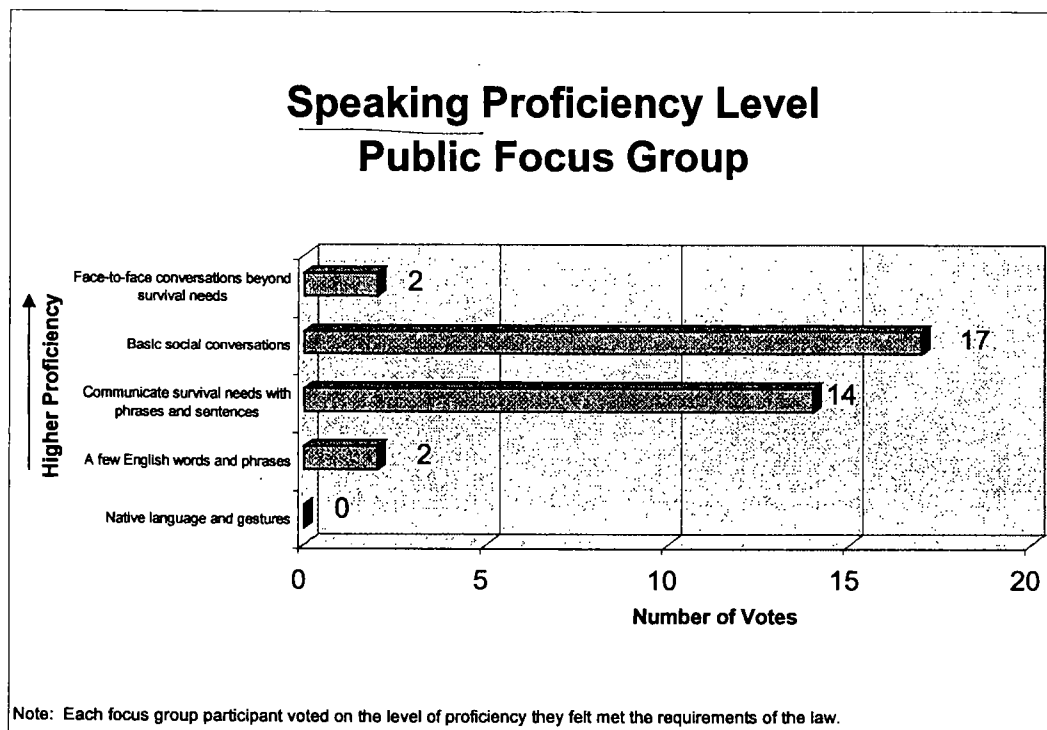
For **reading and writing**, we also presented five options:

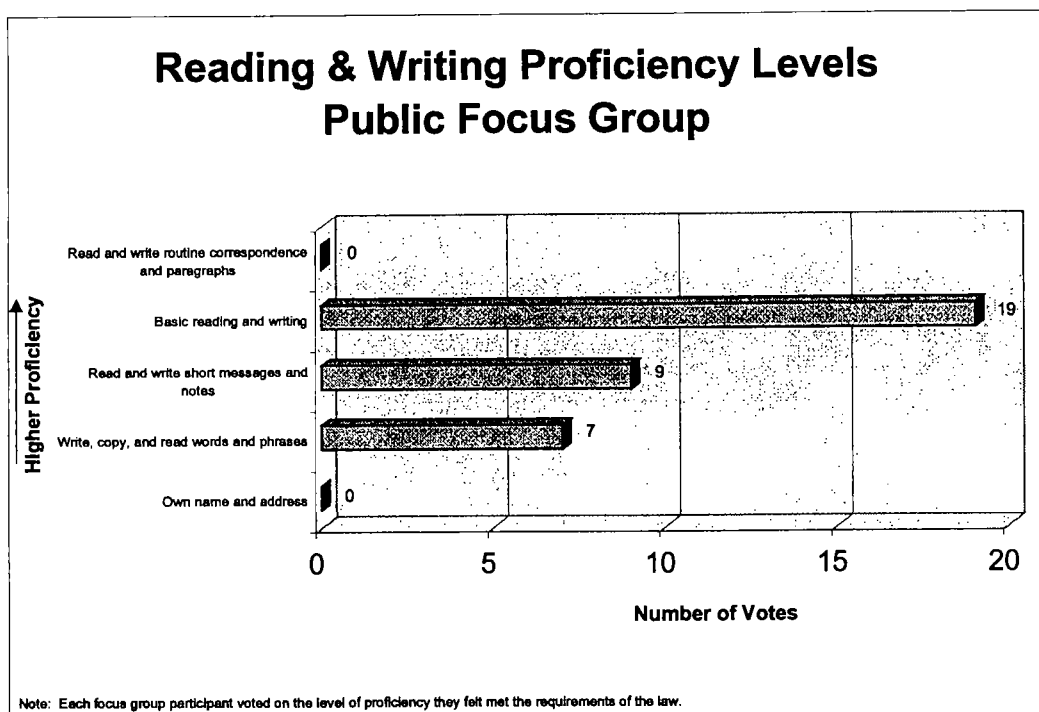
²⁹ We randomly chose these guidelines and have no preference for them as compared to the other standards described herein.



- Reading and writing name and address;
- Reading, writing, and copying words and phrases with limited understanding;
- Reading and writing short messages and notes covering familiar subjects;
- Reading and writing basic information on familiar subjects; and
- Reading authentic materials and writing about previously discussed topics.

The results of our polls follow:





6.4.6 Congress. Initially, congressional staff members felt that applicants should have enough English proficiency to do such things as to participate in the interview, participate in society politically and economically, and read *USA Today*. However, they recognized that the law might require a lower level of English proficiency.

6.5 CONCLUSION

The law requires “an understanding of the English language, including an ability to read, write, and speak words in ordinary usage in the English language...provided, that applicant[s] can read or write simple words and phrases to the end that a reasonable test of his literacy shall be made.” There are a variety of existing resources that INS can draw upon when developing a test of English proficiency. Each of the resources described in this section was developed through a credible process and, with some adaptation, could serve as a basis for English testing.

In order to help INS determine the level of proficiency that should be required, we solicited input from key stakeholders. In terms of speaking, the Consultation Group believed that applicants should be able to participate in basic daily interactions (e.g., shopping, simple social conversations). CBOs felt that applicants should be able to hold everyday conversations about basic needs. The public expected a higher level of speaking proficiency. They felt applicants should be able to participate in basic conversations like those that occur in routine social situations. INS staff had mixed feelings, ranging from requiring basic skills to requiring more complex proficiency.



