

West, H., Surface, E. A., & Hamlyn, H. (2010, November). *The ACTFL OPic[®] - Using validation evidence for test improvement*. Symposium presented at the ACTFL 2010 Annual Convention and World Languages Expo, Boston, MA.

The ACTFL OPic[®] – Using Validation Evidence for Test Improvement

Hollie West
ACTFL

Eric A. Surface
SWA Consulting Inc.

Helen Hamlyn
LTI



NOVEMBER 2010

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

Copyright Notice

This document and its content is copyright ©1997-2010 of SWA Consulting Inc. All rights reserved.

Any redistribution or reproduction of part, or the entire document in any form is prohibited except for: (1) you may print or download to a local hard disk extracts for your personal and non-commercial use only, and (2) you may copy the content to individual third parties for their personal use, but only if you acknowledge the website and author(s) as the source of the material. You may not, except with our express written permission, distribute or commercially exploit the content, nor may you transmit it or store it on any other website or other form of electronic retrieval system.

ACTFL OPic[®]

Using Validation Evidence for Test Improvement

Hollie West – ACTFL Contract Director
Eric Surface – SWA President
Helen Hamlyn – LTI Vice President

Presentation Outline



ACTFL Oral Proficiency Interview - computer



- OPIC Overview
- Validation Process and Examples from the Field
- Evidence-Based Test Improvements



ACTFL Oral Proficiency Interview - computer



OPIC OVERVIEW

- Internet-delivered assessment of speaking proficiency
- Designed to approximate interpersonal interview and iterative test construct as closely as possible
- Background Survey and Self Assessment
- Content areas and sequencing of tasks is unique for each test taker
- Addresses required functions at each level
- Rated by Certified Raters

OPIc Development Model



ACTFL Oral Proficiency Interview - computer



- 2005 - ACTFL designed and developed the ACTFL Oral Proficiency Interview – computer (OPIc®)
 - Test construct
 - Format / design
 - Delivery Platform
 - Rating Protocol and Systems
- 2006
 - Beta testing/refinement & pilot testing/refinement
 - Rater training & certification
 - Validation & Reliability Study
 - Improvements based on research results
 - Operational launch



ACTFL Oral Proficiency Interview - computer



VALIDATION PROCESS

1. What is validation?
2. Why do you need to validate?
3. How do you validate a test?
4. What decisions do you need to make when choosing a validation process?
5. When is a follow-up validation study needed?
6. Examples from the Field

What is Validation?



ACTFL Oral Proficiency Interview - computer

SWA

Consulting

Validation

Reliability

- Consistency of the test
- Amount of error in the measurements

Validity

- Measures what it should
- Doesn't measure other constructs

What is Validation?



ACTFL Oral Proficiency Interview - computer



Multi-stage process that must
begin on **Day 1** of your test design
and development

Why Do You Need to Validate?



ACTFL Oral Proficiency Interview - computer



- Why should your test be reliable and valid and produce accurate results?
- What problems do you have if the test isn't reliable and/or valid?

Why Do You Need to Validate?



ACTFL Oral Proficiency Interview - computer



Legal

Decision
Making

Professional

Financial

Ethical

How Do You Validate a Test?

