



Interagency Speaking Summits: A community of progress

Dr. Rachel Brooks (FBI) and Dr. Clayton Leishman (DLIELC)

ILR Testing Committee

February 25, 2022

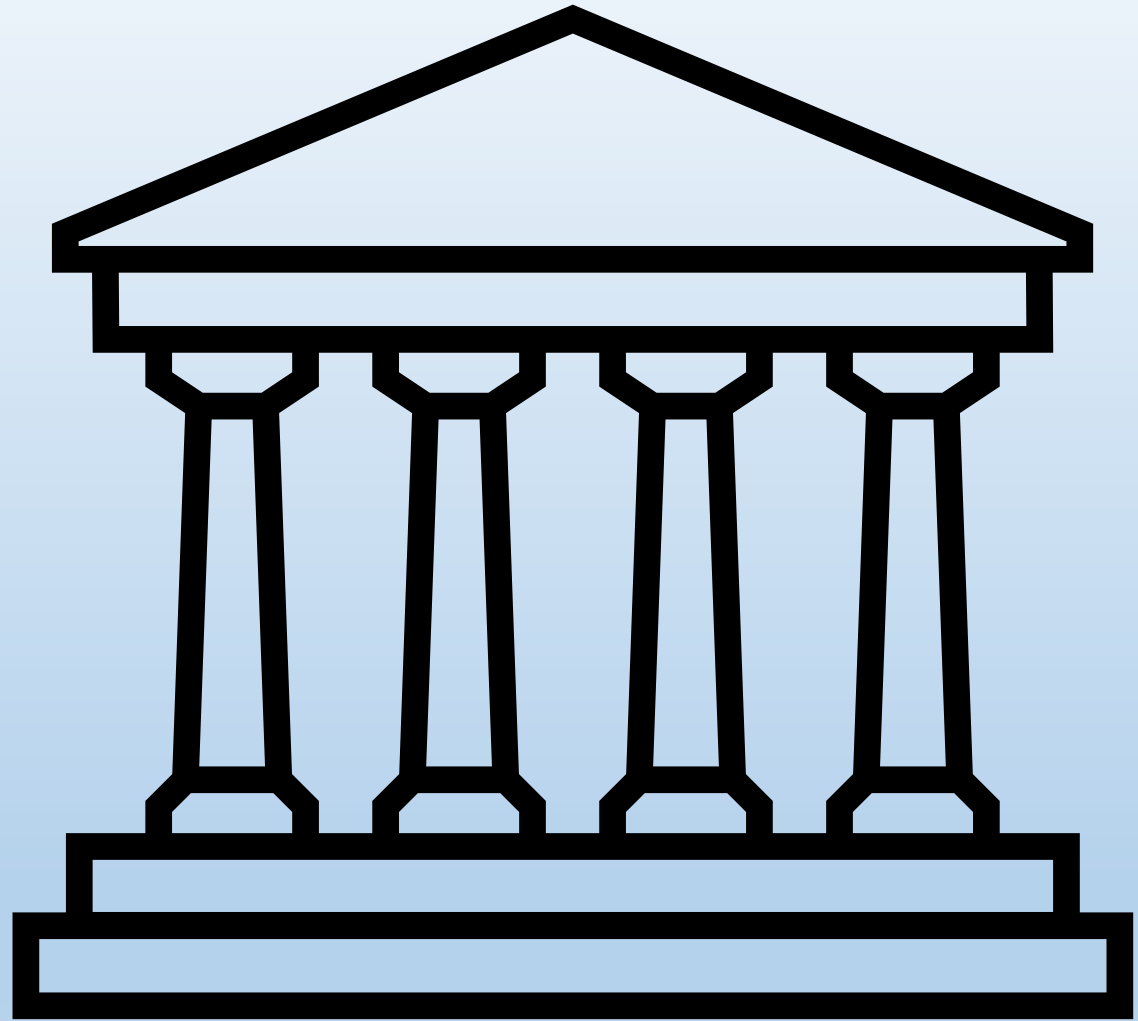