In terms of reading and writing, the Consultation Group believed that the law requires the ability to read and write as necessary to function on a daily basis. CBOs believed that applicants should be able to read and write basic material related to work and personal life, with minimal regard to spelling and grammar. The public felt that the law requires applicants to read and write basic information on familiar subjects.

**PHOTOCOPY
PRESERVATION**



SECTION 7: HOW TO TEST

Once INS has made some determinations with regard to *what* it will test, it must decide what test delivery methods to use, as well as how to design specific test questions. Although *how* to test is not the focus of this report, we hope the following section provides information that INS may use to begin addressing this question.

In November, 1997, PwC submitted a *Testing Options Report* to INS, which focussed largely on how to test, providing options for both test delivery methods and question structure (e.g. multiple choice, short answer, etc.). This section draws upon pertinent information from our November 1997 report, as well as our subsequent research and the input of the Consultation Groups. We recommend that INS work in concert with a contracted test developer to make these decisions.

Section 7.1, below, provides options for test delivery methods, citing advantages and disadvantages of each. Section 7.2 presents options for the structure of the test questions. Because there are many options for structuring the test questions, we first narrow the options using a set of evaluation factors, and then discuss the advantages and disadvantages of only the most viable options. (For a more in-depth discussion of these options, please refer to PwC's November 1997 *Testing Options* report). In Section 7.3 we provide five short testing scenarios that illustrate how naturalization testing could actually be conducted.

7.1 TESTING METHODS

In determining how to test, we recommend INS first address which test delivery method(s) to use. We suggest four methods of test delivery for INS to consider:

- Oral tests;
- Paper tests;
- Computer tests; and
- Profile screening.³⁰ *include completing a course?*

In the following section, we provide a brief description, including a summary of advantages and disadvantages, of each option.

7.1.1 Types of test delivery methods. The test delivery method may vary depending on the objective of the test. Some methods may be more appropriate, for example, in testing English proficiency, whereas others may be more relevant when assessing knowledge of history and government. In the remainder of this section, we describe each of the four options that can be used for citizenship testing.

Oral examination.³¹ During oral examinations, the individual is questioned or interviewed by an examiner (or examiners) in a face-to-face situation. Oral examinations

³⁰ Through profile screening, an applicant would be exempt from some or all of the testing requirements based upon past testing, life experience, etc.



can take many forms, including conversational interviews and "question and answer" sessions. In interviews, the individual is asked a series of questions but the tone/exchange is conversational and often less formal. During a question and answer session, the individual is asked a series of questions and sometimes is instructed to follow a set of directions (point to something, stand up, etc.). The tone is more formal and more closely resembles an academic exam. The following chart cites the advantages and disadvantages of the oral examination:

Advantages	Disadvantages
• Best current option for assessing <u>spoken</u> language • Questions and answers can be clarified • Area of interest can be further examined • Better insight into higher-level thinking • Historic approach to citizenship testing	• Consistent scoring is more difficult • Harder to recall/document exam • More labor intensive • Increased stress for applicant • <u>Personalities can influence results</u> • <u>Requires more training</u> of the interviewer • Less predictable format

Paper examination.³² During a paper-based examination, the applicant is asked to provide a written answer to a question. The length, depth and development of the answer is contingent upon the purpose of examination. Paper-based examinations can be designed to fulfil many assessment purposes, including testing an individual on their ability to recall information, organize thoughts, and apply what they have learned. The pros and cons of this approach are outlined in the chart below.

Advantages	Disadvantages
• Format is familiar • Relatively easy to grade • Easy to document and store answers	• Reliability can be a problem • Cultural bias can be a problem

Computer examination.³³ During a computer examination, an individual is presented questions on a computer screen and is then required to enter a response into the computer terminal by: (1) using a keyboard, (2) using a mouse (point and click), or (3) touching the screen. As the technology has evolved, computer examinations have become a more viable and recognized method of administering a wide range of standardized tests. There are several advantages to this approach, however there are also a number of disadvantages, as evidenced in the following chart.

³¹ Much of the research for this section comes from Clift, J. C. and Imrie, B. W., (1981). *Assessing Students, Appraising Teaching*, New York: Halsted Press.

³² Much of the research for this section comes from Johnson, D.W. and Johnson, R.T. (1996). *Meaningful and Manageable Assessment Through Cooperative Learning*. Edina, Minnesota: Interaction Book Company.

³³ Much of the research for this section comes from the web pages of the Sylvan Learning Center, its subsidiary Sylvan Prometric, and Computer Assisted Testing Service.



Advantages	Disadvantages
• Unique examination can be easily created • Facilitates easy and quick scoring • Enhances customer convenience • Enables electronic transfer of results	• Requires computer skills • Limits type of questions • Higher up-front testing costs • Scoring reliability can be a problem • Cultural bias can be a problem

Profile screening. Through profile screening, an individual is exempted from an exam based upon past testing, experience, or mastery of relevant subject matter. For example, most colleges and universities offer course waivers to incoming undergraduate students who have scored well on a series of exams called Advanced Placement (AP). At the graduate school level, most business schools permit students with significant work experience to "waiver out" of required classes (e.g. an accountant with 10 years of experience is not required to take the basic accounting classes).

There are two screening approaches to consider: (a) alternative accreditation and (b) biographical profiling. Through **alternative accreditation**, applicants would supply the results of another test to establish the fact that they met some or all of the statutory requirements. For example, applicants could submit AP scores in the subjects of U. S. History and/or English. With **biographical profiling**, applicants meeting a certain profile (e.g., based on length of residency, extend of education in the U.S., employment histories, etc.) would receive a waiver for some or all of the testing requirements.

As with the other options, there are both advantages and disadvantages to consider with profile screening. These considerations are outlined in the chart below.

Advantages	Disadvantages
• Eliminates some testing – reduces testing costs • Customer friendly	• Selecting acceptable alternative accreditation may be difficult • Creating and validating profiles consistent with the requirements of the law will be challenging • <u>Diminishing returns</u> (cost to administer vs. savings realized) • <u>Possibility of fraud</u>

7.1.2 Test delivery method - evaluation factors. Each of the delivery methods described in the previous section has strengths and weaknesses. In order to begin to assess these options, INS should consider, at a minimum, the following factors:

- **Ease of scoring.** Ease of scoring should account for both the time involved in scoring a test and the level of scoring subjectivity. For example, multiple choice questions generally take less time to score than essay questions. For well designed questions, multiple choice responses are either right or wrong -- there is little discretion required on the part of the scorer. The scoring of open-ended interview questions, however, is significantly more subjective.



