



TOEFL

Quality Beyond Measure.

Don't Sing the Blues. Get the English Language Proficiency Scores You Can Use

Kathy Pruner

Associate Director, Global Client Relations

ETS

ETS — *Listening. Learning. Leading.*[®]





Introduction

Today we'll talk about how the scoring of English language proficiency tests can impact the reliability of the scores



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Overview

Human Scoring

- Centralized Scoring & Consistent Scoring Standards
- Multiple Raters
- Quality Control Measures for Raters
 - Certification and Calibration
 - Monitoring

Automated Scoring

- Balanced Use of Human and Automated Scoring

Services For Score Users and Teacher



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]





Colorless green ideas sleep furiously.



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Have you considered...

- Whether the scoring is kept separate and distinct from the test administration process?



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Why is this important?

- Ensures the security and integrity of scores
- Provides fair, objective scoring



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]

Centralized Scoring & Consistent Scoring Standards – What the TOEFL® Test Does

- Scoring is done centrally for security and score integrity,.
- The TOEFL program's Online Scoring Network (OSN) ensures consistent and unbiased scoring standards
 - Raters see and hear worldwide responses, not just from one country
 - Raters do not know the examinee
 - Scoring is anonymous
- Because of distribution of responses among raters, rater does not know anything about the rest of the test taker's performance, avoiding the "halo effect".



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

Have you considered.....

- Whether more than one rating contributes to the score?



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Why is this important?

- Using multiple raters increases score reliability
- Using multiple raters minimizes rater bias

Multiple Raters – What the TOEFL® Test Does

- The TOEFL program uses multiple raters to score each test
- Writing
 - Independent writing task:
 - one human rater + one automated rating
 - Integrated writing task:
 - two human raters
 - future plans – one human rater + one automated rating
- Speaking
 - Six tasks
 - Three to six human raters



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

Multiple Raters – What the TOEFL® Test Does

- The Online Scoring Network (OSN) organizes all test taker responses by question type after the test.
- OSN tracks which raters have rated which test takers.
- Responses are distributed for scoring
 - A maximum number of times a rater can rate responses from the same candidate is set



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

Have You Considered...

- Whether human raters are calibrated and monitored frequently and consistently?



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Why Is This Important?

- Provides quality control to ensure accurate, consistent scoring



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Quality Control Measures for Raters – Certification and Calibration – What the TOEFL® Test Does

- Online training and certification
- Raters must pass calibration test each day or with each new question type in order to score responses
 - Includes task familiarization, guidance on scoring the task and practice on a range of responses.
- Must review exemplar responses following a break to refocus attention



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

Quality Control Measure for Raters – Monitoring – What the TOEFL® Test Does

- Continually monitored by Scoring Leaders
- Scoring Leaders > Content Scoring Leaders > ETS Expert
- Real time monitoring of raters online on a daily basis for score accuracy and consistency



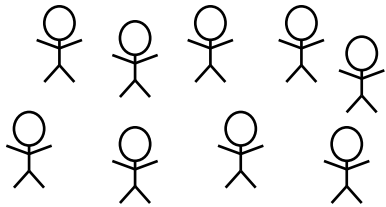
Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