Objectives



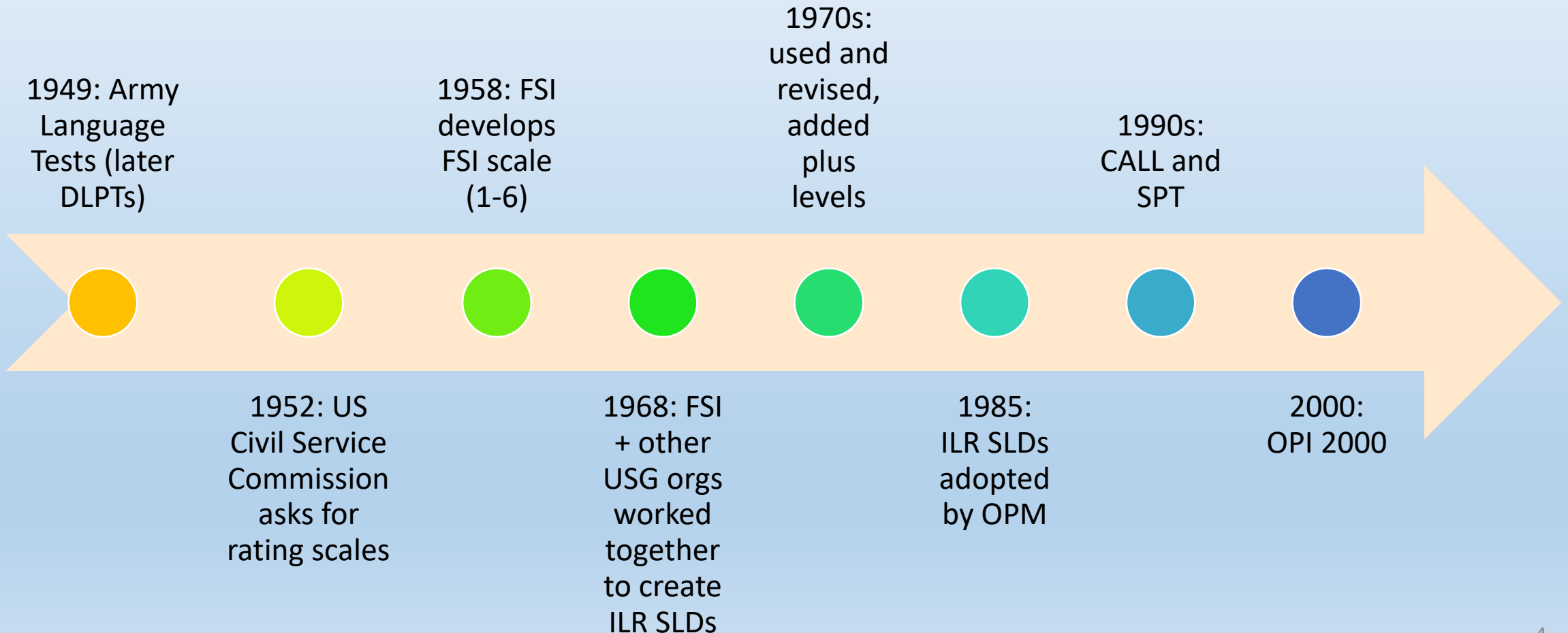
- Give overview of Interagency Speaking Summits
- Report on event of ISS 2021
- Discuss the way forward for ISSs



