

Swender, E., Surface, E. A., & Wilcox, S. L. (2010, November). *Identifying effective ACTFL oral proficiency interview-computer OPIC<sup>®</sup> raters: Preliminary findings from ACTFL OPIC<sup>®</sup> training*. Symposium presented at the ACTFL 2010 Annual Convention and World Languages Expo, Boston, MA.

# **Identifying Effective ACTFL Oral Proficiency Interview- Computer OPIC<sup>®</sup> Raters: Preliminary Findings from OPIC<sup>®</sup> Training**

Elvira Swender  
ACTFL

Eric A. Surface  
SWA Consulting Inc.

Sheila L. Wilcox  
SWA Consulting Inc.



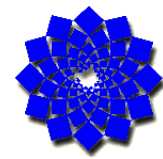
**NOVEMBER 2010**

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.



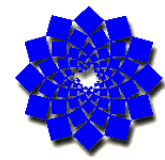
# Identifying Effective ACTFL Oral Proficiency Interview-computer OPIC<sup>®</sup> Raters

## Preliminary Findings from ACTFL OPIC Training

November 19, 2010  
ACTFL Annual Convention 2010

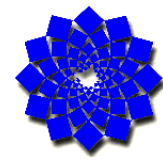
Dr. Elvira Swender  
ACTFL

Dr. Eric A. Surface  
Ms. Sheila Wilcox  
SWA Consulting Inc.



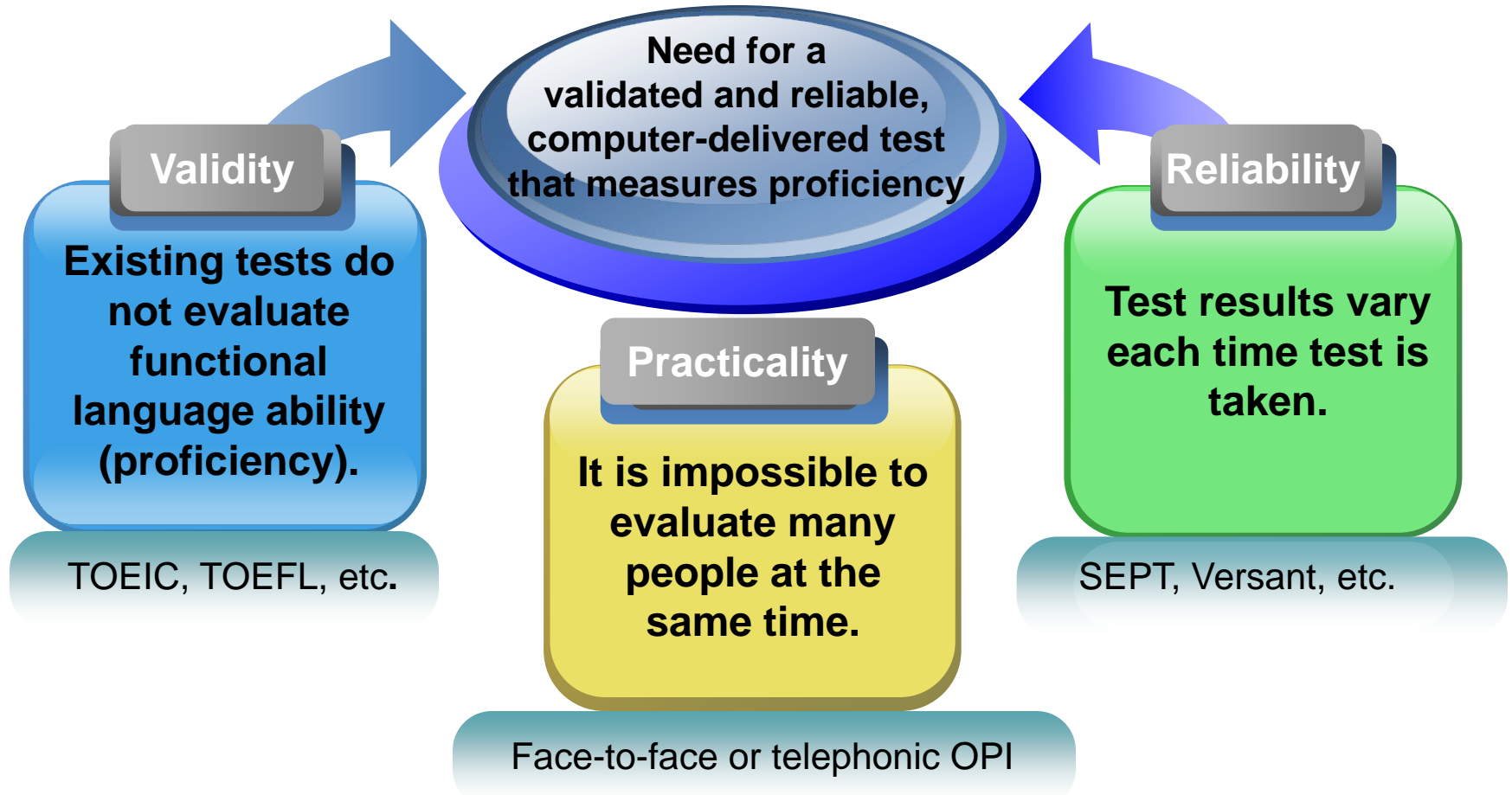
- History and Background of the ACTFL OPic<sup>®</sup>
  - OPic<sup>®</sup> Rater Training and Certification Protocols
  - OPic<sup>®</sup> Raters on the job
  - Rationale for Current Study
- Research Background & Objectives
- Study Methodology
- Findings
- Conclusions
- Future Directions
- Questions and Answers