- **Time for test.** This factor should reflect how long it takes for an applicant to take a test. It is important for two reasons: (1) the applicant may be overwhelmed if the test is too long, but (2) the test must be comprehensive enough to meet statutory requirements.
- **Resources required.** Resources are required to administer and score exams. This evaluation factor should address both.
- **Ease of documentation.** How easy will it be to document the questions asked and answers given? It is important to ensure a record exists should evidence ever be needed after testing is complete. Also, data on performance at the question and test level is important for quality control and continuous improvement reasons.
- **Suitability for naturalization testing.** This factor should consider the appropriateness of each method vis-à-vis the several components of naturalization testing: history and government, reading, writing, and speaking.

7.1.3 Conclusion. As discussed above, there are four test delivery alternatives. Based on our research, the evaluation factors above, and input from the Consultation Groups, we believe all four options are viable for **history and government testing**. **Reading and writing** can also be tested via any of the four options (although if done orally, paper or computer support will be required – i.e., the applicant will need to be given something to read and something to write on). At this point, we believe that **oral testing** is the only way to assess **speaking** proficiency. In time, computer-based testing may be a suitable option, but given the current state of technology, it is not presently viable. Ultimately, the final decision on what testing method to use should be made after INS has determined the levels of English proficiency and civics knowledge that will be required of naturalization applicants.

Every test (whether oral, paper-based, computer-based, etc.) is comprised of a set of questions. The options described in this section involve *test delivery methods*. Below, we present options for the *type of questions* that might be used to construct naturalization tests.

7.2 HOW TO STRUCTURE TEST QUESTIONS

In this chapter, we provide options for question structure. Specifically, we present ten different question structures:

- Multiple-choice;
- True-false;
- Matching;
- Short answer;
- Completion/Cloze;
- Dictation;
- Interpretive;
- Short and long essay;
- Interview/Question & Answer; and
- Reading Aloud



After we provide brief descriptions for each option, we assess these options against a set of evaluation factors. We then provide advantages and disadvantages for only the most viable options.

7.2.1 Types of test question structures.³⁴ The types of questions used vary depending upon the objective of a test -- what is being assessed and at what level. Below we describe a variety of question types that can be used for citizenship testing.

Multiple-choice. Multiple-choice items consist of direct questions or incomplete statements (called the stems) followed by two or more possible answers (called responses or options), only one of which is correct.

True-false. With true-false questions, test-takers identify the correctness of facts, statements, definitions, or principals. Although true-false questions can be answered quickly, more questions are needed in order to overcome the possible effects of guessing.

Matching. Matching questions are comprised of a list of "concepts" and a second list of "responses". A test-taker matches one response to one concept. Matching can also be an oral exercise in which the individual chooses the picture(s) that matches a word spoken aloud by the examiner.

Short answer. During short answer tests, test-takers supply a short answer to a question. The answer is usually a name, word, phrase, symbol, or sentence. In most cases, a list of potentially correct answers is not provided.

Completion/Cloze. With completion questions, the test-taker supplies one or more words that have been omitted from a statement. When the correct word or words are provided, the statement is complete and true. Fill-in-the-blank questions generally fit into this category. The *Cloze* question is a type of completion question. It is used to assess a person's comprehension of written material. The individual reads a short passage that has been altered by removing a number of words (determined by a counting process, e.g. every nth word is removed). The individual is evaluated on their ability to understand the context of the passage and fill in the missing words.

Dictation. During dictation testing, test-takers listens to a sentence or short passage read out loud and then write what they heard.

Interpretive (with visual cues). Interpretive questions employ visuals like pictures, graphs, and maps. Individuals are required to view the visuals and describe what they see. Interpretive questions can be answered in either written or verbal form.

Short and long essay. Essay questions require the individual to prepare a written essay in response to a question. The essay can elicit a wide variety of responses -- from defining

³⁴ Much of the research for this section comes from Clift, J. C. and Imrie, B. W., (1981). *Assessing Students, Appraising Teaching*, New York: Halsted Press. and Johnson, D.W. and Johnson, R.T. (1996). *Meaningful and Manageable Assessment Through Cooperative Learning*. Edina, Minnesota: Interaction Book Company



terms, to comparing key concepts and events, to developing in-depth themes with supporting positions. The key distinction between short and long essays is the length and depth of the response provided. The long essay requires a more in-depth analysis of the topic.

Interview/Question & Answer. Oral exams can take the form of interviews, where the purpose is not necessarily to determine if an answer is correct, but rather to determine what thought process was involved in developing the answer. This is in contrast to interviews that utilize a more close ended format (Q&A questions). In Q&A testing, the purpose is to determine whether the interviewee knows the correct answers to questions.

Reading aloud. This “question structure” is used to test reading proficiency. The test-taker is asked to read aloud from prepared materials (words, phrases, sentences, passages, etc.). Unlike many of the above types of questions, which can be used to test reading by determining if the test-taker provides a correct answer to a question or an appropriate description of a passage they just read, this approach tests reading more directly through actual demonstration.

7.2.2 Question structure –evaluation factors. As with the test delivery method, there are several evaluation factors that can be used to narrow and assess the options for question structure. These factors are the same as those used to assess the test delivery method. The factors are:

- Ease of scoring;
- Time for test;
- Resources required;
- Ease of documentation; and
- Suitability for naturalization testing.

Regardless of the test delivery method or types of questions INS decides to use, it will be important to ensure all tests are developed and tested for validity, reliability, and minimal bias.

7.2.3 Viable options. Each type of question has strengths and weaknesses. Some types are clearly better than others, depending upon the testing situation and information sought. For example, multiple choice questions are comparatively easy to administer and can be constructed to assess a reasonable range of knowledge. However, multiple choice questions are not generally effective tools to determine an individual’s ability to speak English.