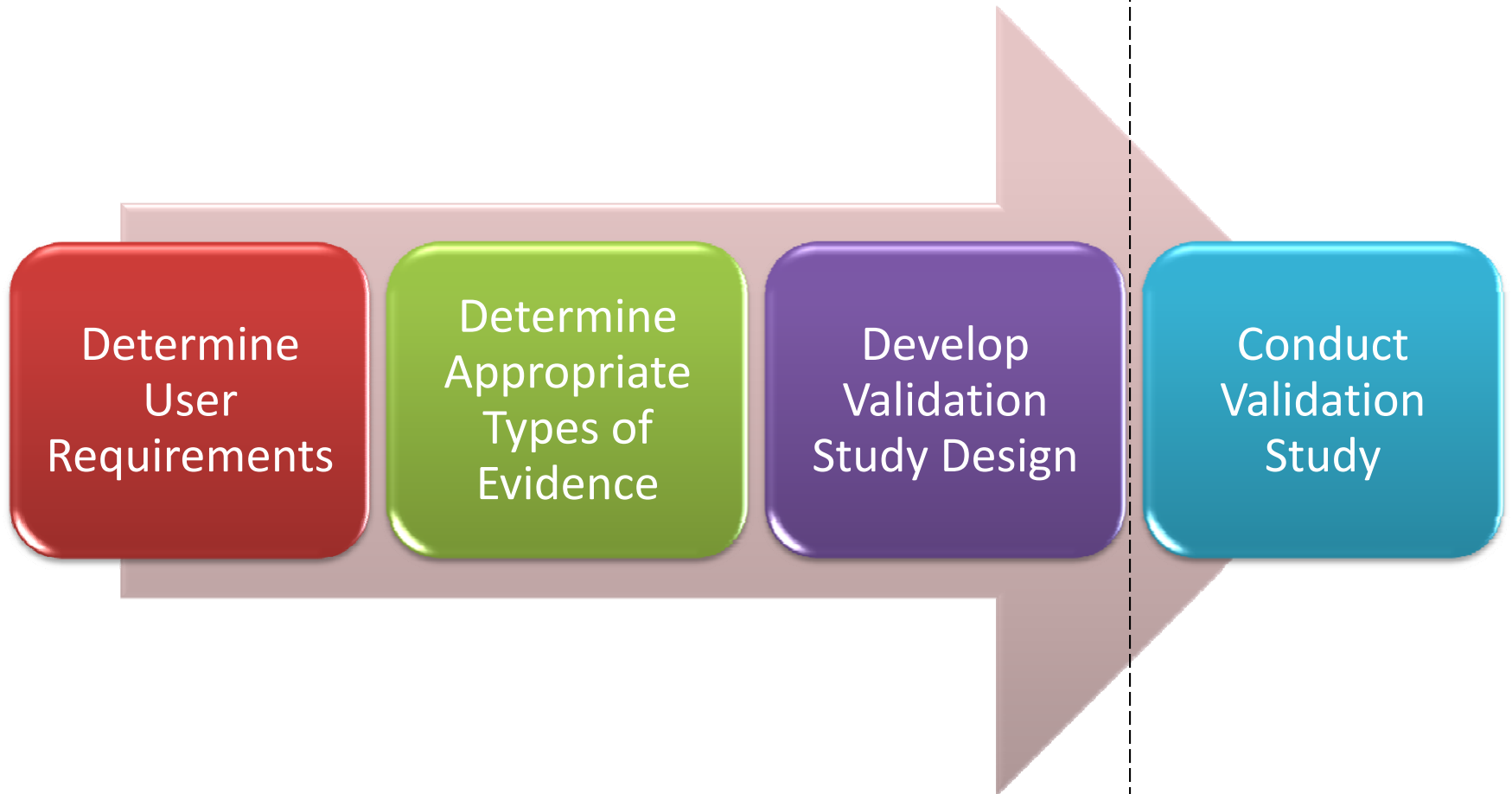


ACTFL Oral Proficiency Interview - computer



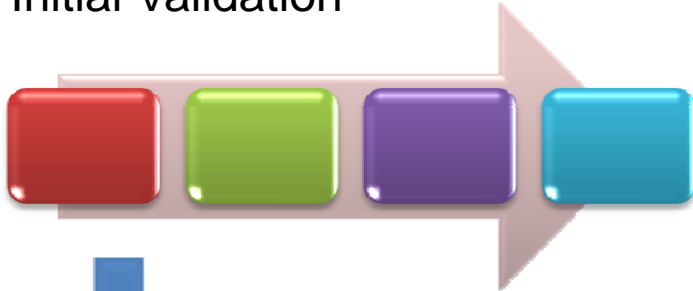