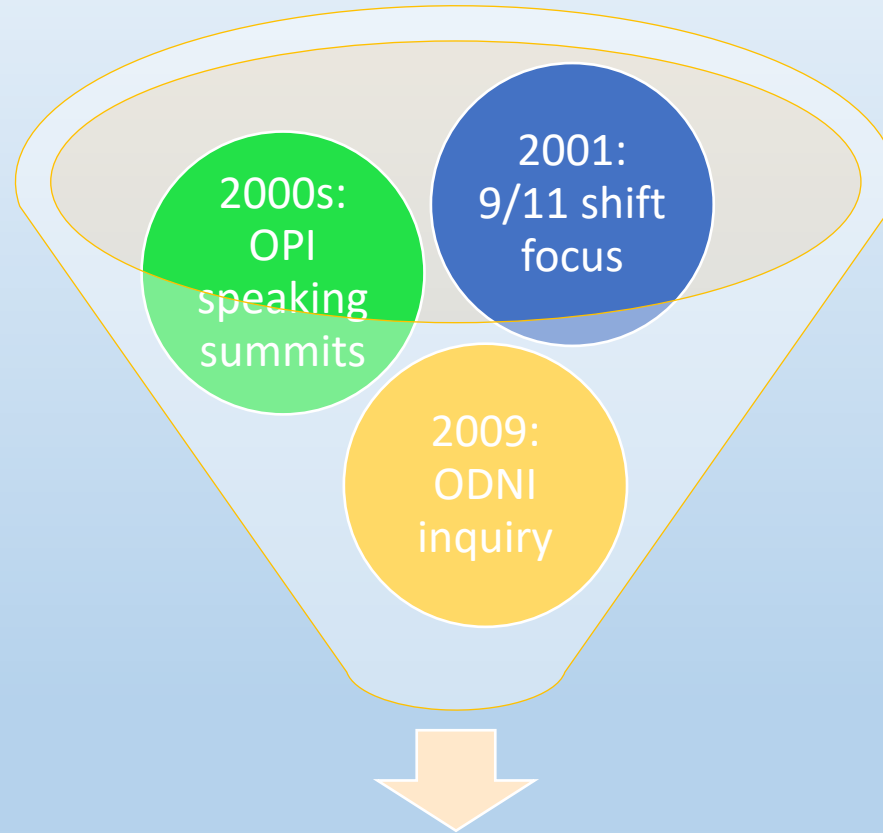
Origins



US Government testing history



US Government testing history



2010: USG Comparability study

ODNI

FLPO

FLEXCOM

TAEG

Study Overview



- Designed in support of Implementation Goal 42 to compare IC language tests for interagency reciprocity
- Led by FBI in 2010-2011
- Three languages were included (Arabic, Chinese, Spanish)
- Three agencies contributed testers
- Three agencies contributed test takers