Based on our previous work, the evaluation criteria above, and input from the Consultation Groups, the following question structures seem most promising for the relevant components of naturalization testing:



TEST COMPONENT	TEST QUESTION STRUCTURE					
	Multiple-choice	Matching	Short Answer	Dictation	Interpretive	Reading Aloud
History/Government	✓	✓	✓			
Reading	✓	✓	✓			✓
Writing			✓	✓	✓	
Speaking			✓		✓	

Though we narrowed the alternatives for the test question structure to the most viable for naturalization testing, there are distinct advantages and disadvantages to each:

Option	Advantages	Disadvantages
Multiple Choice	• Easy to score • Guessing is generally difficult • Format is familiar • Test wide area of knowledge	• Writing good items can be difficult and time consuming • More difficult to test understanding • Perception of "multiple guess"
Matching	• Cover a large amount of material in a small amount of space • Easy to score	• Test only factual information • Possibility of guessing the correct answers
Short Answer	• Requires recall of information • Minimizes guessing • Can be used to assess all components of natz testing	• Measures only lower-level understanding
Dictation	• Easiest test of writing • Format is familiar	• Does not test understanding
Interpretive	• More structured measurement • Good for speaking assessments	• Can be time consuming • Significant training requirements
Reading Aloud	• Direct method for testing reading proficiency	• Can be time consuming • Does not necessarily test understanding

7.2.4 Conclusion. Based on the various options and the associated pros and cons illustrated in the chart above, INS may want to consider a mix of test questions structures, testing different competencies with the most appropriate type of question. In the following section, we present a few sample testing scenarios, using a combination of test delivery methods and question structures.



SECTION 7.3 TESTING SCENARIOS

In the previous sections, we presented options for both test delivery and question structure. During the consultation group meetings, we asked the members to suggest possible test administration scenarios, using the previously discussed options on how to conduct testing and structure the test questions. We compiled their suggestions into the following five testing scenarios. Note: time estimates were provided by the English Consultation Group based on their experience conducting similar tests.

Scenario I

Estimated Time: 12 minutes

The tester engages in a limited front end conversation with the applicant. During this conversation, the tester orally gathers personal information from the applicant, simultaneously testing English **speaking** skills. In order to test the applicant's **history and government** knowledge and English **reading** skills, the tester gives the applicant the history and government questions (structured as multiple choice, matching, or short answer questions) on a piece of paper for the applicant to read (silently or verbally) and answer orally. The last component of the test is dictation of two phrases/sentences to assess English **writing** skills.

Advantages	Disadvantages
• Test design imitates ordinary life/usage • Provides immediate test results • "Doubles up" on Section 312 Requirements (only 3 steps)	• <u>Tester burnout</u> (2-3 min) • Must ensure inter-rater reliability Human Factor

Scenario II

Estimated Time: 12 minutes

This option assesses English speaking and writing in the same fashion as Scenario I, but differs in the way it tests history and government knowledge and English reading. In this scenario, **history and government** questions are asked and answered orally. In order to test **reading** skills the tester has the applicant read 2-3 thematically related sentences aloud.

Advantages	Disadvantages
• Less rigorous method of testing English reading • Provides immediate test results	• Requires one more step than Scenario I (requires 4 steps) • <u>Tester burnout</u> ? • Must ensure inter-rater reliability

Scenario III

Estimated Time: 12-15 minutes

Scenario III is also similar to Scenarios I and II. In this approach, visual cues (e.g., interpretive questions using pictures, etc.) are used to stimulate conversation and



accommodate applicants with different learning styles. The visual cue is used to assess English **speaking** proficiency by asking the applicant to orally describe the picture. The visual cue is also used to assess **writing** proficiency by asking the applicant to write a description of the visual cue. As with Scenario II, **reading** skills are tested by asking the applicant to read 2-3 thematically related sentences aloud. **History and government** knowledge is tested via a set of written or oral questions (structured as multiple choice, matching, or short answer questions).

Advantages	Disadvantages
• Accommodates applicants with different learning styles • Visual cues make the test more interesting • Provides immediate test results • "Doubles up" on Section 312 requirements	• Visual cue may require greater English skills • <u>Tester burnout</u> • Must ensure inter-rater reliability

Scenario IV

Estimated Time: 10 minutes

Scenario IV is a computer-based test where the computer serves as the tester. English **speaking** is tested through voice recognition software using short answer or visual cues. English **reading** and **history and government** knowledge is tested through a short answer or multiple choice test on a touch-pad computer screen. **Writing** could be tested by using short answer questions for the history and government test or by using a computer "voice function" that dictates phrases/sentences to the test-taker.

Advantages	Disadvantages
• <u>Lowest operating cost</u> • Can do test in groups • Questions can be asked in applicant's accent • Less inter-rater reliability issues (scoring only)	• Requires computer/typing skills • Because computer scoring for speaking and writing is not yet available, <u>immediate test results are not possible</u> • <u>Technology not readily available or proven--</u>

Scenario V

Estimated Time: 15-20 minutes

This approach is a combination of an oral interview to test English reading, writing and speaking, and a computer-based test to assess history and government knowledge. English **speaking** would be assessed through a conversation between the tester and testee. English **writing** would be tested through a dictation exercise. **Reading** and **history and government** would be a multiple choice or short answer test conducted on a touch-pad computer screen.

Advantages	Disadvantages
• Requires only limited computer/typing skills • Can do computer portion of test in groups • May result in lower labor costs • Fewer inter-rater reliability issues	• May be more time consuming (2 phases – move from interview to computer) • Must still ensure inter-rater reliability • Unlikely to provide immediate test results



These scenarios provide several viable examples of a new way of testing. We recommend that the final decisions on test delivery method and question structure be made only after INS has determined the levels of English proficiency and history/government knowledge that will be required of naturalization applicants.

**PHOTOCOPY
PRESERVATION**



SECTION 8: NEXT STEPS

8.1 Introduction

The previous seven sections of this report address various options for test content and testing methods – the what and the how of testing. In order to implement the testing process envisioned in the naturalization reengineering project, several other decisions are required. Specifically, there are five decisions INS needs to make as part of the testing redesign:

What **What.** Using the information provided in this report and a test development contractor, we recommend INS develop a test program with clear objectives, meaningful credible content, and well-defined passing standards. *options for standards?*

How **How.** Using the information provided in this report and a test development contractor, we recommend INS develop a standard testing program that uses the same test designs and protocols nationwide.

When **When.** We recommend INS approve PwC's recommendation to conduct testing early in the process before applicants file their applications.