Prior to test development

Test development begins

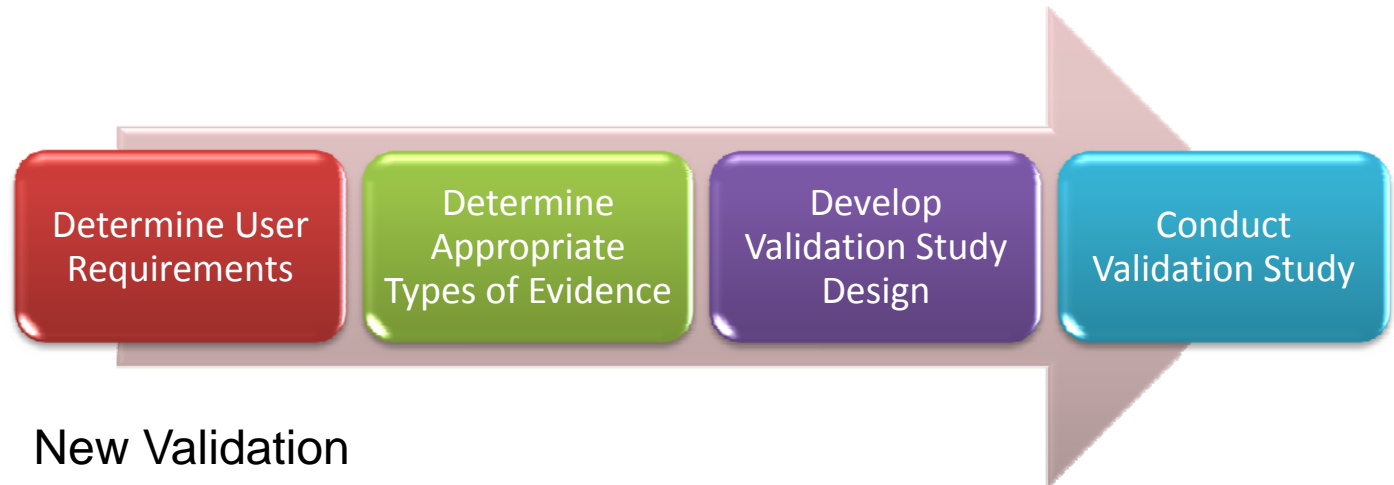


How Do You Validate a Test?

