Student-Centered Language Assessment: Lowering Barriers and Increasing Opportunities

Ryan Munson, Lindsay Addington, and Marivi Walker
Agenda

- Introductions
- Duolingo English Test (Overview)
- Duolingo English Test (Validity & Reliability)
- Graduate Data (Trends from 2021)
- Institutional Perspective: Florida Tech
- Use in Practice (Students & Institutions)
- Q&A
Introductions
Presenters

Ryan Munson
United States Engagement
ryan@duolingo.com

Marivi Walker
Florida Tech
mvwalker@fit.edu

Lindsay Addington
United States Engagement
lindsay@duolingo.com
Use technology to lower barriers and increase opportunities for English language learners everywhere.
duolingo english test overview

Accessible - Test has been taken in 200+ countries and costs $49 USD

Comprehensive - Includes scores, video interview, and writing sample

Fast - Certified results available in 2 days

Reliable - Research supports use in high-stakes decisions

Secure - Remote proctoring, AI, browser lockdown, adaptive test, and more
Access Program 2020-21

70+ countries
150+ counselors / orgs
3,000+ direct use
10,000+ offered to universities

englishtest.duolingo.com/access
NEW! Duolingo English Test partnership with UNHCR

On World Refugee Day, the Duolingo English Test announced our latest access initiative—hiring a dedicated university counselor to directly support UNHCR as they work with refugees seeking higher education. And today, we’re proud to begin the search for the individual who will lead in this role.

Together, we’ll seek to reinforce UNHCR’s 19by30 initiative, which aims to increase refugee enrollment in higher education to 15% by 2030. Currently, only 3% of all refugees ever enroll at a university. This initiative and this counselor’s future work will be transformative in creating a framework for refugees’ access to higher education—advising 25 refugee students from targeted regions of the world, guiding them through the university admissions and financial aid process.

We’re reaching out with hope that you or your team is connected with a passionate university counselor who may be interested in applying to lead this critical initiative.

More information as well as the application can be found at go.duolingo.com/AccessProgramCounselor

We look forward to continuing to work with universities as we launch this partnership, and are ever grateful for your support.

The Duolingo English Test team

Questions from you or your applicants about the Duolingo English Test? Reach out to your Duolingo contact or email Institutional@duolingo.com
International acceptance

3,000+ U.S. programs
200+ Canadian programs
150+ UK programs
50+ Irish programs
30+ Australian programs
What we’re seeing

+2,000
New accepting institutions in 2020. Over 3,500 programs now accept the Duolingo English Test

+2,000%
YTD increase in test volume. Large increases coming from China, Korea, Brazil, Mexico, India, Kazakhstan
test centers per million people
duolingo test takers per million people
Validity & Reliability
More researchers, more research

Dr. Alina von Davier  
Chief of Assessment  
- PhD in Mathematics from Otto von Guericke University Magdeberg  
- President of the International Association of Computer Adaptive Testing  
- Board Director, Assoc. of Test Publishers

Dr. Antony John Kunnan  
Principal Assessment Scientist  
- PhD in Applied Linguistics, UCLA  
- Honorary Professor, Univ. of Hong Kong  
- Author of Evaluating Language Assessments (Routledge, 2018), co-editor of Language Assessment Quarterly

Dr. Jill Burstein  
Principal Assessment Scientist  
- PdD in Linguistics, CUNY  
- Previously developed automated writing evaluation technologies; conducted writing analytics/achievement research  
- Specializes in natural language processing and computational linguistics

Dr. J.R. Lockwood  
Principal Assessment Scientist  
- PhD in Statistics, CMU  
- Previously conducted research at RAND and ETS  
- Specializes in statistics, psychometrics, latent variable modeling, and data analysis
Dr. Geoff LaFlair  
Senior Assessment Scientist  
- PhD in Applied Linguistics, NAU  

Dr. Yigal Attali  
Principal Assessment Scientist  
- PhD in Cognitive Psychology, The Hebrew Univ. of Jerusalem  
- Extensive experience in development of innovative edu technology solutions, automated scoring, and automated item generation

Ramsey Cardwell  
Assessment Scientist  
- PhD candidate in Educational Measurement, UNC Greensboro  
- Previously interned at College Board and Medical Council of Canada  
- Presented at AERA, NCME, AAAL, and AALA

Dr. Sarah Goodwin  
Assessment Scientist  
- PhD in Applied Linguistics, Georgia State University  
- Developed English and Spanish items for the University of Michigan and Michigan State University  
- Specializes in item development, examinee response processes, corpus linguistics, listening, and writing

Dr. Yena Park  
Assessment Scientist  
- PhD in Second Language Studies, Indiana University  
- Research interests include listening, speaking, and integrated tasks in L2 assessment  
- Extensive experience in test development
technical advisory board

Micheline Chalhoub-Deville

- Professor, Educational Research Methodology, UNC at Greensboro
- President, International Language Testing Association
- Formerly, served on the TOEFL® Committee of Examiners, the TOEFL® Policy Board, and the British Council Assessment Board