# Why Develop an OPIc<sup>®</sup>?

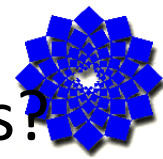


SWA

Consulting



# How is OPIC<sup>®</sup> different from other speaking tests?



## Interview Format

- ❖ Uses the format of the OPI interview
- ❖ Is like a conversation with a live tester
- ❖ Produces a sample of speech that can be rated

**Conversational  
format**

## Adaptive Features

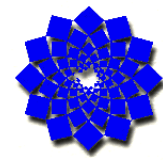
- ❖ Each test is unique
- ❖ Background survey answers determine topic choices
- ❖ Self Assessment determines test level

**Each test is  
individualized**

## ACTFL Certified Rating System

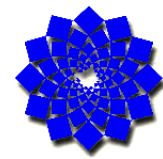
- ❖ ACTFL trained and certified raters
- ❖ Ongoing Quality Assurance

**Secure  
Valid  
Reliable**



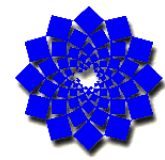
- Internet-delivered assessment of speaking proficiency
- Designed to approximate interpersonal interview and iterative test construct as closely as possible
- Background Survey and Self Assessment
- Content areas and sequencing of tasks
- Addresses required functions at each level
  - Test structure provides follow up on each topic to elicit maximum discourse
- Rated by Certified Raters
  - Rated against same criteria as the OPI

# History of the ACTFL OPic<sup>®</sup>

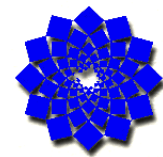


- 2005: ACTFL/LTI create and develop ACTFL English OPic for Korean client
  - January - April 2006: Pilot testing and validation study
  - August 2006: OPic launch / live testing
- Summer 2007: 200 raters recruited and trained
- 2007: Development of ACTFL Spanish OPic
- September 2007: DLI contract to develop ILR OPic in 7 languages
  - Arabic, Chinese Mandarin, Korean, Persian, French, Russian, Bengali
- October 2008: DLI contract continuation to develop
  - Pashto, Tagalog, and Indonesian
- 2008: Development of ACTFL Jr. OPic in English for Korea
- 2008: ACTFL OPic through Superior
- Development of ACTFL OPic K-12 Novice / K-12 Intermediate in English and Spanish for U.S.
- ACTFL German OPic in production
- More than 300,000 OPics conducted to-date

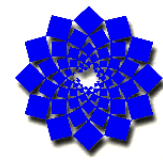
# Certified OPIc Raters



- Training
  - Face-to-face training
- Online rating practice and certification
  - Practice Round – get official ratings and rationales
  - Certification Round
- Quality Assurance procedure
  - Q/A trainers manage rater training and certification
- Must meet criteria for certification
  - Demonstrated ability to consistently rate reliably

