Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

Index

*Page numbers with *f* indicate figures; page numbers with *t* indicate tables.

Access: educational, 106; as fairness issue, 106, 113–115, 133–134; financial, 113, 118; geographical, 106, 113–114, 118; personal, 114
Accessibility of tests, 117–118
Accommodations for test-takers, 18, 72, 118
ACT®, 134
Alternative assessment, 18, 19, 147
American Council on the Teaching of Foreign Languages (ACTHL), performance scales developed by, 42
American Educational Research Association (AERA), 136
American Psychological Association (APA), 136
Analytic rubrics, 85–86, 87–88f. *See also* Rubrics
Analytic scales, 44f, 45
Assessing Writing Journal, 22
Assessment: alternative, 18, 19, 147; authentic, 50; classroom (*See* Classroom assessment); collaboration in development of, 21; in curriculum planning, 33f; development and use of, in educational contexts, 9–11; diagnostic, 5; diversity in, 119–120; embedded, 26; formative (*See* Formative assessment); future directions for, 148–150; in-class, 24; inclusive thinking on, 21; independent skills, 8; integrated reading-writing, 75; integrated skills, 8, 79–80, 81; as just writing tests and using statistics, 13–22; language (*See* Language assessment); learner-centered, 149; on-the-fly, 25; paper-based, 3, 147; peer and self, 13, 26, 29, 37; performance (*See* Performance assessment); purpose of, 93–94; social contexts of, 15; static, 27; student involvement in, 36–38; summative (*See* Summative assessment); teacher competence in, 14, 15–19, 16f; transparency in, 116–117
Assessment for Learning (Afl), 27–30, 140, 149: feedback in, 29; implementing, 32; in second language writing classrooms, 29
Assessment literacy, 15, 150: building though learning communities and collaboration, 21–22; developing, 20–22; on test preparation, 134–135
Assessment Reform Group, 27–28, 140–141
Audiolingual approach, 76

Bachman, L.F., 51, 68
Backwash, 123. *See also* Washback
Bailey, K. M., 17
Banerjee, J., 98–99, 103

Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

Beddow, P.A., 128, 131
Bias, 9, 10–11, 108, 109, 116: content-related, 109; in education, 116; in item format, 110; language-related, 109; reducing, 53; research on, 110. *See also* Fairness
Black, P., 24, 26
Braden, J.P., 128, 131
Broussard, M., 114–115, 118
Brown, J.D., 17, 47, 132, 145

Canale, M., 40, 94
Carson, J., 50, 78, 79
Chalhoub-Deville, M., 41, 47
Chapelle, C.A., 91, 97, 98, 147
Checklists, 51, 101: exit, 98, 101; as formative assessment, 25f; validity of, 98–99
Cheng, L., 111, 112–113, 123, 126, 135
Clang associations, 65
Classroom assessment, 7, 19, 140: evaluation of process and product in, 3; fairness in, 111–113, 118–119; for language teaching, 26–32
Code of Fair Testing Practices Joint Committee on Testing Practices, 138
Cohen, A. D., 43, 97
Colby-Kelly, C., 26, 28–29
Cole, N. S., 107–108
Collaboration: on assessment development and dilemmas, 21; building assessment literacy through, 21–22
College English Test--Spoken English Test (CET-SET), 47
Commercially designed tests, versus teacher-designed tests, 145–146
Common European Framework of Reference (CERF), 43, 94
Communicative adequacy, 94, 96; holistic scoring of, 95–96
Communicative competence, 16, 40, 41f, 43, 94–95, 96; assessment of, 40; defined, 40
Communicative language approach, 13, 17, 97
Comprehensive final exams, 23–38; research on, 24–32; what to do in the classroom, 33–38
Computer-based version TOEFL® (CBT), 114, 147
Constructed-response items/tasks, 7, 59, 105
Construct irrelevant, 109, 119
Constructs, 10, 13, 24–25, 93
Construct underrepresentation, 58–59
Construct validity, 92, 93–97
Content-based instruction, 42–43
Content validity, 92, 93, 96
Criterion-referenced assessment, 18
Criterion validity, 92, 93
Crocker, L., 126, 131
Cronbach, L., 93
Curriculum alignment, 125, 126
Curriculum planning, assessment in, 33f
Curriculum validity, 92