Initial Validation

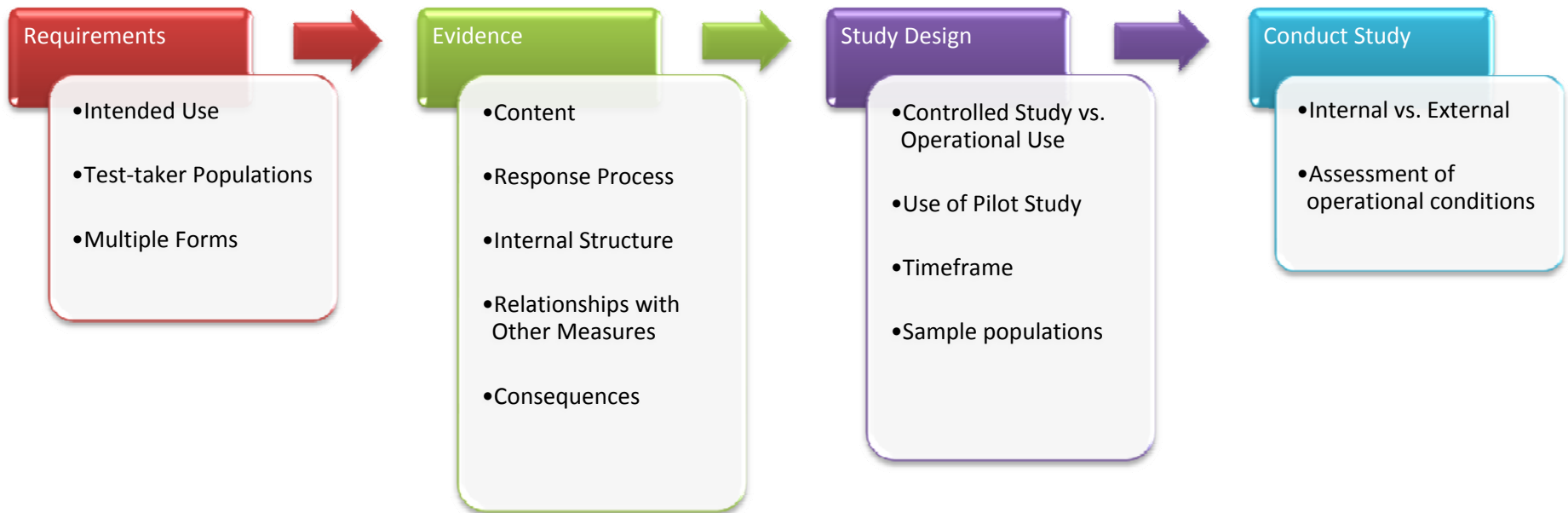


- Changes made to test **OR**
- Test extended to different use **OR**
- Test used in different populations or contexts



New Validation

How Do You Validate a Test?



How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- Professional standards, such as the *Standards for Educational and Psychological Testing (Standards)* and ASTM, provide some guidance on the process
- But, individual professional judgment is needed to choose the specific approach for your test

How Do You Validate a Test?



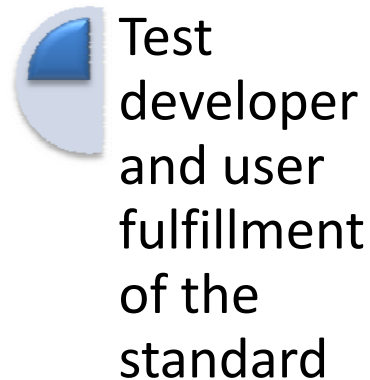
ACTFL Oral Proficiency Interview - computer



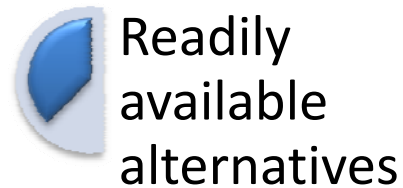
- The *Standards* recommends 4 considerations for evaluating the acceptability of a test:



Professional judgment



Test developer and user fulfillment of the standard



Readily available alternatives



Feasibility of meeting the standard

How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- The *Standards* proposes 5 types of validity evidence
- Keep in mind that validity is the accumulation of evidence that your test measures the intended construct, so these are referred to as types of *validity evidence* rather than types of *validity*

How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- Evidence based on test content
 - Test content:
 - Themes/wording/format of the items/tasks/questions on a test
 - Guidelines for procedures regarding administration and scoring
 - Expert judgment on the match between test content and intended construct

How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- Evidence based on response process
 - Provides evidence of the match between the test construct and the process test takers actually undergo when responding
 - Can be done by:
 - Questioning test takers about their strategies and/or responses to items
 - Documenting physical observations during testing, such as eye movement, response times, or facial expressions
 - Can study both test takers and any personnel involved with delivering or scoring a test (e.g., observer, rater, judge)

How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- Evidence based on internal structure
 - Looking at the relationship between test items/components (e.g., ratings) conform to the test construct
 - Test construct may be either single- or multi-dimensional

How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- Evidence based on relations to other variables
 - Establish relationship between test construct and other tests of similar constructs (*convergent validity*)
 - Establish lack of a relationship between test construct and other tests of different constructs (*discriminant validity*)
 - Establish relationship between test scores and relevant criterion (i.e., outcome variable)

How Do You Validate a Test?



ACTFL Oral Proficiency Interview - computer



- Evidence based on consequences of testing
 - Tests are often used to receive a benefit (e.g., job, placement, award)
 - A fundamental purpose of validation is to indicate whether these specific benefits are likely to be realized
 - Example: Test publisher claims that using their language proficiency test to place students into appropriate course level. Therefore, need evidence of differential benefits to students and the university if alternative placements were used

Consider your test, its intended purpose, and your needs to select the most appropriate types of validity evidence. Remember that validation is a cycle that requires multiple steps.

What Decisions Do You Need to Make When Choosing a Validation Process?



ACTFL Oral Proficiency Interview - computer



- What is the objective of the validation?

Intended Use

- Affects the quality and quantity of validity evidence needed
- High-stakes tests require more validity evidence that meets more rigorous standards than low-stakes tests

Intended Population

- Validations need to be done for each intended population (e.g., students, employees)

What Decisions Do You Need to Make When Choosing a Validation Process?