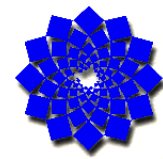


- Native/Superior level speakers of language
- Language professionals
  - Teachers, linguists, translators, etc.
- Independent contractors for ACTFL Testing Office (LTI)
  - 350 OPIC<sup>®</sup> Raters
    - All protocols, all languages
  - 250 English Raters – ACTFL Protocol
- Must maintain high inter-rater reliability

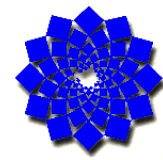


- Test volume activity is variable
  - Depends on test volume and rater availability
  - Surges and dry spells
- Flexibility of “work” hours
  - No scheduled rating times
- Can work from any quiet location with high speed Internet access
- Rater Support
  - Enhancement site
  - Live Trainer support
  - Full time OPIc Rater Q/C Manager
- Interesting and challenging
  - No two tests are identical
  - No two speaker profiles are identical
  - It’s amazing what people will say when communicating with an Avatar

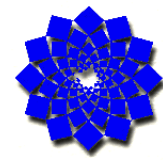
# Rationale for Current Study



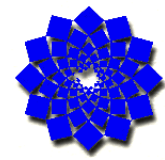
- ACTFL is interested in continuous improvement of its training and certification process
  - *For whom is the training effective?*
- Impact on training design
  - *How can training be optimized for different learners?*
- Impact on recruiting successful trainees
  - *Critical for meeting the demand for certified raters in the profession*
- Impact on selecting successful trainees for OPIc and other testing programs
  - *ILR Tester training*



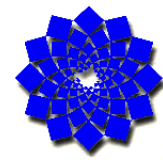
# Research Background and Objectives



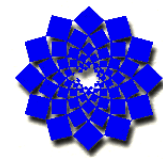
- The effectiveness of any rater-based assessment depends on the extent to which all the raters share a standardized ***mental model*** of proficiency and a standardized protocol for rating that proficiency and apply them consistently.
  - ***Having an effective training and certification process is paramount to having an effective assessment.***
- Previous Research has investigated OPI Tester Training.
  - Dierdorff, Surface & Brown (2010), *Journal of Applied Psychology*
  - This is a different training and job context to investigate
  - Comprehensive program of research on raters and testers



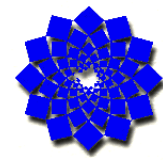
- To become a successful OPIC rater for ACTFL, there is a gated process.
- Multiple hurdles must be navigated:
  - Training
  - Application for certification
  - Certification
  - Decision to work as a rater
  - Performance working as a rater
- Who are the people who best navigate the process?



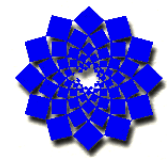
- What factors are related to certification outcomes?
  - What types of people become certified? Decide to work as a rater? Perform well on-the-job?
    - More focus on performance as an OPIc Rater
  - What factors predict:
    - *Obtaining certification?*
    - *Deciding to work as a rater?*
    - *Performance on-the-job?*



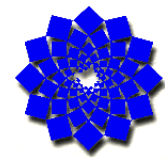
- Individual Differences
  - Motivation, Personality, Demographics, Cognitive Ability, Cognitive Complexity
- Training Outcomes
  - Performance during Certification Process
- *The goal of training is to standardize raters on these training outcomes criteria, so there may be no variability (i.e., No prediction)*



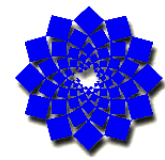
- Training conducted in summer 2007
- What SWA Collected:
  - Pre-training survey
- What ACTFL Collected:
  - Performance during certification
  - Final certification status
- What LTI Collected:
  - Which certified raters went to work
  - Performance on-the-job (August 2007 – September 2008)
  - Current rater status



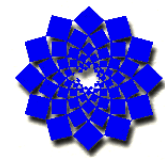
- Transfer (e.g., Baldwin & Ford, 1988):
  - Near Transfer
    - Becoming a certified rater
    - Performance during certification process
  - Far Transfer
    - Performance working as a rater



