Frameworks, Standards, and Scales in the USG
Role of the ILR Testing Committee

- Provide forum for discussion
- Build network among government, academe and industry
- Develop understanding of Skill Level Descriptions
Characteristics of the ILR Scale

- Non-negotiable frame
- Mastery principle
- Proficiency not performance or achievement
- Global rating
- Apex is well-educated native speaker
Overview of Presentation

- Use of the ILR Self Assessments
- OPI Summits
- First Listening Summit
- ASTM Standard Practice
Frameworks, Standards and Scales

Impact of Internet-based Speaking Self-Assessment Tool on Department of State New Hires
Setting

Until September 2007 – one pencil and paper Foreign Service Examination/year resulting in:

- Approximately 400 applicants/year tested for language proficiency in languages considered useful to the Foreign Service
Since September 2007 – 4 online Foreign Service examinations/year resulting in:

- Approximately 400 applicants/year tested for language proficiency in languages considered useful to the Foreign Service

plus

- Approximately 1000 additional tests in Super Critical Need Languages: Arabic, Chinese, Hindi, Urdu, Dari, Farsi
Setting: expectations

Expected threshold level for testing in:

- **World languages:**
  - Speaking Proficiency 3
- **All other languages:**
  - Speaking Proficiency 2
Challenges

- High testing numbers with strict deadlines
- Low passing rates
Solution

Internet-available Speaking Self-Assessment Tool on the ILR Homepage

http://www.govtilr.org/
Results

- **Before Speaking Self-Assessment Tool:**
  - Applicants tested: 85%
  - Passing rate: 42%

- **After Speaking Self-Assessment Tool:**
  - Applicants tested: 63%
  - Passing rate: 63%
Additional Results

- More realistic expectations
- Happier examinees
- Test administration resource savings.
Frameworks, Standards and Scales

Oral Proficiency Interview Summits

Listening Summit
Impetus for Speaking and Listening Summits

- Started in June 2008 as a result of ACTFL OPI work for DLI at Mid & Higher levels
- Questions about comparability at the higher levels in Speaking
- Success of Speaking Summits led naturally to Listening
- Opened up to the wider ILR community and non-USG
Summit Overall Goals

- Create common understandings
- Achieve comparable scores
- Supplement the ILR Skill Level Descriptions
- Long-term goal to update the SLDs within the ILR Framework
OPI Summit Highlights

- Highlighted similarities and differences in approaches to testing Speaking
- Recognize that agencies’ tests are different, but all aim to measure speaking in the ILR context
- Uncovered differences in interpretations of the SLDs
- Discussion on keeping the Educated Native Speaker at the Apex of the Scale
- Agreement to continue to tighten wording and standards
OPI Summit Outcomes

- Work underway to write explanatory appendix of terms
- TAEG-Sponsored study of OPI comparability has been designed and is waiting approval
Listening Summit Goals

- Surface areas in the SLD that need our attention
- Build a common understanding on the challenge of listening proficiency
Listening Summit Outcomes

- Identified participative and non-participative listening as a major issue
- Generated a list of terms that need to be further defined and/or clarified
- Recognized need to bring in new research in the field
Next Steps

- OPI Summit work is underway to draft explanatory notes for Speaking Skill Level Descriptions
- Next Listening Summit in the winter 2010
Frameworks, Standards and Scales

Development of the
ASTM Standard Practice
What is ASTM?

• ASTM: American Society for Testing and Materials
  ➢ 30000+ industry standards
  ➢ Develops but **does not enforce** standard Independent, private sector, not-for-profit

➢ Mission is to provide a system for experts to directly participate in developing market relevant, high quality, international standards

➢ Standards developed in an open process involving all interested parties (government & industry)

➢ ASTM does not enforce standards

➢ Standards may be referenced in contracts
The ASTM Standard

- Many ILR users are competent test developers and can design their own test specifications

- **BUT**
  - Some ILR users need a test but do not have the expertise to develop it or to determine what existing tests might suit their need
  - ILR users might differ in their interpretations of the ILR for testing; these differences might not be apparent in test specifications
The ASTM Standard

- Having an overarching standard allows multiple test developers all referencing the ILR to have a common reference point.
- A formal standard is transparent and allows contractors and government agencies to know what is required.
- Standard is a basis for quality assurance.
- Standard guide vs. standard practice vs. standard specification.
Challenges

- The ASTM Standard Practice is **NOT** an interpretation of the ILR
- As of now, there is no enforcement mechanism for the standard
- The standard is modular, since different testing needs require different types of tests and test development processes—not lockstep standardization
Looking forward

- The process of articulating the standard has brought about unprecedented dialogue among different agencies and contractors about what is expected and needed.

- The requirements for needs analysis and frameworks that are stated in the standard practice raise awareness among stakeholders that will lead to better-fitting tests.
Conclusions

- Broaden access to the ILR framework
- Improve common understanding of the ILR scale
- Update Skill level Descriptions