Purpose of the Study: Research Questions



1. Will an examinee receive the same ILR proficiency level rating when tested at three different agencies?

If yes:

2. How can these differences be characterized?
3. Is it possible to link the different ratings across agencies in order to retain useful information about an examinee's proficiency?

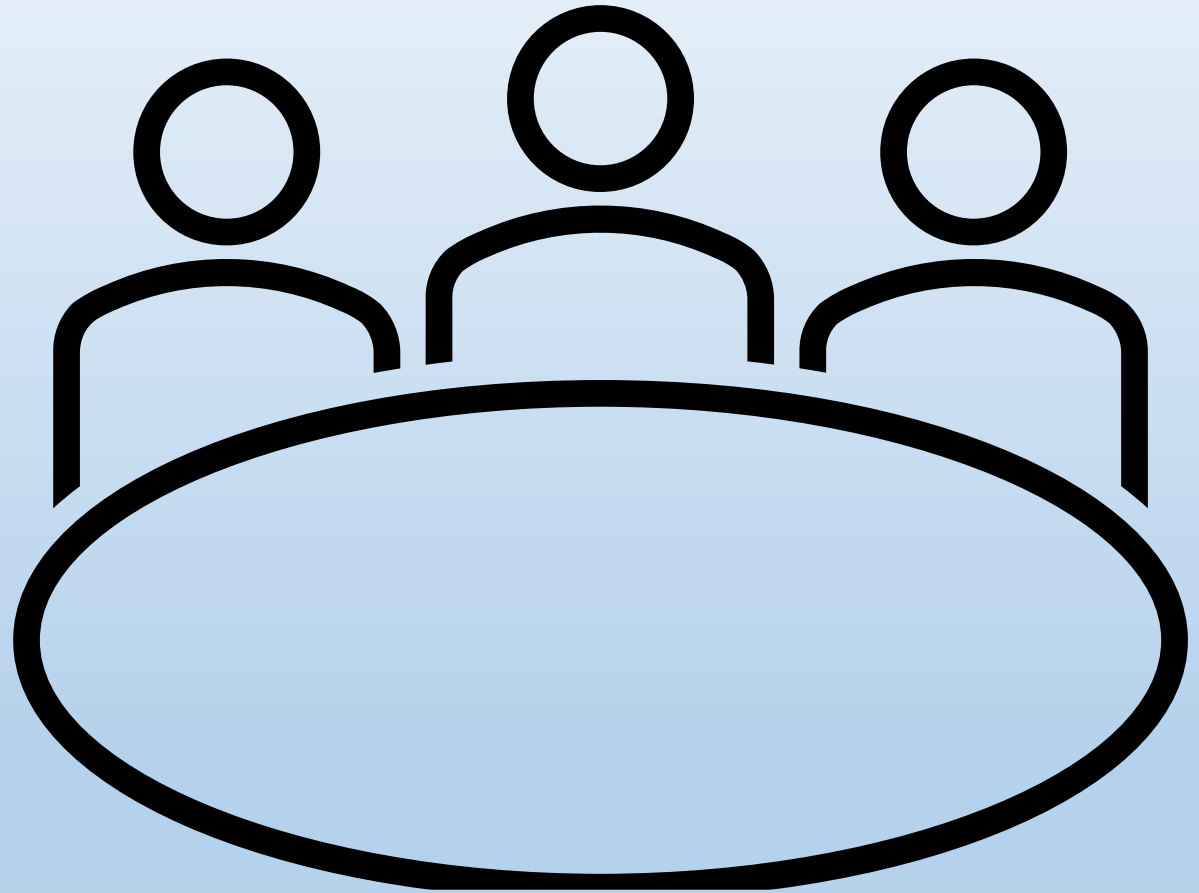
Recommendation leading to a USG workshop