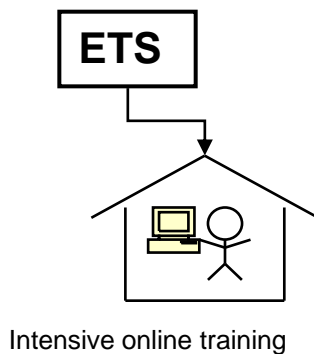


How Scoring for the TOEFL® Test Works

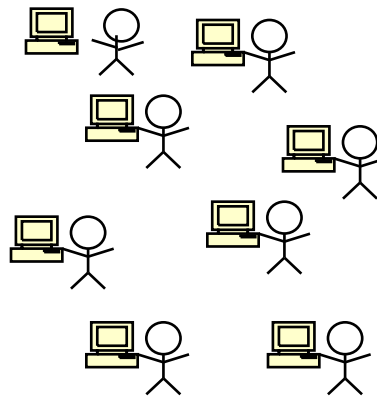
1. Apply



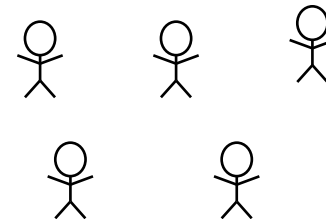
2. Training



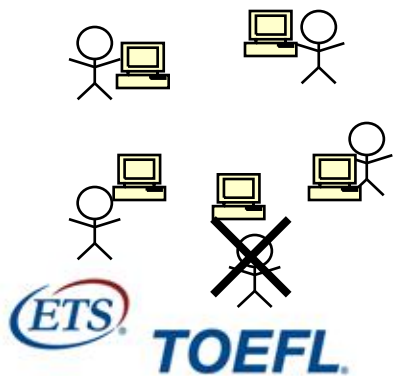
3. Test



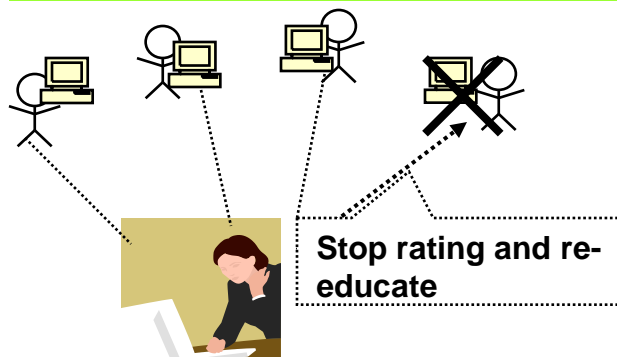
4. Certified as a rater



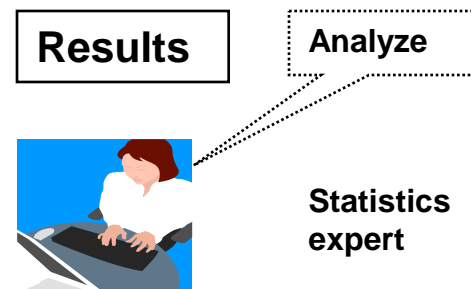
5. Calibration Test



6. Monitoring by scoring leader



7. Analyze results



Quality Beyond Measure.

Have You Considered...

- Whether scoring is done by human raters, an automated process, or both?



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Why Is This Important?

- A combination of rating methods offers a complete and accurate picture of an applicant's ability
- Automated scoring alone measures the mechanical aspects of writing sufficiently, but content is measured only at a basic level



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]

ETS is the World's Leader in Automated Scoring

- ETS automated scoring capabilities
 - *e-rater*[®] Scoring Engine: essays
 - *SpeechRater*SM Scoring Engine: spontaneous speech
 - *c-rater*TM Scoring Engine: correctness of short responses
 - *m-rater*SM Scoring Engine: mathematical expressions, figures and plots



Quality Beyond Measure.

ETS — Listening. Learning. Leading.[®]



ETS is the World's Leader in Automated Scoring

How is e-rater[®] used by ETS?

A few examples:

- High Stakes (in combination with human rater)
 - The GRE[®] Test
 - The TOEFL[®] Test
- Placement & Practice (automated scoring only)
 - TOEFL[®] Practice Online
 - SAT[®] Online
- Learning (automated scoring with feedback)
 - *Criterion*[®] Service



Quality Beyond Measure.

ETS — Listening. Learning. Leading.[®]

ETS is the World's Leader in Automated Scoring

- Increases scoring efficiency and reduces cost
- Increases score reliability when used in conjunction with human scorers
- Offers potential for diagnostic feedback
- Does an excellent job of evaluating linguistic features of language use, such as
 - Grammar
 - Structure
 - Vocabulary
 - Fluency
 - Pronunciation

but.....



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



...there is more to **communicative** academic English proficiency at the university level than the linguistic aspects of language that computers can currently analyze



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]





Colorless green ideas sleep furiously.



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Balanced Use of Human and Automated Scoring:

What the TOEFL® Test Does

- TOEFL combines the judgment of humans for content and meaning and the consistency of automated scoring for linguistic aspects of language



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®



Balanced Use of Human and Automated Scoring – What the TOEFL® Test Does

- TOEFL Writing section has two tasks:
 - Independent writing task – 1 human + 1 automated rating
 - Integrated writing task - 2 human raters
 - Will use e-rater® in future, with 1 human +1 automated rating
- Extensively researched by ETS and reviewed by external TOEFL Committee of Examiners



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

Balanced Use of Human and Automated Scoring – What the TOEFL® Test Does

- TOEFL integrated questions require students to combine skills and formulate responses to summarize, evaluate, compare and synthesize information from multiple sources
 - Automated scoring alone cannot assess the coherence and logic of language demanded for university study



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®



Balanced Use of Human and Automated Scoring – What the TOEFL® Test Does

- Combines human and automated scoring
 - Does not put constraints on what we want to measure
 - Does not limit test content to simple, structured test questions that can be accurately recognized and readily machine scored
 - Does not put the cart before the horse



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®

Balanced Use of Human and Automated Scoring – What the TOEFL® Test Does

- Combines human and automated scoring
 - Questions that use only automated scoring do not promote the practice of real communication skills
 - Tests that use automated scoring alone are
 - More coachable
 - Highly vulnerable to test gaming



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®





TOEFL

Quality Beyond Measure.

