



**TOEFL**

*Listening. Learning. Leading.*

# TOEFL<sup>®</sup> iBT Program Update

**Kathy Pruner**  
**Associate Director**  
**TOEFL Client Relations**

# One Test – Different Formats

- The TOEFL® Internet-based test (iBT) – 95% of all TOEFL testing
  - Offered in 1,436 cities in 139 countries
  - Administered in 4,471 certified test centers by trained personnel
- The TOEFL Paper-based test (PBT) - used in countries where iBT is not yet available and to supplement iBT in a limited number of areas
- The TSE® (*Test of Spoken English™*) - available upon request in PBT locations
- The TOEFL Computer-based test (CBT) ended in September 2006 – scores not reported after September 2008.
- All TOEFL scores are valid for two years

The TOEFL® Test — Quality Beyond Measure



# Unmatched Quality

## Responding to Your Needs

- Authentic – tasks simulate university classroom and life; developed in collaboration with universities
- Integrated skill questions emphasize ability to use English for communication
- 100% academic – based on extensive research
- Highly reliable - 0.94
- Pricing
  - Differential pricing allows testing in poorest countries below ETS's costs
  - Less expensive than IELTS in almost all countries

The TOEFL® Test — Quality Beyond Measure



# Unmatched Quality Focus on Scoring

- Unbiased, objective scoring
  - Expert raters score Speaking and Writing responses through Online Scoring Network (OSN)
    - Do not see or know test takers
    - Rate responses from around the world
    - 3 – 6 different raters score Speaking tasks
    - 2 different raters score each of 2 Writing tasks
- Raters tested and monitored
  - After initial training, take calibration test each time they begin scoring
  - Continually monitored by Scoring Leaders
  - Scoring Leaders can contact Content Scoring Leaders, who are in contact with assessment experts at ETS

The TOEFL® Test — Quality Beyond Measure

# TOEFL iBT 2007 Total Population Scaled Section and Total Score Means

	iBT Mean
Reading	19.1
Listening	19.8
Speaking	19.2
Writing	20.2
Total	78.0

The TOEFL® Test — Quality Beyond Measure

# TOEFL® iBT 2007

## Total Population Score Means by Test Taker Purpose

Group	Reading	Listening	Speaking	Writing	Total
<b>Graduate</b>					
Mean	20.6	21.0	19.8	21.0	82.0
<b>Undergraduate</b>					
Mean	17.3	19.3	19.2	19.7	76.0
<b>Professional</b>					
Mean	20.6	21.9	21.0	21.0	84.0

The TOEFL® Test — Quality Beyond Measure

# TOEFL® iBT 2007

## Total Population Score Means by Sample Countries

Native Country	Reading	Listening	Speaking	Writing	Total
Korea (ROK)	20	20	18	20	77
Japan	16	16	15	18	65
China	21	19	18	20	78
India	20	21	21	22	84
Taiwan	17	18	18	19	72
Turkey	19	20	19	20	77
Germany	23	26	24	24	97
Italy	22	19	18	19	78
France	23	23	21	22	88

The TOEFL® Test — Quality Beyond Measure

# Using TOEFL iBT Speaking to Screen International Teaching Assistants

- Study conducted with 4 universities to determine if TOEFL iBT Speaking can serve as a pre-arrival initial screening measure for ITAs
  - Drexel University (Drexel)
  - University of North Carolina, Charlotte (UNCC)
  - University of California, Los Angeles (UCLA)
  - University of Florida at Gainesville (UF)
- Study goals:
  - Determine the association between scores on TOEFL iBT Speaking and on local ITA tests
  - Set TOEFL iBT Speaking score requirements for ITA screening for these institutions

The TOEFL® Test — Quality Beyond Measure

# Using TOEFL iBT Speaking to Screen International Teaching Assistants

- Success of ITAs depends on numerous factors
- What do you intend to measure with tests?
  - English proficiency
  - Teaching ability
  - American culture
- Local test rubrics varied at each institution
- TOEFL iBT Speaking would provide information on English proficiency

The TOEFL® Test — Quality Beyond Measure



# Trade-off of Different Classification Errors

- Students classified into non-pass, provisional pass, clear pass
- Study cut scores established by schools were high to eliminate false positives
  - False positive: unqualified TAs predicted as qualified
    - May have more serious impact from the perspective of undergraduate education
    - No easy way to rectify false positives
    - Difficult to reverse assignments and pull funding
  - False negative: qualified TAs predicted as unqualified
    - Can retest in future

The TOEFL® Test — Quality Beyond Measure

# Correlations Between TOEFL iBT Speaking and Local ITA Test Scores

	UCLA TOP Test	UNCC non-content presentation test	Drexel DIP Test	UNCC Content-based Presentation Test	UF Teach Evaluation
TOEFL iBT Speaking	.78 <b>.84</b>	.78 <b>.93</b>	.70 <b>NA</b>	.53 <b>.58</b>	.44 <b>.72</b>

Disattenuated\* correlations are in bold face numbers (lower row)

\*Corrected for unreliability in the scores for each test

The TOEFL® Test — Quality Beyond Measure

# Major Findings

- Relationships between TOEFL iBT Speaking test and local ITA test scores moderately strong
  - Depends on the extent to which the local test engages and evaluates non-language abilities
  - TOEFL Speaking scores more strongly related to linguistic aspects measured by the local ITA tests
- Results provide guidance on using TOEFL iBT Speaking scores for ITA screening purposes and setting cut scores
- TOEFL Speaking scores
  - Significant predictors of success when used for TA assignments

The TOEFL® Test — Quality Beyond Measure

# TOEFL Information

- TOEFL website: [www.ets.org/toefl](http://www.ets.org/toefl)
- Write to us at: [TOEFLNews@ets.org](mailto:TOEFLNews@ets.org)
- TOEFL iBT Help Line: 1-609-683-2008
- Sign up to be a test center: [www.ets.org/ctas](http://www.ets.org/ctas)
- Contact Client Relations:
  - Terry Axe: [taxe@ets.org](mailto:taxe@ets.org)
  - Kathy Pruner: [kpruner@ets.org](mailto:kpruner@ets.org)
  - Eileen Tyson: [etyson@ets.org](mailto:etyson@ets.org)

The TOEFL<sup>®</sup> Test — Quality Beyond Measure