- ...discuss training and rating procedures at the different agencies employing the ILR guidelines for rating purposes, with the aim of establishing best practices for raters and tester trainers.



Meetings





TAEG speaking sub-group: ISS

Participants

- TAEG members with speaking programs
- Other interested TAEG members

Location

- In person
- Host varies

Dates

- First full week of March
- 3-4 days

Interagency Speaking Summits



2015

- Define proficiency
- Protocol or SLD understanding?
- Norming

2016

- Norming
- Interagency task list
- Protocol descriptions
- Role plays

2017

- Updates and norming
- Tester/rater trainings
- Rating categories
- Proficiency vs performance

2018

- Updates and norming
- Plus levels
- Joint training modules
- Tester recertification

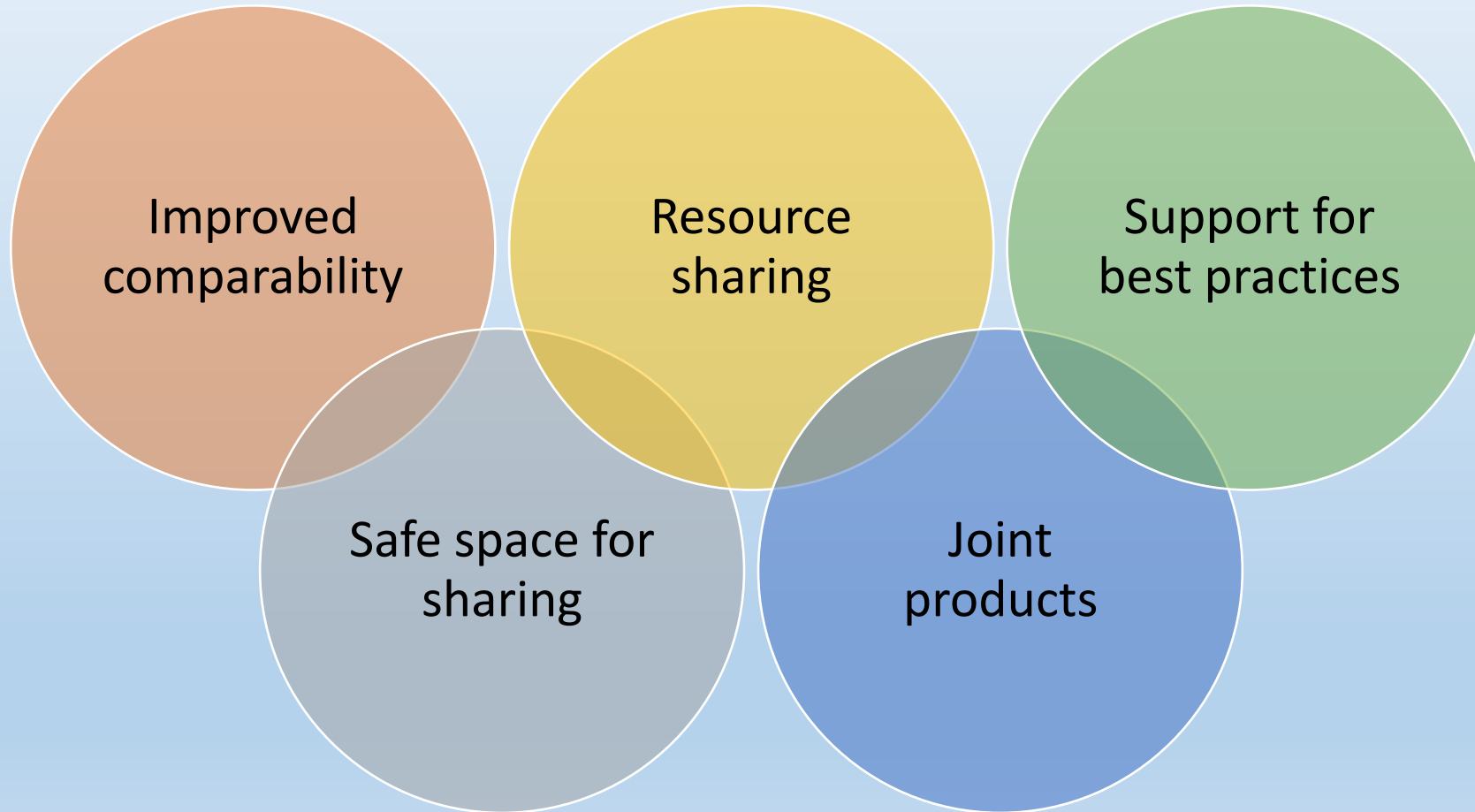
2019

- Updates and norming
- Rating processes
- Quality assurance
- Russian pilot study