Comparison of TOEFL and IELTS Total Scores

ARCHITECTURE CANADA BUSINESS LAW ENGINEERING KOREA
 GERMANY MEDICINE NURSING COMPUTER SCIENCE JAPAN
 BANGLADESH ART POLITICAL SCIENCE AUSTRALIA PHYSICS

INDONESIA UNITED STATES UK MATHEMATICS INDIA FINANCE TEACHING CAMBODIA
 COMMUNICATIONS HONG KONG SINGAPORE ECONOMICS FRANCE ITALY

ETS — *Listening. Learning. Leading.*[®]



Why ETS Created a Comparison Table

- Concordance tables and similar documents are published in various media
- Some of the information on these tables is inaccurate and misleading
- Universities have asked for a way to compare TOEFL iBT™ and IELTS™ scores

Research to Support the Comparison Table

- ETS collected more than 1,100 sets of TOEFL iBT™ and IELTS™ scores from individuals who had taken both tests
- An equipercentile approach was applied to each IELTS band score to develop the comparison table
 - Ex: a TOEFL iBT score of 115 represents the 90th percentile of the test takers in the sample, the same as an IELTS score of 8.5

TOEFL iBT™ and IELTS™ Score Comparison

TOEFL iBT	IELTS
35	5.0
46	5.5
60	6.0
79	6.5
94	7.0
102	7.5
110	8.0
115	8.5



Quality Beyond Measure.

Research to Support Using the Table

- The comparison table was created as an initial step for universities that wanted to compare TOEFL iBT™ and IELTS™ scores
- ETS recommends the table be used for purposes of general understanding and not for establishing cut scores
- It is not appropriate to use the comparison table to project an applicant's IELTS or TOEFL iBT score
- Follow up validity studies are recommended to validate the score comparisons



TOEFL

Quality Beyond Measure.

Feedback from Institutions

ARCHITECTURE CANADA BUSINESS LAW ENGINEERING KOREA
 GERMANY MEDICINE NURSING COMPUTER SCIENCE JAPAN
 BANGLADESH INDIA

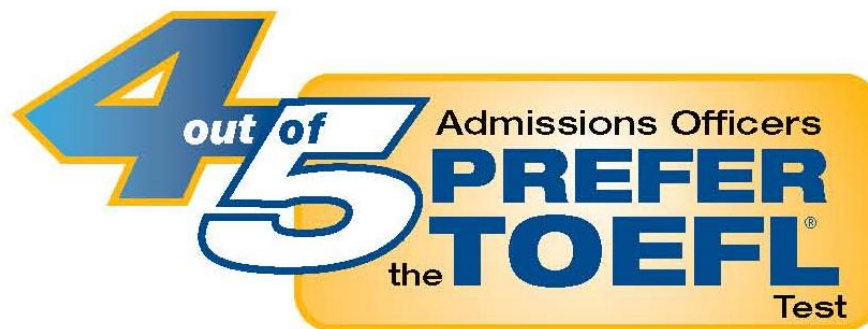
INDONESIA UNITED STATES UK MATHEMATICS INDIA FINANCE TEACHING CAMBODIA
 COMMUNICATIONS HONG KONG SINGAPORE ECONOMICS FRANCE ITALY

ETS — *Listening. Learning. Leading.*[®]



What Do Institutions Think?

- 4 out of 5 admissions officers, surveyed and who stated a preference, prefer the TOEFL test over other tests



Source: Survey of 263 admission officers at U.S. universities, of which 212 accept both the TOEFL test and the IELTS™ test and 152 state a preference



Quality Beyond Measure.

ETS — Listening. Learning. Leading.®



What Do Intensive English Language Professionals Think?