Data-driven scales, 46, 53; in discriminating writing proficiencies, 45
Differential item functioning (DIF), 110, 113
Direct tests, comparison of multiple choice tests with, 59
Discourse competence, 41f
Discourse synthesis, 99, 143
Diversity in assessments, 119–120
Downing, S., 60, 65
Dynamic Assessment (DA), 27, 30–32, 140; implementing, 32

Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

Educational contexts, development and use of assessment in, 9–11
Educational Testing Service (ETS), 42; development of TSE® by, 42; development of TWE® by, 42; *ETS Guidelines for Fairness Review of Assessments,* 108, 119
Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS), 124
Elder, C., 47, 106, 109
Embedded assessment, 26
English for Academic Purposes (EAP) program, 28, 75; skill integration in, 78–79
Enright, J. M., 97, 98
Essay writing, 3, 7–8
Ethics, test preparation and, 131–132, 132f
ETS Guidelines for Fairness Review of Assessments, 108, 119
Exit checklists, 98, 101

Face validity, 92
Fairness, 1, 8, 55, 104–120, 133, 144; access and, 106, 113–115, 133–134; in the classroom, 111–113, 118–119; defined, 106, 107; final comments on, 116;
justice and, 106; learning opportunities and tests and, 107–110; outcomes,
111–112; procedural, 111; research on, 106–116; reviewing tests for, 108–110; validity and, 106; what to do in classroom, 116–120
Fairness in Testing Manual of the Data Recognition Corporation, 120
Feedback, 28–29; in Assessment for Learning, 29; checking on clarity of, 36; emphasis on, in assessment, 35–36; grading on improvement from, 36; requiring revisions based on, 35–36
Ferris, D., 79
Fill-in-the-blank test items, 3, 40
Financial access, 113, 118
Fitzgerald, J., 75, 77
Formative assessment, 7, 24–25, 140; best practices in, 33; characteristics of, 25; defined, 24; examples of, 25f; goal for, 25; need for continuity between summative assessment and, 33; recent developments in, 26; types of, 25–26
Fulcher, G., 20, 41, 68

Gebril, A., 59, 75, 79, 81, 84, 124, 149
Geographical access, 106, 113–114, 118
Gilabert, R., 94–95, 96
Grammatical competence, 41f
GRE® test, 18
Guessing, control of, in multiple choice tests, 59, 64

Harrison, C., 24, 26
High-stakes/standardized testing, 7, 51
Hill, K., 47, 106
Holistic scales, 44f, 95–96
Horowitz, D., 50, 78

ILTA Code of Ethics & Code of Practice, 137–138
Inbar-Lourie, O., 14, 15, 16
Independent skills assessment, 8
Integrated reading-writing assessment, 75; benefits of, 75
Integrated skills assessment, 8, 81; student performance on, 79–80

Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

Integrated tasks: paying attention to, in scoring, 85–86, 87–88f; writing effective, 81, 82–83f, 84–85
Interagency Language Roundtable (ILR), 41
International English Language Testing System (IELTS®), 6, 98
International Language Testing Association (ILTA), 138
Interviews, oral, 3, 7–8, 41, 42, 100
Item analysis, 9, 13. *See also* Multiple choice tests
Item discrimination, 68–72
Item response theory (IRT), 60, 110

Jamieson, J., 97, 98
Jeong, H., 19, 139
Johns, A. M., 75
Justice, 115; fairness and, 106

Keith-Spiegel, P., 111, 112
Kettler, R. J., 128, 131
Kim, Y., 47, 109
Knoch, U., 45, 53
Kozulin, .A., 31–32
Kuiken, A., 94–95, 96
Kunnan, A., 106, 113