Eunice Eunhee Jang

- Professor, Department of Applied Psychology and Human Development, University of Toronto
- Specializations in diagnostic language assessment and technology-rich learning and assessment.
- Current research includes AI-infused assessment systems supporting literacy development and international student success

Bryan Smith

- Associate Professor, English, Arizona State University
- Researcher in computer-assisted language learning
- Editor of CALICO Journal, an international journal devoted to research and discussion on technology and language learning

Stephen Sireci

- Distinguished Professor, Psychometrics and Director of the Center for Educational Assessment, UMass Amherst
- Researcher in educational test development particularly issues of validity, standard setting
- Vice President of the National Council of Measurement in Education (NCME)

Brian Junker

- Professor of Statistics and Data Science, Carnegie Mellon University
- Extensive experience in applying modern statistics and statistical psychometrics to challenging problems in assessment
- Co-chair, Design and Analysis Committee, National Assessment of Educational Progress
<table>
<thead>
<tr>
<th>Scale</th>
<th>CEFR</th>
<th>Descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>A1</td>
<td><strong>Can understand very basic English words and phrases.</strong> <strong>Can understand straightforward information and express themselves in familiar contexts.</strong></td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td><strong>Can understand very basic English words and phrases.</strong> <strong>Can understand straightforward information and express themselves in familiar contexts.</strong></td>
</tr>
<tr>
<td>60</td>
<td>B1</td>
<td><strong>Can understand the main points of concrete speech or writing on routine matters such as work and school.</strong> <strong>Can describe experiences, ambitions, opinions, and plans, although with some awkwardness or hesitation.</strong></td>
</tr>
<tr>
<td>90</td>
<td>B2</td>
<td><strong>Can fulfill most communication goals, even on unfamiliar topics. Can understand the main ideas of both concrete and abstract writing.</strong> <strong>Can interact with proficient speakers fairly easily.</strong></td>
</tr>
<tr>
<td>120</td>
<td>C1</td>
<td><strong>Can understand a variety of demanding written and spoken language including some specialized language use situations.</strong> <strong>Can grasp implicit, figurative, pragmatic, and idiomatic language.</strong> <strong>Can use language flexibly and effectively for most social, academic, and professional purposes.</strong></td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td><strong>Can understand a variety of demanding written and spoken language including some specialized language use situations.</strong> <strong>Can grasp implicit, figurative, pragmatic, and idiomatic language.</strong> <strong>Can use language flexibly and effectively for most social, academic, and professional purposes.</strong></td>
</tr>
</tbody>
</table>
## Alignment to IELTS & TOEFL

### CORRELATION

<table>
<thead>
<tr>
<th>Relationship</th>
<th>r</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>*TOEFL ↔ IELTS</td>
<td>.73</td>
<td>1,153</td>
</tr>
<tr>
<td>DET ↔ IELTS</td>
<td>.78</td>
<td>991</td>
</tr>
<tr>
<td>DET ↔ TOEFL</td>
<td>.77</td>
<td>2,139</td>
</tr>
</tbody>
</table>

*ETS, 2010*
## Scale and concordance

<table>
<thead>
<tr>
<th>Duolingo</th>
<th>IELTS</th>
<th>TOEFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>85-90</td>
<td>5.5</td>
<td>150-160</td>
</tr>
<tr>
<td>95-100</td>
<td>6</td>
<td>145</td>
</tr>
<tr>
<td>105-110</td>
<td>6.5</td>
<td>140</td>
</tr>
<tr>
<td>115-120</td>
<td>7</td>
<td>135</td>
</tr>
<tr>
<td>125-130</td>
<td>7.5</td>
<td>130</td>
</tr>
<tr>
<td>135-140</td>
<td>8</td>
<td>125</td>
</tr>
<tr>
<td>145-150</td>
<td>8.5</td>
<td>120</td>
</tr>
<tr>
<td>155-160</td>
<td>9</td>
<td>97-102</td>
</tr>
</tbody>
</table>
Computer Adaptive Testing

Note: item types are equally represented in the test, to represent all skills. Each format is shown 3–7 times (6 on average).
Reliability

- **Internal consistency**: how well the questions work together to measure the component of language
- **Test-retest**: two different test scores for the same people within 30 days

<table>
<thead>
<tr>
<th>Subscore</th>
<th>Internal consistency</th>
<th>Test-retest reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literacy</td>
<td>0.89</td>
<td>0.82</td>
</tr>
<tr>
<td>Conversation</td>
<td>0.93</td>
<td>0.80</td>
</tr>
<tr>
<td>Comprehension</td>
<td>0.95</td>
<td>0.78</td>
</tr>
<tr>
<td>Production</td>
<td>0.76</td>
<td>0.83</td>
</tr>
</tbody>
</table>

*Note: The closer the value of these are to 1.00 the better. Internal consistency was calculated using split-half reliability methods.*
Digital-first means unique security advantages

- Large item bank means we’re not worried about item exposure
- No supply-chain risks (e.g. bribed printers)
- ML assistance and behaviour database grows with every test taker
- Adaptive digital delivery eliminates test theft problems
- We can easily collect test session and test taker video/audio recordings
Human proctoring