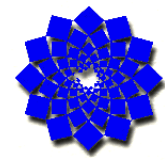
- Pre-Training Questionnaire
  - $N = 173$ 
    - Caution because of low sample size
  - Background with foreign languages, teaching, computers, taking proficiency tests, and rating proficiency tests
  - Why attended training
  - Task-Specific Self-Efficacy (OP1c)
  - Psychological
    - Goal Orientation, Learning Self-Efficacy, General Self-Efficacy, Personality, Core Self-Evaluation
  - Wonderlic
  - Cognitive Complexity



- Participant Demographics:
  - Highest educational degree obtained
    - BA/BS = 28%, MA/MS = 59%, Ph.D./Ed.D. = 6%, Other = 7%
  - Have you served as an ACTFL OPI rater?
    - Yes = 2%, No = 98%
  - Have you ever served as a rater for another assessment of language proficiency?
    - Yes = 29%, No = 71%
  - Native language
    - English = 97%, Spanish = 1%, Other = 2%

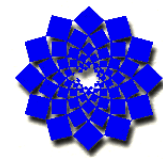


- Participant Demographics (cont.)
  - Occupation
    - Professor = 16%, Other = 23%, K-12 Teacher = 52%, Student = 9%, Researcher = 1%
  - Why in Training
    - Flexibility that job offers = 44%
    - Professional/Career Development = 38%
    - Extra Income = 21%
    - Personal Interest = 48%
    - Aligns with Current Skill Set and Previous Experience = 40%
    - Help Others with Language Proficiency = 6%
    - Interest/Belief in ACTFL Mission/Goals/Assessments = 9%
    - Activity for Retirement = 5%
    - Technology Aspect to Language Teaching/Assessment = 3%

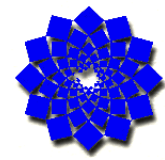


- **Pre-Training Predictors:**

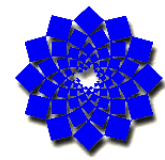
- Previously Assessed with OPI; Previously Taken Another Assessment of Language Proficiency
- Is Highest Degree Related to Foreign Language Teaching or Research
- Have You Ever Taught a Language Course
- Served as Rater for Other Language Proficiency Test
- Years Using Computer (Computer Experience); Internet Use at Work; Internet Use at Home
- Why in Training: Flexibility that Job Offers; Professional/Career Development; Extra Income; Personal Interest; Aligns with Current Skill Set and Previous Experience; Improve as a Teacher



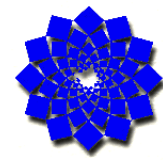
- **Pre-Training Predictors (cont.):**
  - Personality: Extraversion; Agreeableness; Neuroticism; Openness to Experience; Conscientiousness
  - Goal Orientation: Learning; Prove; Avoidance
  - Core Self-Evaluation
  - General Self-Efficacy; Learning Self-Efficacy; Task-Specific (OPIc) Self-Efficacy: Learning; Applying; Results
  - Wonderlic
  - Cognitive Complexity



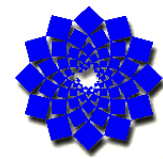
- Certification Criterion:
  - Becoming a Certified Rater
- Certification Predictors:
  - Agreement During Practice Round 1
  - Agreement During Rounds 2-4
  - Times to Criterion
  - Agreement for each Proficiency Level
  - Overall Agreement



- Job Criteria:
  - Deciding to Work as Rater
  - Overall Number of Ratings
  - Overall Agreement
  - Current Status

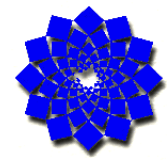


# Preliminary Findings

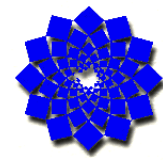


# Preliminary Findings

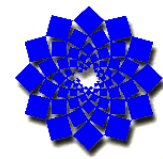
Who completes certification?



