Status of Language Testing Reforms:
Briefing for SLS Instructional Staff
National Academies Report

A Principled Approach to Language Assessment: Considerations for the U.S. Foreign Service Institute

FSI Review

2020 – Publication and Internal Review

2021/22 - Language Testing Task Force

• Reviewed study
• Made recommendations
• Implementation started in July 2022

Three Initial Task Force Deliverables (2021)
• Statement of Test Purpose, Use, and Impact
• Selection of 13 approaches
• Long-term institutional structure to guide and support language testing

Follow-on Deliverables (2022)
• Generate guidance for test design team
• Review design work and make recommendations for adoption
Statement of Test Purpose, Use, and Impact

Task-based constructs

- Integrated language skills, including listening, speaking, and reading
- Typical and high-frequency work interactions
- Formal and informal settings
- Building relationships and social capital
- Conducting official U.S. government business
- Gathering and reporting information
- Communicating U.S. policy publicly and privately

Criteria:

- Capitalize on the strengths and diversity of the workforce
- Employ a comprehensive approach
- Ensure tests and scores are relevant, fair, useful, and realistic
- Assess integrated language skills in speaking, listening, and reading
Task Force Recommendations from 13 NAS Study Approaches

- **In implementation:**
  - More transparent scoring criteria
  - More detailed score reports
  - QA by using additional scorers
  - Scenario-based assessment
  - Multiple measures
  - Scoring listening on the speaking test

- **Pending further consideration:**
  - Utilizing recorded listening tasks
  - CAT for reading and listening
Test Design Parameters:

- Keep tests roughly the same length as current tests.
- Administer any future computer-administered reading/listening test preferably on the same day as the speaking/listening test, or the day before or after the speaking/listening test, if it better fits the examinee's schedule.
- Embed any task preparation or topic priming (e.g., introductory reading or listening passages) in the same test session.
- Ensure smooth flow of the test without abrupt interruptions, segmentations, or transitions to different parts of the test.
- Explore use of language supports for task preparation and topic priming while ensuring test security.
Gap Analysis of Speaking Test

- Address potential for bias (EEO protected categories)
  - ✓ Professional rather than personal introduction
- Eliminate presentation / explanation of topic in Speaking Part II
  - ✓ Scenario-based task to recast Work-related Exchange
- Improve flow of test to make conversation feel more natural and authentic
  - ✓ Pre-conference / pre-meeting scenario throughout test
- Reweight scoring criteria to give listening equal weight with speaking
  - ✓ New integrated speaking and listening task rubrics
- Reweight scoring criteria to give greater weight to ‘communicative effectiveness’ and ‘interactional competence’ over ‘grammar’ and ‘vocabulary’
  - ✓ New order of main abilities: 1) Functional Ability, 2) Content Meaningfulness, 3) Contextual Appropriateness, 4) Precision of Forms and Meanings
Gap analysis of Reading Test

✓ Maintain current format with oral report in English
✓ Explore CAT test, but not a high priority
✓ Eliminate focus on esoteric language use, e.g., use of poetry and testing of literary / cultural knowledge for upper proficiency bands

Scoring both Speaking and Reading

✓ Cap scoring at ILR 4, in line with Department requirements
✓ Develop integrated scoring rubrics specific to Foreign Service domain, and map to the ILR framework
Washback

Language use at post

Instruction

Testing

General proficiency

Foreign Service context
Reviewed >100 suggestions from Test Design Competition and Test Design Team

Developed prototypes

Conducted >80 testing sessions

Current version of revised speaking test
Social Event

Pre-conference or pre-meeting scenario

[a] Setup
[1] Social Conversation
[2a] Q&A
[2b] Topical Conversation
Pre-conference / Pre-meeting Scenario

**Set Up**
- Introduce test format
- Select material

**1. Social Conversation**
- Ease examinee into the test context
- Determine proficiency band—low, mid, high

**2a. Q&A**
- Transition conversation to agenda
- Establish listening working level

**2b. Topical Conversation**
- Establish speaking working level and ceiling

**3. Gathering and Reporting**
- Establish listening ceiling

Task purpose
speaking test structure

set up
- introduce test format
- select material

1. social conversation
- ease examinee into the test context
- determine proficiency band—low, mid, high

2a. q&a
- transition conversation to agenda
- establish listening working level

2b. topical conversation
- establish speaking working level and ceiling

3. gathering and reporting
- establish listening ceiling

replacement for speaking part ii: work-related exchange
Speaking Test Structure

**Set Up**
- Introduce test format
- Select material

**1. Social Conversation**
- Ease examinee into the test context
- Determine proficiency band—low, mid, high

**2a. Q&A**
- Transition conversation to agenda
- Establish listening working level

**2b. Topical Conversation**
- Establish speaking working level and ceiling

**3. Gathering and Reporting**
- Establish listening ceiling

**Integrated speaking and listening skills**
- Communicating U.S. policy publicly and privately
- Building relationships and social capital
- Gathering information
- Typical and high-frequency work interactions
- Conducting official U.S. government business
• Written in English

• Agenda:
  • 5 topics per career track
  • 2 career tracks

• Career tracks:
  • Consular
  • Economic
  • Diplomatic Security
  • Management
  • Political
  • Public Diplomacy
  • International Development
  • IT

• Example set of agendas:
  • Consular/IT
  • Economic/Political
  • Management/Public Diplomacy
  • Diplomatic Security/International Development
Speaking Test
Part I: Social Conversation

- **Purpose**
  - Determine proficiency band—low, mid, high
  - Ease EE into the testing context

- **Scenario**
  - Pre-conference/meeting gathering
  - Examinee: conference/meeting attendee
  - Examiner: English-speaking colleague of Examinee
  - Tester: conference/meeting attendee who speaks the language of the test

- **Components**
  - Professional introduction
  - Casual conversation
  - More complex conversation
Speaking Test
Part IIa: Q&A

- **Purpose**
  - Transition conversation to agenda
  - Establish listening working level

- Testers asks the Examinee about information on the agenda.
Speaking Test
Part IIb: Topical Conversation

- **Purpose**
  - Establish working level and ceiling for speaking
  - The Tester discusses 2-3 professional topics with the Examinee in the language of the test.
Speaking Test
Part III: Gathering and Reporting Information

- **Purpose**
  - Establish ceiling for listening.

- Examinee selects 1 topic and interviews the Tester in the language of the test.

- The Examinee reports the Tester's responses to the Examiner in English.

- The Examinee asks 4 main questions (not including follow-up questions).
**Speaking Test Structure**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Introduce test format</td>
<td>• Ease examinee into the test context</td>
<td>• Transition conversation to agenda</td>
<td>• Establish speaking working level and ceiling</td>
<td>• Establish listening ceiling</td>
</tr>
<tr>
<td>• Select material</td>
<td>• Determine proficiency band—low, mid, high</td>
<td>• Establish listening working level</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Questions


LTA Websites
OpenNet:
http://fsi.state.gov/SLS/6951

SharePoint:
https://usdos.sharepoint.com/sites/SLS/LTU/SitePages/Home.aspx