1:25

Varies depending on the test and center, but on average, 1 proctor is tasked with supervising ~25 test-takers

Multiple:1

Each test session is individually reviewed by multiple highly trained proctors after test completion
2021 Graduate Data
test-taker data & insights

Snapshot: Number of Duolingo English Test taker cities in a range of countries and regions

12,000+

Test Taker Cities
India insights

India insights for 740+ cities.
Test-taker intent: Academic Level
September - October 2021

- Graduate (41.6%)
- Undergraduate (38.2%)
- None (10.2%)
- Not applying (5.9%)
- Secondary school (3.1%)
- Other (1.0%)
Graduate test-taker intent: Field of Study
September - October 2021

- Business & Management (23.6%)
- Computer Science & Math (18.1%)
- Engineering (12.1%)
- Science - Life (3.0%)
- Other (5.6%)
- Other (37.6%)
Institutional test share reports - New in 2022

- End of admission cycle test-taker summary
- Top countries of test takers
- Overlapping institutions for score sends
Institutional Perspective: Florida Tech
Institutional Overview

FLORIDA INSTITUTE OF TECHNOLOGY, MELBOURNE FL
....on Florida’s Space Coast
Why the Duolingo English Test?

- Accessibility
- Low cost to student
- Ease for student
- Quick turnaround of results
- Fee waivers available
- High level security
- Excellent technical support and customer service
Florida Tech began to accept DET in Spring 2019 after visiting the Pittsburgh office and subsequent approval by English & Languages Division, School of Arts & Communication.

First “test” group were members of our Crew team. They had been admitted but were competing in the European championships and could not schedule alternative testing. DET made it very easy for them to test over the summer while on the road for the Fall 2019 cycle.
Using the Test & Score Criteria

Scores received via:

Slate integration

- Scores electronically uploaded, ELP automatically satisfied if DET score of 100 or 120 for GSA

DET dashboard allows verification and ensures secure certified results for institutions
<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Test Scores Received (UG &amp; Grad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-19</td>
<td>2</td>
</tr>
<tr>
<td>2019-20</td>
<td>10</td>
</tr>
<tr>
<td>2020-21</td>
<td>281</td>
</tr>
<tr>
<td>2021-22 (to date)</td>
<td>1556</td>
</tr>
<tr>
<td>2022-23 (to date)</td>
<td>34</td>
</tr>
</tbody>
</table>
Lessons Learned

- No difference in academic performance for those who provided DET vs. any other ELP score - did not affect retention

- Glad we pushed DET approval - signed on early - 100% ready to offer international students an easy solution in 2020 and our AY 2020-2021 enrollments did not suffer, quite the contrary (remote access + on-campus)

- You can never have too many methods of demonstrating ELP.
Test Data Reports - Coming in 2022
Use in Practice:
Test Taker Journey
Traditional Testing

Students need to take an English test

Book a date

Travel to test center—sometimes to a different city or country

+3 hours to take a test
Duolingo English Test

Students need to take an English test

Take the test any day you want, any time you want

Test anywhere you have internet access and a computer
Use in Practice: Institutional Experience
Institutional resource center

englishtest.duolingo.com/dashboard
## DASHBOARD

### Receiving Results

<table>
<thead>
<tr>
<th>Surname(s)</th>
<th>Given Name(s)</th>
<th>Test Taken</th>
<th>Score Sent</th>
<th>Score</th>
<th>Interview</th>
<th>Certificate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kachmar</td>
<td>Pavlo</td>
<td>Jul 15, 2019, 12:24 PM</td>
<td>Jul 17, 2019, 10:29 AM</td>
<td>125</td>
<td>VIEW</td>
<td>VIEW</td>
</tr>
<tr>
<td>Ogita</td>
<td>Malachi</td>
<td>Jul 15, 2019, 12:14 PM</td>
<td>Jul 17, 2019, 10:29 AM</td>
<td>125</td>
<td>VIEW</td>
<td>VIEW</td>
</tr>
<tr>
<td>Mandsanka</td>
<td>Tanvi</td>
<td>Jul 15, 2019, 12:28 PM</td>
<td>Jul 17, 2019, 10:29 AM</td>
<td>100</td>
<td>VIEW</td>
<td>VIEW</td>
</tr>
<tr>
<td>Gito</td>
<td>Christian</td>
<td>Jul 15, 2019, 11:57 AM</td>
<td>Jul 17, 2019, 10:29 AM</td>
<td>125</td>
<td>VIEW</td>
<td>VIEW</td>
</tr>
</tbody>
</table>

*Email: christianjto.freelancer@gmail.com*  
*DOB: 1989-10-26*  
*Coupon: VERSIONTHREE*
Describe the last place you visited with your family or close friends. What did you like most about this place? How could the place be improved for the people who visit there?

Money isn’t always the only motivation for having a job. Other important reasons include the desire to be productive and to contribute to society. Many people truly enjoy utilizing their skills and serving people who need their assistance. I would keep a job even if I didn’t need one to live so that I could remain connected to the world.
Slate and other CRM integrations
thanks!

questions?

institutional@duolingo.com