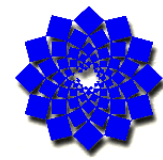
- Participants who took pre-training survey: 173
  - Certified: 129 (75%)
  - Recommended for Retraining: 39 (23%)
  - Incomplete (Didn't finish certification process): 4 (2%)



- Significant ( $N = 138$ )
  - Why in Training: Personal Interest ( $r = 0.18$ )
    - If this was a reason why attended training, then did become certified
  - Openness to Experience ( $r = 0.19$ )
    - Trainees higher on openness to experience became certified
  - Task-Specific (OPIc) Self-Efficacy: Applying ( $r = -0.19$ )
    - Trainees higher on OPIc SE Applying (confident they can use the skills from training to assign ratings) did not become certified

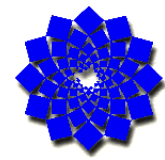


- Approaching Significance ( $N = 138$ )
  - Why in Training: Professional/Career Development ( $r = -0.16$ )
    - If attended training for professional career development, then did not become certified
  - Wonderlic ( $r = 0.16$ )
    - Trainees who scored higher on Wonderlic became certified
  - Years Using Computers ( $t = -1.95$ )
    - Trainees who became certified had fewer years of experience using computers
  - Cognitive Complexity ( $t = -1.81$ )
    - Trainees who became certified more cognitively complex

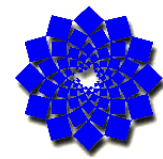


# Findings

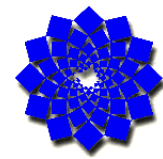
Who Goes to Work as a Rater?



- Certified Raters = 129
  - Worked as Rater = 123 (95%)

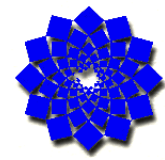


- Significant
- Approaching Significance ( $N = 97$ )
  - Previously Assessed with OPI ( $r = 0.18$ )
    - Certified raters who had not been assessed with an OPI went to work as a rater
  - Why in Training: Improve as Teacher ( $r = -0.18$ )
    - Certified raters who attended training to improve as a teacher did not go to work as a rater

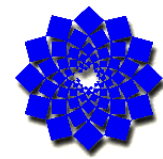


# Findings

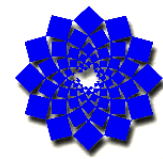
Who does well on-the-job?



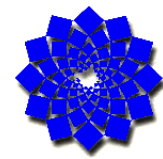
- Multiple indicators of performance on-the-job:
  - Overall Number of Ratings
  - Overall Agreement
  - Current Status



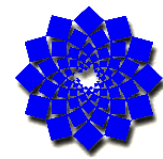
- Overall Number of Ratings ( $N = 91$ )
  - Significant
    - Previously taken another assessment of language proficiency ( $r = -0.22$ )
      - Raters who had taken another proficiency test rated more interviews on-the-job
    - Task-Specific (OPIc) Self-Efficacy (Results) ( $r = 0.21$ )
      - Raters higher on OPIc Self-Efficacy: Results (more confident they can become certified rater and work as a rater) rated more interviews on-the-job
    - Wonderlic ( $r = 0.24$ )
      - Raters who scored higher on Wonderlic rated more interviews on-the-job



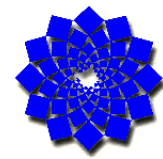
- Overall Number of Ratings (cont.)
  - Approaching Significance
    - Agreement for Intermediate High interviews during Certification ( $r = 0.18$ )
      - Raters who more accurately rated intermediate high interviews during certification rated more interviews on-the-job



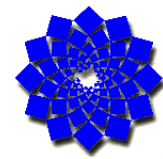
- Overall Agreement ( $N = 91$ )
  - Significant
    - Previously assessed with OPI ( $r = -0.22$ )
      - Raters who had previously taken an OPI had higher agreement on-the-job
    - Why in Training: Professional/Career Development ( $r = 0.36$ )
      - Raters who attended training for professional/career development had higher agreement on-the-job
    - Core Self-Evaluation ( $r = 0.21$ )
      - Raters higher on core self-evaluation had higher agreement on-the-job



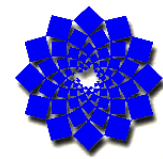
- Overall Agreement (cont.) ( $N = 91$ )
  - Significant
    - Agreement during Certification Round 1 ( $r = 0.24$ )
      - Raters who had higher agreement during round 1 of certification had higher agreement on-the-job
    - Agreement on Novice Mid interviews during Certification ( $r = 0.31$ )
      - Raters who had higher agreement on NM interviews during certification had higher agreement on-the-job
    - Agreement on Intermediate Mid interviews during Certification ( $r = 0.26$ )
      - Raters who had higher agreement on IM interviews during certification had higher agreement on-the-job
    - Agreement on Intermediate High interviews during Certification ( $r = 0.22$ )
      - Raters who had higher agreement on IH interviews during certification had higher agreement on-the-job
    - Overall Agreement during Certification ( $r = 0.24$ )
      - Raters who had higher agreement during certification had higher agreement on-the-job