- Who should do the validation?

Internal Evaluation

- Internal evaluators can apply their inside knowledge of how the test should be functioning to identify problems and possible adjustments

External Evaluation

- External evaluators can provide a more objective examination of the test's reliability and validity

What Decisions Do You Need to Make When Choosing a Validation Process?

- Which study design is most appropriate?

Operational Use

- Results may be more generalizable
- Less control generally increases the amount of error or bias in the results

Controlled Study

- Minimizes amount of error and bias in the results
- Can limit the generalizability of findings

What Decisions Do You Need to Make When Choosing a Validation Process?

- Other Considerations

Resources

- Time
- Money
- Participants

Assessment Types

- Computer vs. Human delivery
- Ratings vs. Multiple-Choice

Logistics

- Data Collection
- Test Scheduling
- Obtaining Appropriate Sample Size

What Decisions Do You Need to Make When Choosing a Validation Process?



ACTFL Oral Proficiency Interview - computer



Need to consider several issues when designing your validation process, including: (1) objective of the validation; (2) who should conduct the validation; and (3) study design.

When is a Follow-up Validation Study Needed?



ACTFL Oral Proficiency Interview - computer



- Test should be re-validated if:
 - Changes are made to the test **OR**
 - The test is extended to different uses **OR**
 - The test is used in different populations or contexts
- Regular follow-up studies should be conducted to monitor the test's reliability and validity

Examples from the Field



ACTFL Oral Proficiency Interview - computer

SWA

Consulting

ACTFL
English OPIC®

ACTFL/ILR
OPIC®

ACTFL English OPIC®

- English
- Korean employees

ACTFL/ILR OPIC®

- Arabic, Chinese Mandarin, French, Korean, Russian, Persian Farsi
- Military and university language students

ACTFL English OPIC®

- Two studies (Study 1 and Study 2)
- Test revisions following 1st study
- Pilot study included in Study 1

ACTFL/ILR OPIC®

- Two studies (Study 1 and Study 2)
- Test revisions following 1st study
- Pilot study included in Study 2

ACTFL English OPIC®

- Study 1: Pre-Survey → OPI/OPIC → OPIC/OPI → Post-Survey → OPIC (for some)
- Study 2: Pre-Survey → OPIC → OPI

ACTFL/ILR OPIC®

- Study 1: Pre-Survey (for some) → OPI/OPIC → OPIC/OPI → Post-Survey
- Study 2 – Military: OPIC → Post-Survey → OPI → Post-Survey
- Study 2 – University: Pre-Survey → OPI/OPIC → OPIC/OPI → Post-Survey

ACTFL English OPIC®

- Reliability: rating consistency, rating concordance, test-retest
- Validity: internal structure, relation with other variables (OPI)
- Test Taker Reactions: provide insight to reliability and validity results

ACTFL/ILR OPIC®

- Reliability: rating consistency, rating concordance
- Validity: internal structure, relation with other variables (OPI)
- Test Taker Reactions: provide insight to reliability and validity results

ACTFL English OPIC®

- Results provided strong reliability and validity evidence
- On-going projects to monitor reliability

ACTFL/ILR OPIC®

- Results coming next year in Denver

EVIDENCE-BASED TEST IMPROVEMENTS


- Changed the User Interface of the ACTFL OPIC[®]
- Program and Delivery Upgrades
- Test Format and Prompt Refinement
- OPIC Rating Protocol and OPIC Rater Training



Research based OPIC Improvements

Changed the user interface of the ACTFL OPIC®

Test page presentation, navigation and Avatar interaction



OPIC System Check. Is your system ready?

[Click here to go to test](#)

Criteria	Requirements	Your System Information	Meets Requirements
			✓ Yes ✗ No ? Unable to Detect
Browser	Windows Internet Explorer 6.0 or FireFox 3.0	Microsoft Internet Explorer 8.0 on Win32	✓
Screen Resolution	1024 x 600 with thousands of colors	1,259 x 535 with millions of colors	✗
JavaScript	JavaScript 1.1	JavaScript 2.0	✓
Bandwidth	100 kilobits per second per tester	2,401 kilobits per second - Test Your Bandwidth	✓
Flash Player	Flash Player 10	Flash Player 10	✓
Microphone	Headset with External Microphone	Start Microphone Test	?