- Does the best job in assessing the English-language proficiency of their applicants
- More accurate scores
- More available and accessible to test takers
- Provides better customer service to institutions
- Provides more updates



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



See How the TOEFL Test Measures Up To Your Standards

- To find out more download our publication “How to Select the Best Academic English-language Test” at www.ets.org/toefl/trust



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]





TOEFL

Quality Beyond Measure.



ARCHITECTURE CANADA BUSINESS LAW ENGINEERING KOREA
GERMANY MEDICINE NURSING COMPUTER SCIENCE JAPAN
BANG

Services for Score Users and Teachers



ETS — *Listening. Learning. Leading.*[®]



INDONESIA UNITED STATES UK MATHEMATICS INDIA FINANCE TEACHING CAMBODIA
COMMUNICATIONS HONG KONG SINGAPORE ECONOMICS FRANCE ITALY

Promote Your Institution

- Record a brief video testimonial about your institution and the TOEFL® test at www.ets.org/toefl/video
- Your video will be viewed by more than a million students each month — valuable exposure at no charge

The TOEFL Channel on the YouTube® website

(www.youtube.com/TOEFLtv) The TOEFL Program and YouTube invite teachers and students to send short videos with tips on:

- Learning English
- Preparing for the TOEFL test
- Studying abroad



Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*®



Resources for Universities



EducationUSA — U.S. Department of State-affiliated educational advising centers in more than 400 overseas locations

- Reach prospective international students through this well established global network (no fees)
- Gain access to a variety of services:
 - Collaborate with Regional Educational Advising Coordinators (REACs)
 - Send your university information to advising centers for distribution or weekly e-newsletter
 - Participate in higher education fairs
 - Link your international admissions website to EducationUSA's to direct prospective students to their centers for on-the-ground, in-country support for U.S. institutions



<http://www.educationusa.info/>

Quality Beyond Measure.

ETS — *Listening. Learning. Leading.*[®]



Resources for Teachers and Learners

Teacher/classroom resources

- **Free teacher webinars:** helps you build communicative activities
- **Propell™ Manual for the TOEFL iBT™ Test:** helps you build an curriculum based on TOEFL iBT learning objectives
- **Criterion® Online Writing Evaluation Service –**
www.ets.org/criterion/ell
 - 35 independent essay prompts for TOEFL test practice
 - Instructor can also create their own prompts
 - Student-submitted essays are scored instantly and students are provided feedback so they can revise and resubmit their essays
- **Academic Connections™** (Pearson Longman)
New! – instruction and assessment program that focuses on critical academic skills



Quality Beyond Measure.

ETS — Listening. Learning. Leading.®



Resources for Teachers and Learners

- **Lexile Measures** www.Lexile.com/TOEFL
 - Match **TOEFL iBT™** reading scores to **Lexile®** measures
 - Find books of interest at the appropriate reading level
- **TOEFL® Practice Online** (available at the ETS Store www.ets.org/store)
 - Simulates TOEFL iBT testing experience
 - Provides instant scores and performance feedback
- **The Official Guide to the TOEFL® Test, 3rd Edition**
 - Paperback book and CD or e-book formats with test preparation tips, strategies and sample questions
 - Institutional discounts: orders of 50 or more receive 50% discount(contact TOEFLNews@ets.org)
- **Pronunciation in English; Idioms in English; and Writing in English** (available at the ETS Store www.ets.org/store)
 - Interactive multimedia programs



Quality Beyond Measure.