Where **Where.** INS needs to make a policy decision regarding where to conduct testing. Several options are viable. Testing can be conducted at INS offices, ASCs, public schools, community colleges, community buildings, or stand-alone test centers. *know pre case? p.66*

Who **Who.** INS needs to make a policy decision regarding who will conduct testing. Two options are viable. INS employees can conduct testing or a contractor(s) can be competitively selected to manage the testing program. At this stage, we recommend the latter approach. *delayed?*

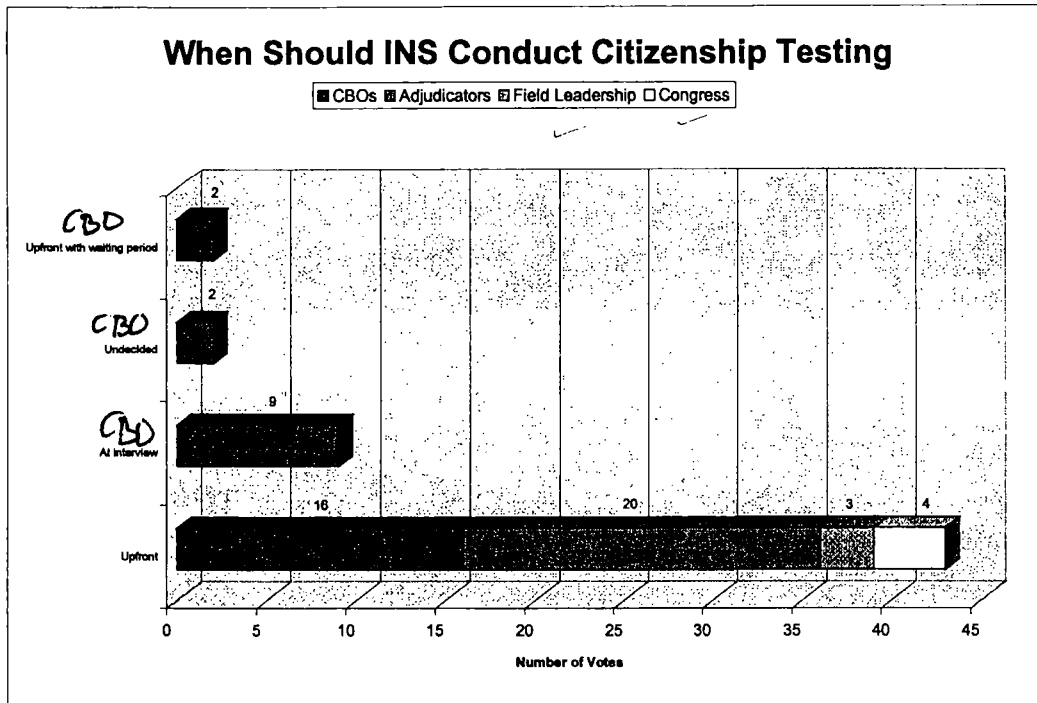
Below we provide a more ^{delayed?} discussion of these decisions, as well as a summary of the specific steps that we believe INS must take to implement a standard testing process.

8.2 Next Steps

What **8.2.1 Make high-level decisions.** Before proceeding further, INS will need to make high-level decisions on English proficiency levels and on history/government test content. In this document, we have provided options regarding both. Although it will not need to make a final decision on test delivery method or question structure when writing a statement of work for a test development contractor, INS will need to provide some high-level guidance on these issues.

How

8.2.2 Develop English test and history/government test. Once the underlying questions that will frame the direction of testing have been answered, PwC recommends that an outside contractor be competitively selected to develop the new naturalization tests. Such



8.2.4 Determine where to conduct testing. INS must decide where to conduct testing. Until a comprehensive feasibility study can be done, PwC does not have a recommendation on testing sites. However, there seem to be six viable venues to consider – ASCs, INS offices, public schools, community colleges, contracted space, and community buildings (e.g., churches, public buildings). Below we provide a table of advantages and disadvantages for each option:



	# OF TRIPS	ADVANTAGES	DISADVANTAGES
ASCs	3 trips (FP, test, interview) to 2 locations (ASC and INS office)	• Convenient locations • Good INS oversight • Designed to facilitate processing of people • Stable location	• Not enough space at some ASCs • Not currently designed to accommodate testing
INS Offices	3 trips (FP, test, interview) to 2 locations (ASC and INS office)	• Good INS oversight • Stable location	• Not enough space at many offices • Not currently designed to facilitate testing • Not as convenient as ASCs or community buildings
Public Schools, Community Colleges <i>mostly just San Fran.</i>	3 trips (FP, test, interview) to 3 locations (ASC, test location, INS office)	• Convenient locations • Credible locations • Less intimidating setting • Capacity for large test administrations	• Must negotiate use, schedules, etc. • Less stable • Limited testing times (e.g., weekends) • Harder to manage logistics • Less security • Harder to ensure INS oversight • Requires 3 locations to complete Natz process
Contracted Space	3 trips (FP, test, interview) to 3 locations (ASC, test location, INS office)	• Convenient locations • Can design to facilitate testing • Can select to ensure adequate space • Less intimidating setting	• Probably more expensive option • Harder to ensure INS oversight • Requires 3 locations to complete Natz process
Community Buildings <i>mostly just San Fran.</i>	3 trips (FP, test, interview) to 3 locations (ASC, test location, INS office)	• Convenient locations • Less intimidating setting • Capacity for large test administrations	• Must negotiate use, schedules, etc. • Less stable • Limited testing times (e.g., weekends) • Harder to manage logistics • Less security • Harder to ensure INS oversight • Requires 3 locations to complete Natz process

In order to address the question of where to conduct testing, we recommend two steps:

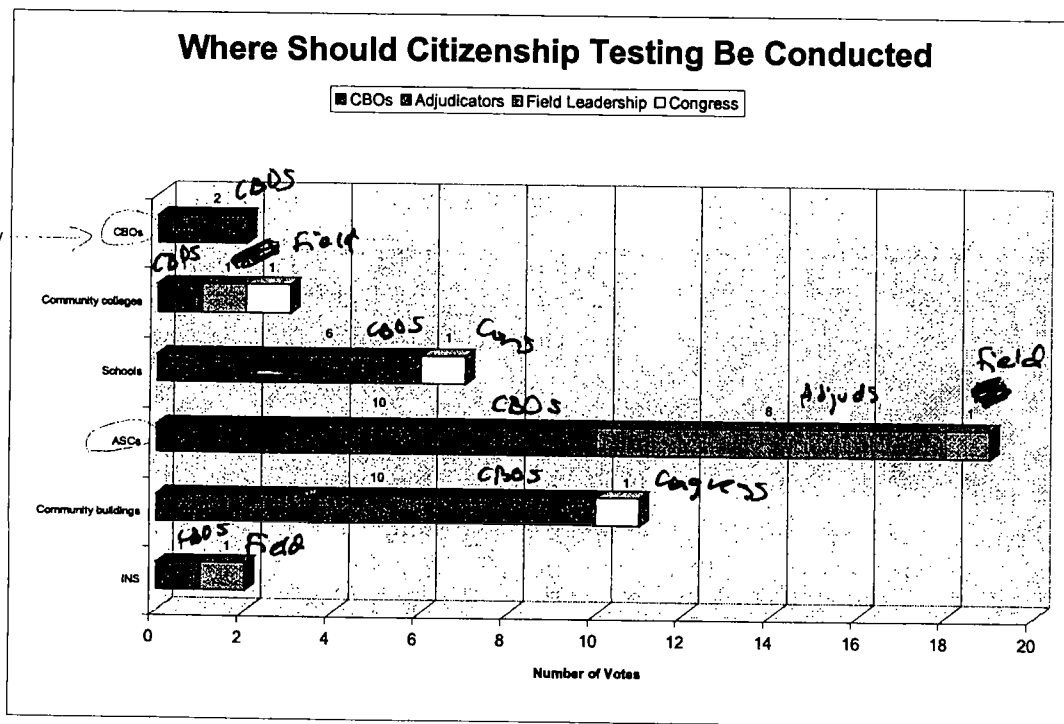
Ask customers. Because customer service is a Commissioner priority and applicant comfort is an important factor to consider vis-à-vis testing, we recommend INS survey a sample of individuals who have recently completed the naturalization process to ask their opinion on testing locations. The results of this survey will help inform INS's decision on where to conduct testing.



Conduct a feasibility study. Determining the best locations to conduct testing will also require a more comprehensive analysis of the options. Specifically, INS will need to conduct (1) an analysis of testing demand by location, (2) an analysis of the space needed to satisfy that demand, and (3) an analysis of the space available for each of the four options (ASCs, INS offices, contracted space, and community buildings). INS will also need to understand the implementation cost and timeline of each option. Once these analyses are completed, a fair comparison of the four options can be made, taking into account the above pros and cons, as well as the costs associated with each.

what
get
public sch.
space

Although stakeholders' opinions vary on the issue of test locations, locations out "in the community" (e.g., schools, ASCs, and community) appear to be the preferred options:



? of
Contracted
space!

red
color

Why can't still not like ASCs (skewed neg.)
why ASC not seem to be INS?
like the 1-stop shopping



Who

8.2.5 Determine who will conduct testing. INS must also decide who will administer the new citizenship tests. INS can administer the tests with INS staff or INS can enter into a no-cost contract with outside testing organizations. The chart below outlines several of the advantages and disadvantage for each option.

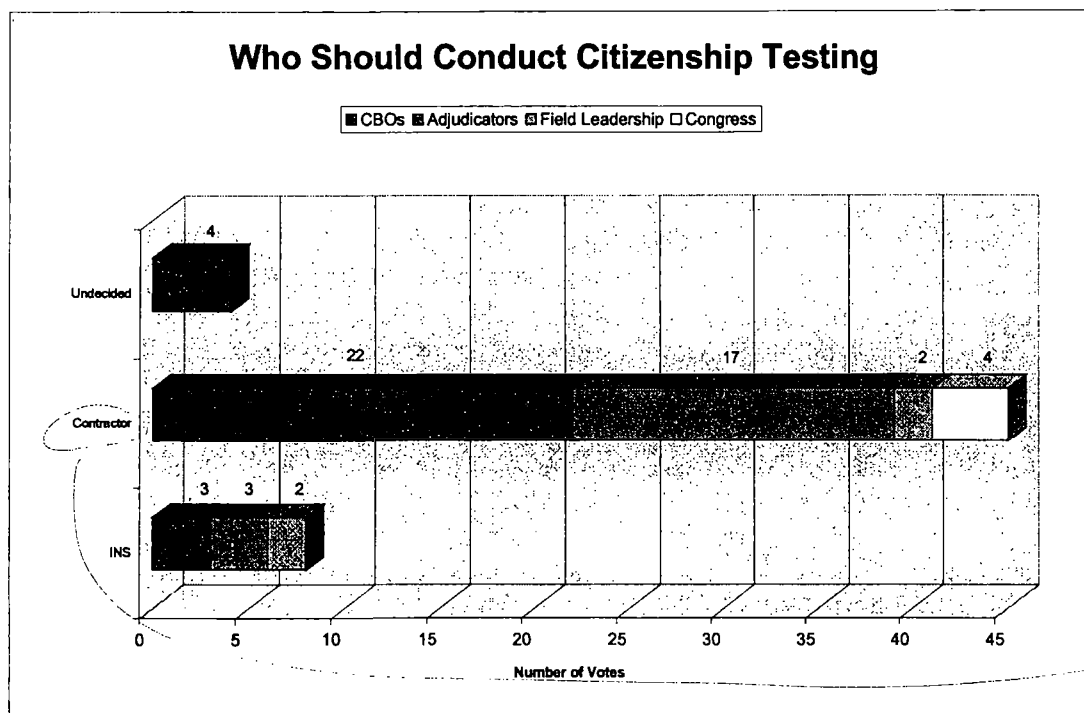
	ADVANTAGES	DISADVANTAGES
INS Administrators	INS will have <u>direct control</u> over process	- Testing is not an INS core business - INS will need to recruit, hire, and train staff - INS will need to handle scheduling, administration, and scoring of tests
Contractor Administrators	- Expertise in test administration - Experience in test administration - Shorter implementation time-frames - INS would not have to hire staff or handle scheduling, administration, and scoring	No direct INS control -- INS will need to have <u>strong oversight</u> and quality control mechanisms

Consistent with the above table and further to our September 1997 report and the recommendations of the Commission on Immigration Reform, we recommend INS contract with two professional testing organizations to administer citizenship testing.³⁷ We recommend using two providers because we believe it will motivate high performance (through comparison and competition). Two providers will also afford INS the option of voiding a contract for poor performance with minimal disruption to the overall program. We also recommend INS utilize performance-based service contracts to ensure high-quality responsive performance. Additionally, the selected vendors should be reputable testing organizations with extensive experience in high stakes testing. Finally, systematic and regular quality control and oversight should be built into the contracts and full-time INS staff should be designated whose exclusive responsibilities will be its management and oversight.