Language: as four separate skills, 76–78, 143; view of, 101
Language assessment: defined, 2–3; developing, 2–3; enthusiasm for performance assessment in, 42–43; growth in field of, 2; power of, 3; reasons for, 2; teaching, 17–18; technology in, 147; uses of, 2
Language assessment literacy, 16–17, 20
Language Assessment Quarterly, 22
Language learning: evaluation in, 2; mediation in guiding, 30–31
Language teaching, classroom-based assessment for, 26–32
Language testing: 1, 5–8, 17, 142
Language Testing Journal, 22
Language Testing Resources (website), 22
Learning communities, building assessment literacy through, 21–22
Learning opportunities and tests, 107–110
Lee, C., 24, 26
Lee, I., 28, 29
Leki, I., 50, 79
Lumely, T., 46–47, 106

Master's in TESOL courses. *See* Teacher training
McNamara, T., 15, 106, 115
Mediation in guiding language learner, 30–31
Messick, S., 93, 123, 145
Multiple choice items/tests, 3, 40, 56–73, 142: anatomy of, 64f; assessing difficulty of, 68–72, 70t, 142; comparison with direct tests, 59; control of guessing in, 59, 64; critiquing, 67, 67f; establishing clear test specifications, 61–63, 63t; higher score reliability of, 59; item analysis on, 13; negative associations with, 58; piloting, 68; planning, 60–61; research on, 59–60; scoring, 59; taking advantage of test accommodation strategies, 72–73, 72t, 73f; validity and, 8, 58–59; what to do in classroom, 60–73; writing effective questions, 63–66, 66f

Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

National Council on Measurement in Education (NCME), 136–137; Code of Professional Responsibilities, 138
Needs analysis, 5, 34-35, 50, 60–61, 101–102
Negative washback, 10, 124, 142

Objective assessment, versus performance assessment, 7–8
One-on-one conferences, 35
Oral interviews, 3, 7–8, 41, 42, 100
Oral Proficiency Interview (OPI), 42
Outcomes fairness, 111–112

Paired terms, 6–8
Palmer, A., 51
Paper-based assessment, 3, 147
Park, J. J., 134, 135
Peer and self-assessment, 13, 26, 29, 37
Peer review, 28, 149
Performance assessment, 3, 39–55, 141: designing rubrics for, 37; enthusiasm for, in language assessment, 42–43; versus objective assessment, 7–8; research on, 40–49;; scoring rubrics for, 141; transforming into scores, 43, 45; what to do in classroom, 50–55
Perkins, D. V., 111, 112
Placement exams, 5–6
Plakans, L., 53, 75, 76, 79, 80, 81, 99, 124, 146
Positive washback, 10, 124
Post-test information gathering, 102
Primary trait scales, 45f
Private tutoring, 124–125
Professional Learning Communities (PLCs), 21–22
Professional standards, upholding, 136–138, 137f
Proficiency tests, 6, 101

Qualitative techniques, 110
Quantitative techniques, 110
Quickwrites, 14, 37

Rating rubrics, 39–40
Rating scales, 9, 13. *See also* Rubrics: analytic, 44f; in assessing writing, 45; data driven, 46; differences between rubrics and, 41, 42f; holistic, 44f, 47; practice based, 46; primary trait, 45f; theory based, 46
Ravitch, D., 108, 120
Reading placement tests, 89, 91
Reliability, 8, 19; of scores, 43, 45; in teacher-based assessment, 20
Revisions: requiring, based on feedback, 35–36; selective, 35–36
Roever, C., 15, 115
Rubrics, 9, 13. *See also* Rating scales: analytic, 85–86, 87–88f; designing for performance assessment, 37; development of, 19; differences between rating scales and, 41, 42f; involvement of students in development of, 37; for performance assessment, 141; responding to trustworthiness of, 54f, 55; selecting, developing, and critiquing, 52–54

Sawaki, Y., 75, 77–78
Scores, 39–55, 43, 45–49: interpretation of, 9–10, 89, 90; reliability of, 43, 45; research on, 40–49; transforming performance into, 43, 45; what to do in classroom, 50–55
Scoring, 9; paying attention to integrated tasks in, 85–86, 87–88f; preparing for, 53, 55
Second language proficiency, 47–48

Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

Second language writing classrooms: Assessment for Learning in, 29; assessment plan for, 34–35, 34t
Selected response tests items, 7, 59
Self-assessment, 146. *See also* Peer and self-assessment
Shanahan, T., 75, 77
Shepard, L.A., 15, 117
Shohamy, E., 111, 115, 134
Single skill testing, 74–88; research on, 76–80; responding to, 81–88
Skill integration, in English for Academic Purposes (EAP) program, 78–79
Stakeholders, educating, 135–136
Standards for Educational and Psychological Testing, 136
State-mandated standardized tests, test preparation for, 127
Strategic competence, 41f
Students: involvement of, in rubric development, 37; involving in assessment, 36–38; language assessment by the teacher, 146; performance on integrated skills assessment, 79–80; reasons for assessing, 101; teacher involvement in preparing for tests, 121–138; research on, 123–132; using multiple measures to evaluate, 102; what to do in classroom, 133–138
Summative assessment, 7, 24–25, 29; defined, 24; examples of, 25f; need for continuity between formative assessment and, 33; use of, in formative ways, 26
Swain, M., 40, 94

Taylor, C., 15, 16, 78
Teacher-designed tests, versus commercially designed tests, 145–146
Teachers: assessment of students' language by, 146; competence in assessment, 15–19, 16f; involvement in assessment, 14
Teacher training, language testing courses in, 10, 17, 56, 148–150
Teaching, relationship between testing and, 123
Teaching to the test, 126, 144–145
Technology in language assessment, 147
Test banks, 146
Testing: alignment of integrated skills with, 81; within the curriculum, 18; purposes of, 5–6; relationship between teaching and, 123. *See also* Assessment
Testing-learning-testing format, 31–32
Test items, bias in, 109
Test preparation: assessment literacy on, 134–135; different approaches to, 126–128, 128t; ethics and, 131–132, 132f; management of, 130; providing students with access to materials, 133–134; teacher involvement in, 121–138; research on, 123–132; terms related to, 123–126; validity and, 128–130; what to do in classroom, 133–138
Tests: critiquing, 102–103; learning opportunities and, 107–110; making accessible, 117–118; reviewing for fairness, 108–110
Test specifications, 18, 19
Test-taking strategies, 128–129; teaching, 133

Assessment Myths
Applying Second Language Research to Classroom Teaching
Lia Plakans and Atta Gebril
http://www.press.umich.edu/5056216/assessment_myths
Michigan ELT, 2015

Thanaweya anna test, 121–122, 123–124, 135
Think-alouds, 46, 51, 100
TOEFL® (Test of English as a Foreign Language), 6, 13, 134
TOEFL® iBT test, 42, 98, 124
Transparency, 116–117
TSE® (Test of Spoken English), 42
Turner, C. E., 26, 28–29, 53, 146
TWE® (Test of Written English), 42

University placement testing, 74–75

Validity, 8, 19, 89–103, 115, 143–144; arguments on, 97–98; of checklists, 98–99; collecting evidence on, 98–101; complexity of, as concept, 91; construct, 92, 93–97; content, 92, 93, 96; criterion, 92, 93; curriculum, 92; defined, 89–90, 91, 100–101; evolving definitions of, 92–93; face, 92–93; fairness and, 106; interpretation of scores and, 89, 90; multiple choice tests and, 58–59; predictive, 92; purpose of test and, 89; research on, 91–101; in teacher-based assessment, 20; test preparation and, 128–130; what to do in classroom, 101–103
Vedder, I., 94–95, 96

Wall, D., 98–99, 100, 123, 126
Wang, X., 111, 112–113
Washback, 10, 18, 123–124, 126; negative, 124, 142; positive, 124
Weigle, S. C., 47, 79, 84
Whitley, B. E., 111, 112
Williams, D., 24, 26
Wittig, A. F., 111, 112
Writing: conferences on, 26; data-driven scales in discriminating proficiencies, 45; essays, 3, 7–8; persuasive, 84; process, 28; rating scales in assessing, 28, 45; in second language classrooms, 29, 34–35, 34t

Xie, Q., 129–131