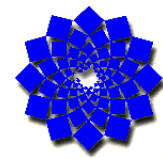
- Overall Agreement (cont.) ( $N = 91$ )
  - Approaching Significance
    - Years Using Computers ( $r = 0.20$ )
      - Raters who have more experience using computers had higher agreement on-the-job
    - Extraversion ( $r = 0.20$ )
      - Raters higher on extraversion had higher agreement on-the-job



- Current Status
  - Overall, 142 certified raters from Summer 2007 training events
    - Not all took our survey
  - Active Rater for LTI: 57 (40%)
  - Rater Dropped Out: 81 (57%)
  - LTI Dropped Rater: 2 (1%)
  - Not Currently Active, but Not Dropped Out: 2 (1%)



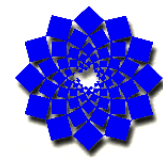
- Current Status
  - Significant
    - Goal Orientation: Learning ( $t = -2.00$ )
      - Raters who dropped out had higher means than Active raters
    - Overall Agreement On-the-Job ( $t = 2.74$ )
      - Raters who dropped out had lower agreement than Active raters
    - Number of Ratings On-the-Job ( $t = 6.47$ )
      - Raters who dropped out made fewer ratings than Active raters
  - Approaching Significance
    - None



Findings

Summary

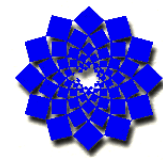
	Certification	On-the-Job
Previously Assessed with OPI		Significant Approaching Sig.
Previously Taken Other Lang. Prof. Assessment		Significant
Years Using Computers	Approaching Sig.	Approaching Sig.
Why in Training: Personal Interest	Significant	
Why in Training: Improve as Teacher		Approaching Sig.
Why in Training: Prof./Career Development	Approaching Sig.	Significant
Core Self-Evaluation		Significant
Task-Specific (OPIc) SE (Results)		Significant
Task-Specific (OPIc) SE (Applying)	Significant	
Openness to Experience	Significant	
Extraversion		Approaching Sig.
Goal Orientation: Learning		Significant
Cognitive Complexity	Approaching Sig.	
Wonderlic	Approaching Sig.	Significant
Agreement on NM during Cert.		Significant
Agreement on IM during Cert.		Significant
Agreement for IH during Cert.		Significant Approaching Sig.
Agreement during Cert. Round 1		Significant
Overall Agreement during Cert.		Significant



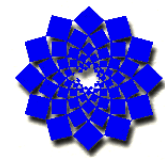
SWA

Consulting

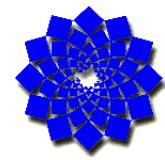
# Conclusions and Future Directions



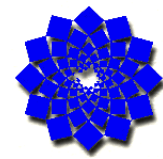
- Several individual difference factors appear to impact rater performance during certification and on-the-job, including:
  - Background experiences
  - Motivation for Attending Training
  - Psychological (e.g., personality, goal orientation, self-efficacy)
  - Cognitive Ability
- More research needed



- Performance during certification related to performance on-the-job
- For this training event, many trainees became certified and nearly all went to work as raters for LTI
  - However, within 3 years of training, over half had dropped out
  - But, those who dropped out were ratings fewer interviews and less accurately than raters who are still active



- Continue monitoring on-the-job performance for these raters
  - Research Question: What predicts long-term performance?
- Continue analyzing data for these participants
  - Add additional performance data
  - More sophisticated modeling
- Conduct similar study including participants trained under current training/certification process
  - Looking for any differences with current process



# Questions and Answers

Please feel free to contact  
Dr. Swender or Dr. Surface  
with additional questions.

## **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](http://www.swa-consulting.com)) or contact Dr. Eric A. Surface ([esurface@swa-consulting.com](mailto:esurface@swa-consulting.com)) or Dr. Stephen J. Ward ([sward@swa-consulting.com](mailto:sward@swa-consulting.com)).