While most of the external stakeholders we spoke with agreed that contractor-administered testing is a good approach, INS employees felt strongly that close INS monitoring, and perhaps even on-site INS supervision will be required to make this work.

site

³⁷ INS sponsored a testing program where five outside testing agencies conducted testing prior to the naturalization interview. This program has now been discontinued. We believe that this program failed because: (1) there was only a limited relationship between INS and the outside providers, (2) there were no established performance standards, (3) there was not enough security and monitoring built into the system, and (4) each provider developed their own test and protocols without much INS involvement.



Color
Why?
GBO's not
that
INS

8.2.6 Plan and prepare for implementation. Once standard tests have been developed and decisions have been made on when to test, who will administer tests, and where tests will be given, the detailed implementation planning can begin. If INS decides to contract with outside contractors to administer testing, INS will be responsible for two primary tasks—selecting the contractors and monitoring/supporting their efforts. In order to provide proper contractor support, INS will need to convene a management team to perform oversight during all stages of the contract. PwC estimates that INS will need a team of five members.

Although we recommend INS contract with professional testing firms to administer testing, in-house testing is still a viable option. This effort would require a full-time INS management team of 15-20 people.

Below, we describe the major planning activities that will be required to roll-out the new testing program. The steps listed are largely the same regardless of whether testing is conducted through contractors or in-house. There are two major differences: (1) if contractors are used, INS will not directly conduct most of the described steps, but rather will supervise the contractors as they perform the work and (2) using contractors will reduce the time required to implement the new program.

2
Advantage

8.2.7 Select two contractors (only if INS opts to out-source testing). If INS opts to conduct testing via contractors, the first step of the process is to craft an RFP and SOW. PwC recommends that the SOW be performance based, with clear outcome-oriented performance standards. Examples of performance standards include “no applicant must travel more than 100 miles to get to a testing location” and “no applicant must wait more than 1 month to be scheduled for a test.” We also recommend that meaningful incentives



for good performance and consequences for inferior performance are built into the contracts. Finally, strict quality control and monitoring provisions must be a key component of the contracts. The contractor selection process will be similar to the one described in Section 8.2.2 above. We estimate it will take 4 months.

8.2.8 Budget for planning phase and, eventually, test administration. If INS conducts testing in-house, it will need to budget for all of the expenses associated with implementation (e.g., equipment, scoring software, etc.) and on-going operations (e.g., proctors, tests, etc.). If testing is outsourced, INS will only need to budget for on-going monitoring expenditures and for possible one-time implementation expenses (e.g., modifying ASC floor plans).

8.2.9 Convene testing management team. The testing management team will be responsible for overseeing the work of the contractors (or for actually performing all of the tasks described below if INS decides to do testing in-house). Ultimately, this team will be responsible for ensuring the new testing program is fully implemented in an effective and well-managed fashion.

8.2.10 Identify specific testing sites. At this point, INS³⁸ will have already decided where to conduct testing (ASCs, INS offices and/or community buildings). As a result of the feasibility study described above, INS will also have a good idea of the space needed for testing on a district-by-district basis. With this information, INS will be able to finalize a list of specific testing sites.

8.2.11 Acquire/modify space. Regardless of what option is selected vis-à-vis testing locations, we expect INS will need to acquire new test sites or modify existing facilities. For instance, if INS decides to conduct testing at ASCs, INS or the contractor will need to compare space needs against current space at each ASC and, where necessary, modify current floor plans or obtain additional space.

8.2.12 Map core processes. While space is being acquired/modified, INS can simultaneously start mapping core processes. It will be necessary to determine how to schedule applicants, collect fees, conduct check-in, give tests, score tests, and notify individuals of results. Generic process “maps” should be developed that can eventually be modified to fit each individual site. It may be necessary to develop several generic models (e.g., small test site, medium test site, and large test site).

8.2.13 Develop standard operating procedures (SOPs). Standard operating procedures will need to be developed based on the processes designed above. The SOPs should provide step-by-step detail for each key task in the process. They should also cover customer service and special accommodations procedures.

8.2.14 Develop floor plans. Once the list of test sites is finalized and core processes are mapped, INS should develop model floor plans. These plans will serve as guides that can

³⁸ In the remainder of this section, we recommend tasks that “INS” should undertake. However, if INS opts to hire contractors to administer testing, much of the work required can and should be managed by the selected contractors.



be tailored to accommodate the requirements and circumstances of each testing site. For example, Helena, Montana will have a process very similar to Los Angeles, however the testing site floor plans may be very different, given the anticipated variations in testing volume.

8.2.15 Conduct staffing analysis. Approximately halfway through the test site identification/modification process, we recommend INS initiate a staffing analysis to determine how many proctors and support personnel will be needed at each center. This analysis should be based on the size and number of test sessions per day, and should follow the best practice of one proctor for every twenty test-takers.

8.2.16 Hire staff. If testing is administered in-house, INS will need to hire the requisite number of staff as determined by the staffing analysis. INS will need to create vacancy announcements, assess/interview candidates, and hire the necessary staff. If a contractor is used, they will be expected to recruit and hire qualified staff, consistent with clear qualification requirements that should be written into the SOW.

8.2.17 Train staff. Once staff have been hired for each testing location, they must be trained on the testing SOPs and on providing quality customer service. If testing is done in-house, INS will need to create training materials and provide the appropriate amount of training to ensure integrity and a quality testing program. If a contractor is used, they will be expected to train the staff.

8.2.18 Create test center management system. INS will need to develop a test center management system. This system will either be a module in or interface with CLAIMS. Through the system, test scores will be stored and transmitted. Performance metrics will also be recorded. Specifically, this computer software will accommodate the following:

- The entry of basic applicant information using bar-codes located on the testing authorization form;
- The capture of applicant test scores;
- The transmission of this information to a CLAIMS; and
- Regular reporting of key performance data.

Once the basic requirements of the system are identified, specific system architecture will be designed. INS will then need to design and ultimately pilot the system. We recommend a 4-month pilot and subsequent system refinement be scheduled to ensure the system is ready in advance of the overall testing pilot.