2020

- Updates and norming
- Speaking pilot study
- Levels 4 and 5
- ILR revisions

Interagency Speaking Summit Benefits



Products



Interagency task list

Speaking rating categories

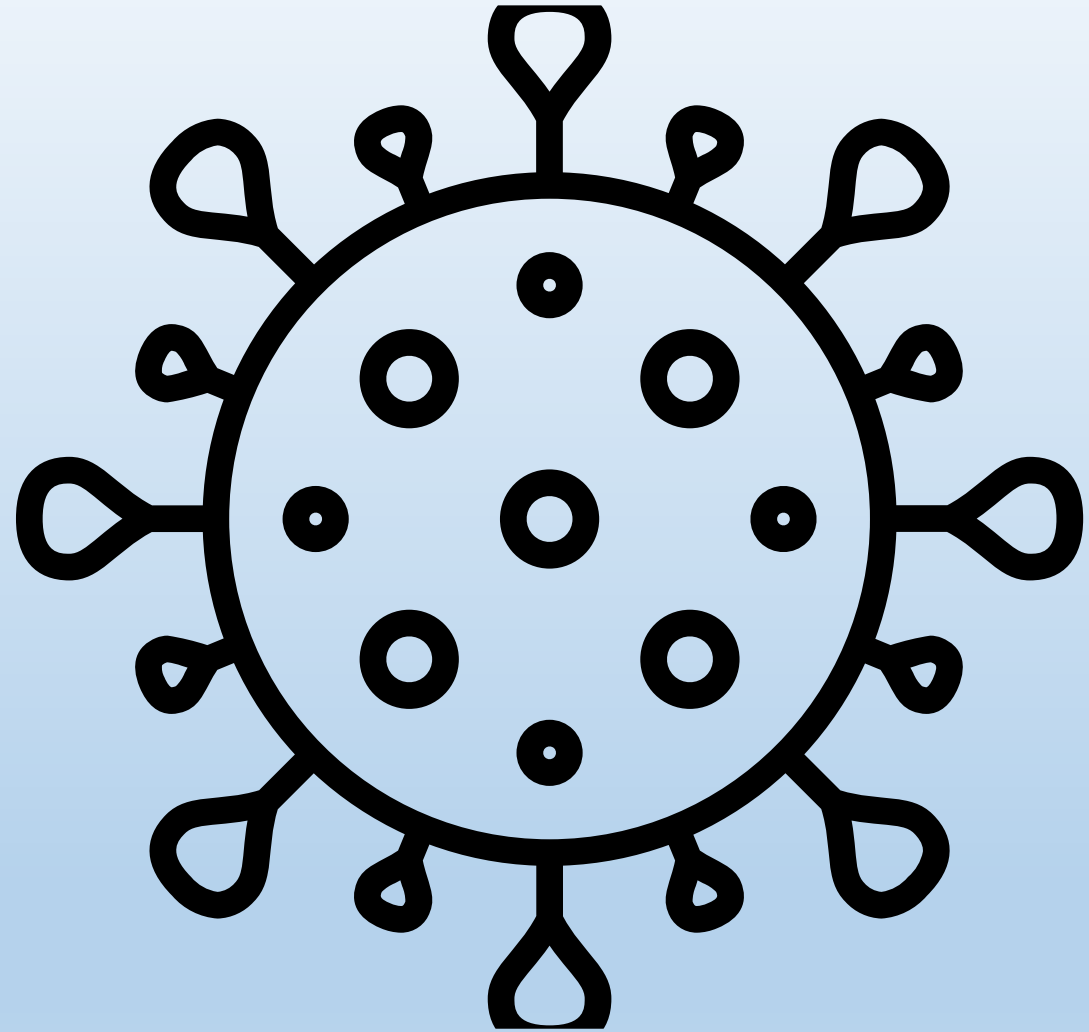
Plus levels statements

Quality assurance best practices white paper

ILR SLD validation framework



2020



Regroup and evolve



Speaking Pilot study

- Move from May to October
- Go virtual

Interagency Speaking Summit 2021

- Move from March 2021 to October 2021
- Go virtual



ILR Speaking Pilot Study

Organizers

- ODNI FLEXCOM TAEG ISS participants in 2018

Purpose

- To help build a validation argument for the ILR SLD speaking revisions.
- To examine whether there is a score shift resulting from the updates of the ILR SLD speaking revisions.

Participants

- (n = 32) testers from CIA, DLIELC, DLIFLC, FBI, FSI

Data

- (n = 860), tests from ILR Levels 0+ - 5

Languages

- English, Mandarin Chinese, Spanish

ILR Speaking Pilot Study Results

Can USG speaking raters reliably identify examinees' levels across the entire ILR SLDs?

Yes! Inter-rater reliability: Weighted Kappa = 0.832



Are USG speaking raters more confident in giving ratings using the revised ILR SLDs for Speaking?

Yes! The rater participants preferred the revisions, saying that they were clearer (92%), more complete (88%) and easier to use (84%).



Do the ILR SLD Speaking abilities and sub-abilities, as outlined in the ILR Crosswalk Matrix, assess separate features of speaking?

Yes! A confirmatory factor analysis showed that the four main abilities fit nicely into model with no need to rate at the sub-ability level



Is there any patterned shift of scores on tests rated on the current scale vs the revised scale?

No! A paired samples t-test showed no significant difference between 1985 SLD scores and the revised SLD scores ($t(859) = -1.66, p = .097$).

ISS 2021



- October 26 – 28, 2021
- Virtual on ZoomGov
- Included international partners
- Over 60 participants



ISS 2021: Agenda

- Organization updates
- Status of ILR SLD adoptions
- Lessons learned from testing during a pandemic
- Tester training modules
- Interagency task list updates
- Training the new SLD abilities
- Organization norming in English, Arabic, and Spanish



ISS 2021: Achievements

- Revised task list
- Hands on practice on how to rate in 3 languages
- Sharing site for documents/training materials, products
- Monthly informal meetings on tester training
- Talked about what is new in each organization
- Strategies for working through COVID and beyond
- Received training on the revised ILR scale