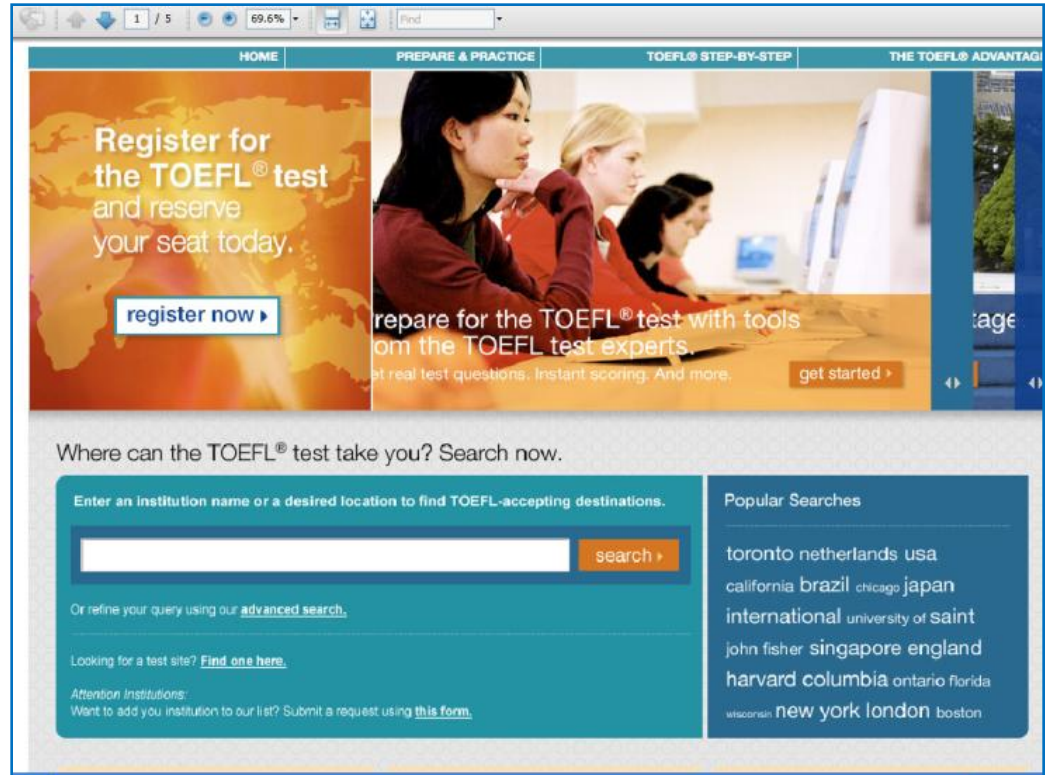
ETS — *Listening. Learning. Leading.*®



TOEFL® Go Anywhere Website

www.toeflgoanywhere.org

- TOEFL test information
- Sample test questions
- Study tips
- Test registration
- TOEFL Destination Directory



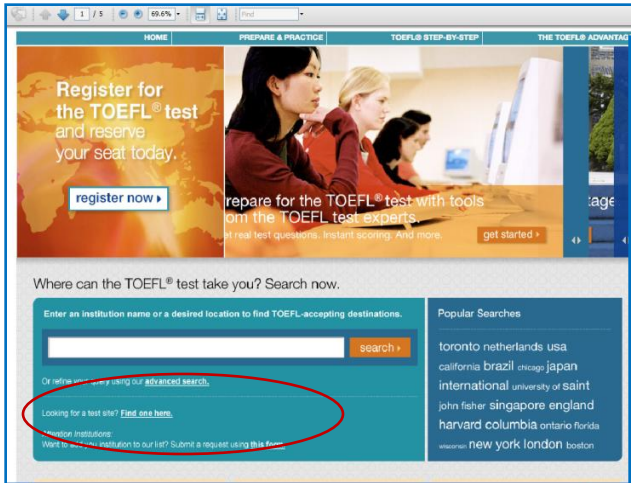
Quality Beyond Measure.

ETS — Listening. Learning. Leading.®



The Advanced TOEFL® Search

- Advanced search provides an institution's score requirements and program information



The Advanced TOEFL® Search

Use the detailed search features below to find institutions that accept the TOEFL® test.

Institution Name

Test Type

TOEFL iBT 25

TOEFL PBT 25

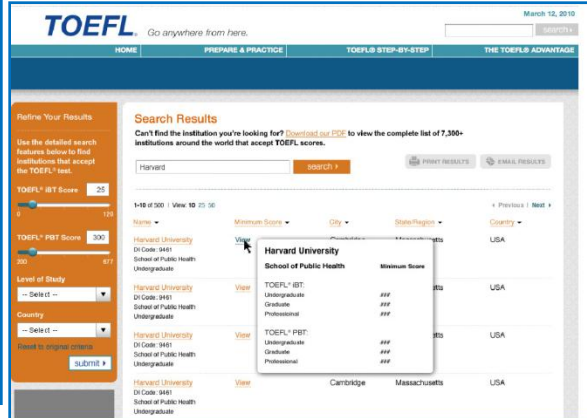
The TOEFL® PBT Test is not offered in all areas.

Level of Study

Country

State/Region

City



Quality Beyond Measure.

ETS — Listening. Learning. Leading.®



We're Here to Help

- TOEFL Website: www.ets.org/toefl
- Write to us at: TOEFLNews@ets.org
- TOEFL Help Line: +1-609-683-2008

- Kathy Pruner – kpruner@ets.org



Quality Beyond Measure.

ETS — Listening. Learning. Leading.®