Copyright © 2009 - 2010 American Council on the Teaching of Foreign Languages (ACTFL). All rights reserved.

Research based OPIC Improvements

Reprogrammed OPIC to overcome delivery issues

Added systems' checks throughout the test to insure local computer/internet performance

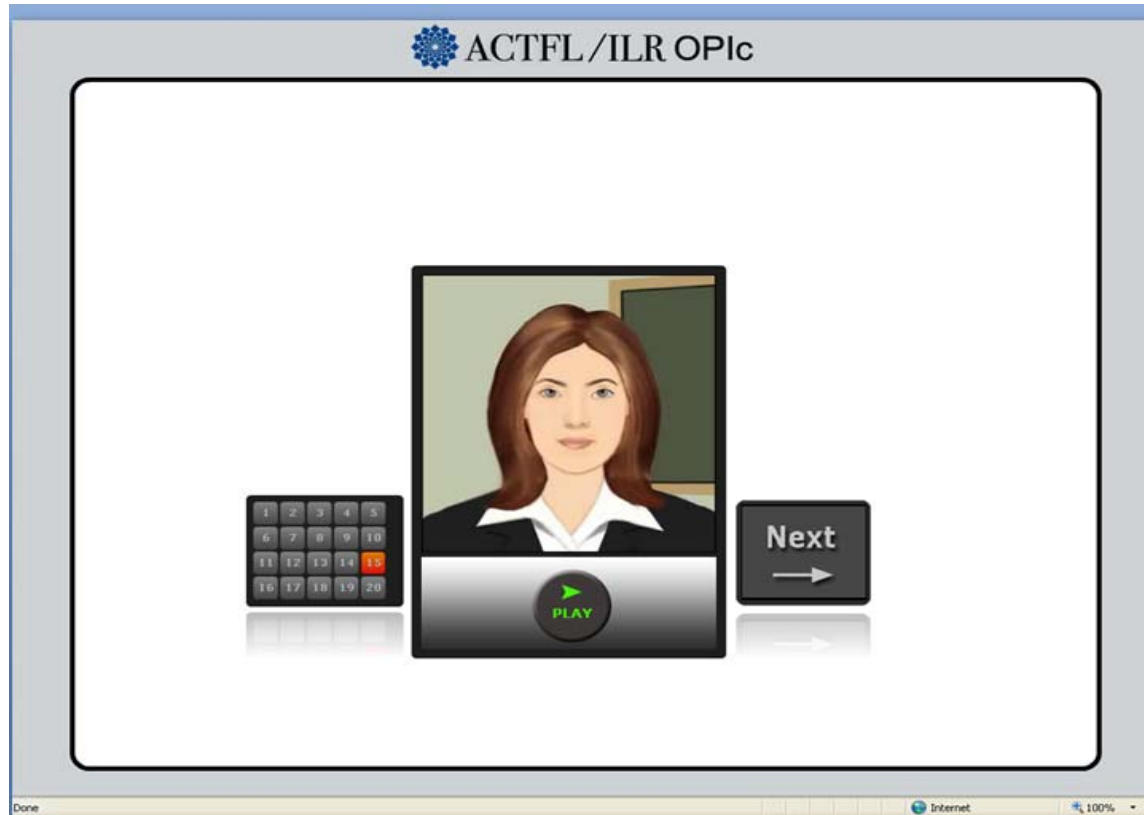
New OPic Products Coming



ACTFL Oral Proficiency Interview - computer



- ACTFL K-12 OPics
- New ACTFL Languages
- ILR Format using an ILR Rating Protocol
- ACTFL OPic and the new ACTFL WPT