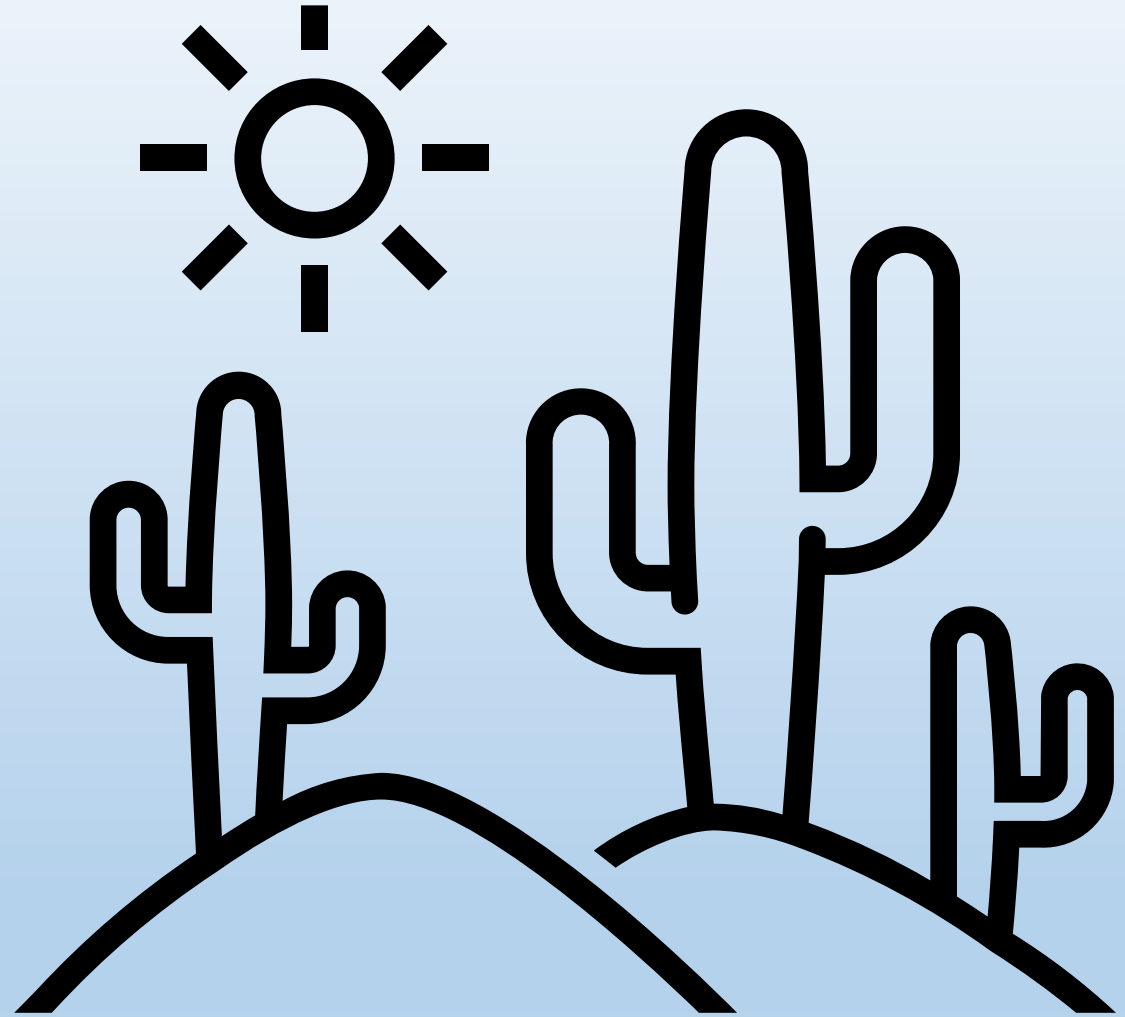


ISS 2021: Parking lot

- How much online/face-to-face proportion is effective in training?
- How do we address mixing of languages in test?
- What is considered a standard language?
- How do we measure appropriateness and acceptability in language tests?
- How do we include the language aspects from other language speaking communities?
- Multilingual competence and translanguaging...
- How do we make sure that we are not testing knowledge cultural knowledge versus culturally appropriate language usage?
- Does notetaking impact examinee performance?
- How can we reduce tester bias?



ISS 2022



ISS 2022



- May 17 - 19, 2022
- Hybrid format: in person and virtual
 - DLIELC: San Antonio, TX

ISS 2022: Agenda



- Organization updates
- Status of ILR SLD adoptions
- Interagency task list updates
- Testing listening skills
- Organization norming in English, Mandarin Chinese, and French
- Contextual appropriateness



ISS 2021: Conclusions

- Meeting together is invaluable to norming and sharing ideas and establishing best practices in a government context
- Virtual-only trainings, at some point face-to-face are important, especially with new testers
- Hybrid models of training seem to be effective and efficient for the future
- Work this year confirms that the conclusion from last year that we should NOT condense levels was correct
- It is important to document our consensus, such as the white paper



Questions?

Thank you!

Building a Validity Argument for the ILR SLDs