8.2.19 ~~Revise the Federal Regulations.~~ Section 312 in the Aliens and Nationality Code of Federal Regulations must be changed to accommodate the testing redesign. Specifically, the text that discusses the content of the tests and the text that describes the test as part of the interview must be modified. We recommend these revisions be completed before the new testing program is piloted. We estimate that this effort will take approximately eight months.



8.2.20 Manage the change. The new testing program will affect three key groups of stakeholders: adjudicators, applicants, and CBOs. The majority of the change management effort will occur late in the process because little change will be realized for up to 12 months. Please see PwC's *Testing Business Case and Implementation Plan* for a more detailed change management plan.

8.2.21 Roll-out new testing program. As INS gets closer to implementation, it will need to determine how to best roll-out the new citizenship testing program. There are several ways INS can roll-out the new program. Testing can be implemented on a geographic basis (e.g. east to west, region by region, etc.). It can also be rolled-out on a district-by-district basis (e.g. first to a representative mix of high and low volume districts, first to the "big five," etc.). INS can also initiate the new program nationwide on a single date. At least six months prior to implementation, we recommend INS develop a detailed system-wide plan that addresses how and when the new citizenship testing program will be rolled-out to the field.

8.2 CONCLUSION

There are significant benefits to be realized by redesigning the naturalization process as described above:

Increased integrity. By the end of this effort, INS will have meaningful and credible tests with consistent, minimum levels of passing nationwide. INS will also have standardized processes and protocols to uniformly and fairly test applicants. Finally, the integrity of citizenship testing will be enhanced because strong security and quality assurance mechanisms will be built into the process.

Improved screening. By testing applicants at the beginning of the process, INS will screen ineligible applicants out of the process before they apply. Eliminating ineligible cases from the process will help free interview slots and ultimately reduce backlogs. Our latest survey results show that 24% of continuations occur when applicants fail the citizenship test during their first interview. By testing prior to application INS can clear ^{could} at least 72,000 interview slots in FY 1999.³⁹ Moreover, eliminating testing from the interview process will allow adjudicators to conduct more focused interviews, centering on the other eligibility requirements that they are better suited to assess. (smiley face)

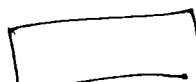
Improved customer service. A fair and consistent process is perhaps the greatest customer benefit of the new testing program. Separating testing from the interview process may also reduce some of the anxiety applicants feel because of the investigatory nature of the adjudication interview. Additionally, testing before applications are filed will allow potential applicants to determine whether or not they are eligible for naturalization before they invest the filing fee. Finally, because naturalization is a fee-based program, those who currently meet the English and civics requirements are paying a higher fee to cover the extra costs associated with continuations. Moving testing to the

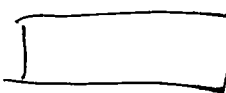
³⁹ 700,000 (receipts in FY 99) * 43% (cases continued) * 24% (of cases continued, those that are continued for testing) = 72,240 cases that no longer require two interviews.

mislaid
can't do
this
FY 1
true?
but
no
12/31/98
back pay to PwC
save up



beginning of the process (and significantly reducing continuations) will mitigate such subsidies.

100 225 

75 222 



APPENDIX A: HISTORY/GOVERNMENT ASSESSMENT CONSULTATION GROUP CHARTER AND ABBREVIATED BIOGRAPHIES

CHARTER

Introduction. As part of a larger reengineering project, Coopers & Lybrand, L.L.P (a precursor firm to PricewaterhouseCoopers) recommended that the citizenship testing program be redesigned. The Department of Justice (DOJ), the Immigration and Naturalization Service (INS), and PricewaterhouseCoopers (PwC) all concur that the first step in this challenging process is to bring together a group of subject matter and testing experts to advise and confer with PwC.

Background. 8 U.S.C. 1423 (Section 312 of the INA) requires applicants to demonstrate “a knowledge and understanding of the fundamentals of the history, and of the principles and form of government, of the United States....” The law is silent regarding *what* the fundamentals of U.S. history are. It is also silent on what new citizens should know regarding the principles and form of U.S. government. Finally, the law does not detail *how* INS should determine whether an applicant meets these requirements.

PwC is interested in learning how experts in the field of American history and government interpret the statutory language, as well as what knowledge of history of government and what level of understanding they believe should be expected of new citizens. Ultimately, this information will support the development of a standard assessment program.

Key Questions. The first step toward developing a standard assessment program is to address the following questions:

1. What are the objectives of the requirements of 8 U.S.C. 1423 (Section 312 of the INA)?
2. Given these objectives, what is the specific body of knowledge and level of understanding applicants should possess about U.S. history and our principles and form of government?
3. What is the best way to determine whether an applicant has met these requirements?

Consultation Group Mission. The Consultation Group will work in an advisory capacity with PwC to begin to address questions 1 and 2 above. Question 3 will also be considered as it informs questions 1 and 2. However, most of the work associated with addressing question 3 will be done through a contractor. The Consultation Group will not develop or recommend a single answer, but rather, will provide options and supporting information to PwC to, in turn, present to DOJ and INS for consideration.

Level of Effort. C&L will conduct approximately three full-day working sessions with the Consultation Group over a three month period.



HISTORY AND GOVERNMENT ASSESSMENT CONSULTATION GROUP ABBREVIATED BIOGRAPHIES

Charles Bahmueller works on special projects for the Center for Civics Education. He worked on the first draft of the National Standards for Civics and Government and is an expert on civics standards. Dr. Bahmueller is also working on the revision of the California History Standards.

Brenda Bell is an associate director of the Center for Literacy Studies at the University of Tennessee. She also directs the Civic Participation Standards Project of Equipped for the Future, an initiative of the National Institute for Literacy.

Noah Pickus is a professor at the Sanford Institute of Public Policy at Duke University. Dr. Pickus was the chair of *Becoming American/American Becoming*, a Duke University workshop on immigration and citizenship. His research interests lie in constitutional law, American political development, immigration policy, and citizenship theory.

David Warren Saxe is a professor of social studies education at the Pennsylvania State University. Dr. Saxe is an expert on state history standards and is the author of the Fordham Foundation report, *State History Standards: An Appraisal of History Standards in 37 States and the District of Columbia*.

K. Scott Wong is a History Professor at Williams College. Dr. Wong's research interests lie in Asian American History, immigration history and the Asian immigrant experience in America.