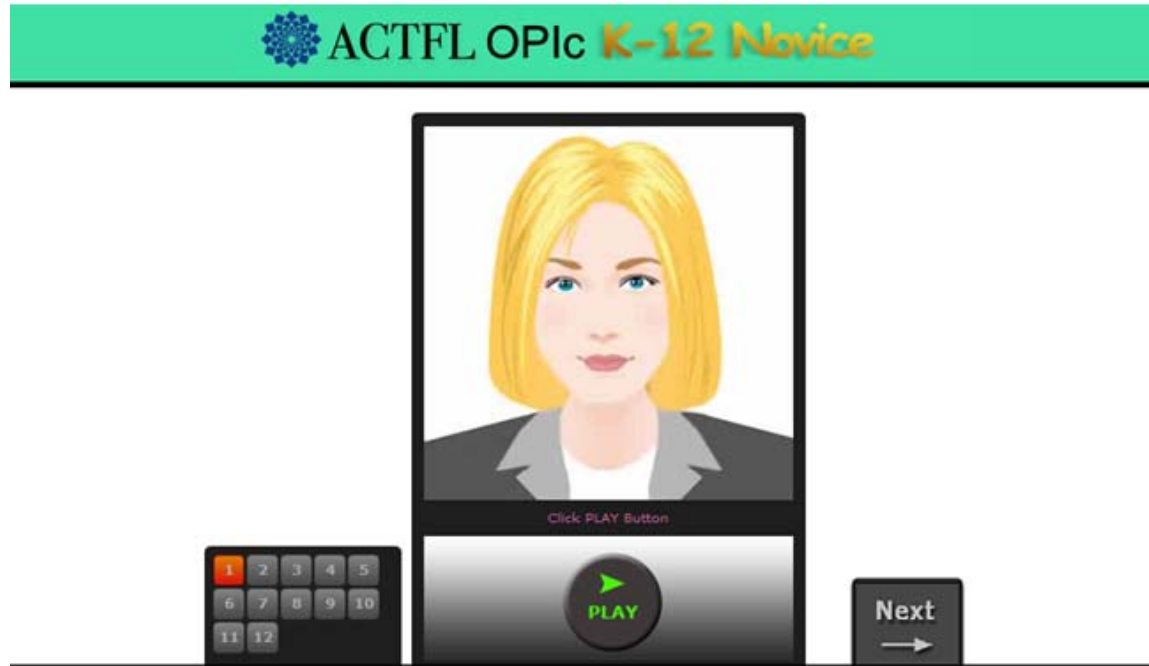


ACTFL ILR OPIC

Format modeled on an ILR OPI. Rated according to ILR protocol

Tests L0+ to L2 range of ILR proficiency

Arabic, French, Mandarin, Persian Farsi and Russian



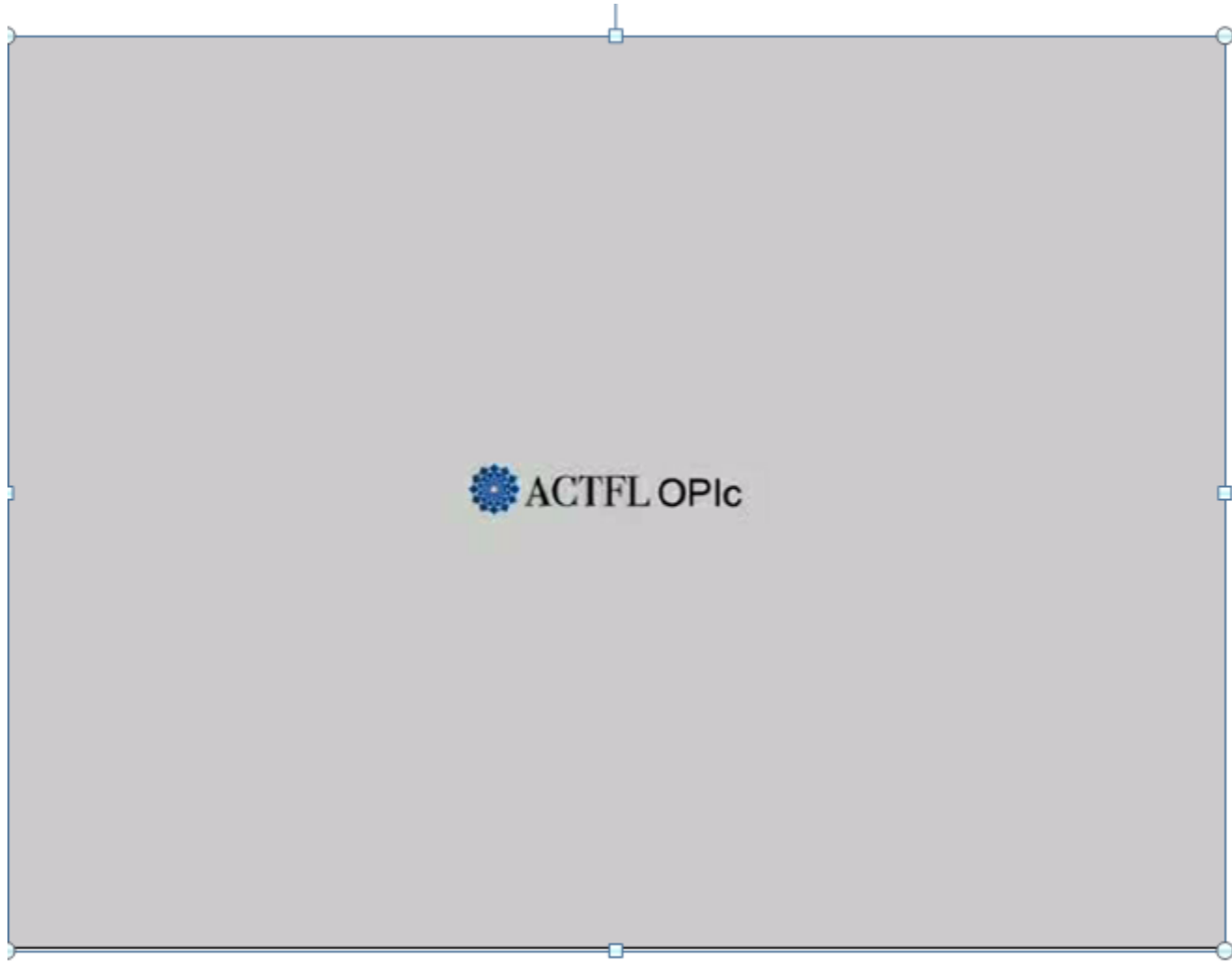
ACTFL K-12 OPIC

Fixed Levels: Novice and Intermediate

Background Survey adjusted to younger learners' experiences and interests

Prompts revised and additional prompts added

English, French, German and Spanish



ABOUT SWA CONSULTING INC.

SWA Consulting Inc. (formerly Surface, Ward, and Associates) provides analytics and evidence-based solutions for clients using the principles and methods of industrial/organizational (I/O) psychology. Since 1997, SWA has advised and assisted corporate, non-profit and governmental clients on:

- Training and development
- Performance measurement and management
- Organizational effectiveness
- Test development and validation
- Program/training evaluation
- Work/job analysis
- Needs assessment
- Selection system design
- Study and analysis related to human capital issues
- Metric development and data collection
- Advanced data analysis

One specific practice area is analytics, research, and consulting on foreign language and culture in work contexts. In this area, SWA has conducted numerous projects, including language assessment validation and psychometric research; evaluations of language training, training tools, and job aids; language and culture focused needs assessments and job analysis; and advanced analysis of language research data.

Based in Raleigh, NC, and led by Drs. Eric A. Surface and Stephen J. Ward, SWA now employs close to twenty I/O professionals at the masters and PhD levels. SWA professionals are committed to providing clients the best data and analysis with which to make solid data-driven decisions. Taking a scientist-practitioner perspective, SWA professionals conduct model-based, evidence-driven research and consulting to provide the best answers and solutions to enhance our clients' mission and business objectives. SWA has competencies in measurement, data collection, analytics, data modeling, systematic reviews, validation, and evaluation.

For more information about SWA, our projects, and our capabilities, please visit our website (www.swa-consulting.com) or contact Dr. Eric A. Surface (esurface@swa-consulting.com) or Dr. Stephen J. Ward (sward@swa-consulting.com).