SYLLABUS Jan. 6, 2023

LLT 808 Assessment for Language Teaching and Research (3 credits)

Michigan State University, Department of Linguistics, Languages, and Cultures

Instructor: Professor Paula M. Winke, Ph.D. Dylan Burton, M.A.
B252 Wells Hall, East Lansing, MI, 48824 B331 Wells Hall, East Lansing, MI 48824
Office phone: 517-353-9792 jdburton@msu.edu
Cell: 517-775-0156
winke@msu.edu

Mondays,* 9:10 AM to noon, 1/9/2023 - 4/30/2023, A304 Wells Hall
*Asynchronous class (due to AAAL conference): Monday, March 20, 2023
- Final exam materials due by 12:00 Noon Eastern, Monday, May 1, 2023

Paula Winke's office hours: 8 to 9 AM Wednesdays, and 2 to 3 PM Fridays, or by appointment (online or in person). Due to my work load this semester that involves travel, please schedule with me ahead of time to come to office hours. Online office hours will be at this location: https://msu.zoom.us/my/winke
You can sign up using my MSU Bookings at the following link. Choose a 15-minute, 30-minute, or 60-minute time slot first, and then the date, to check for availability:
https://outlook.office365.com/owa/calendar/Bookings.PaulaWinke@booking.msu.edu/bookings/ If you don't see openings that work for you, see me before or after class and we can put a meeting on my calendar. That works well, as I can open my calendar and "pencil" you in.

Classroom Management System: LLT 808 Teams site

Course Description/Aims & Goals/Instructional Objectives:

In this course, we review current theories and methods in L2 assessment for classroom-based language teaching, programmatic language assessment for monitoring or tracking growth, and for second language acquisition and applied linguistics research (data collection) purposes. The main goals and instructional objectives are as follows:

- To have students understand and be able to apply Classical Test Theory (CTT) to test data for the purposes of creating, reviewing, or redesigning test items or an overall test.
- To have students be able to critically define the constructs they want to test, and meanwhile ensure those constructs stem from defensible theories within applied linguistics.
- To have students who can create assessments that reliably measure the defined constructs, and that prove, over time, to be valid measures of the constructs (produce scores that are interpretable, meaningful, appropriate, and credible; come from tests that "do no harm" and are ethical and just).
- To have students who can write good multiple-choice and other discrete-point items (including Cloze tests, C-tests, etc.), but who also understand the limits on what those items can measure. It is my hope that students will learn to use them sparingly or for specific, comprehension-checking (socially beneficial, lower-stakes) purposes.
- To have students who can design a variety of assessments and assessment components, such as self-assessments, portfolio assessments, peer-review rubrics/checklists and procedures, performance assessment rubrics, observation sheet/checklists, rater training procedures and materials, etc.
- To have students who can critically evaluate the power of tests, and who can become advocates for appropriate test design and test-score uses.

In class, students will review and discuss current Instructed SLA (ISLA), SLA, and cognitive-psychology theories, hypotheses, and principles that highly influence L2 assessment, such as:

- The Critical Period Hypothesis (CPH) & age effects in L2 learning
- Pronunciation ("native"-speakerism ideologies), identity, and sociocultural theories
- CAF (complexity, accuracy, and fluency) in L2 production

- Motivation and anxiety in SLA
- Skill acquisition theory; practice in SLA
- Cognitive differences in SLA (working memory, short and long-term memory, exceptionalisms, etc.)
- Interaction hypothesis, input hypothesis, and multimodal input in SLA
- Communicative and task-based language teaching

As the final exam, students will do one of the following:

- Join an on-going or collaborative language testing research project. People who may have projects you can join are, for example, me, Kadidja Kone, Monique Yoder, Dylan Burton, or Dan Reed.
- Create and trial or pilot a language assessment for a degree milestone (your thesis, QRP, or dissertation). Here, the test, survey, or rubric you create is an instrument or material you will use in your thesis, QRP, or dissertation. For the final, you would write up the description of the material, describe the SLA theories it is based on, turn in the material, and write any information on the piloting or trailing of the material.
- Design your own language assessment research project that could be your thesis, QRP, or dissertation. Here, a test, test condition, test effect, or testing procedure is the object of study. This can be anywhere from a research proposal (lit review, RQs, methods section) to a full study.
- Help an outside organization with their current language testing needs. This is a la carte and would be discussed more if anyone wants to do this. I will describe potential options.

Auditors may participate in all assignments and obtain feedback from me if they would like, although the assignments are not required for auditors.

CLASS SCHEDULE& TOPICS

Please note the schedule and topics are subject to change. If there are changes, I will post a revised syllabus in Teams, and notify you of the changes and new posting of the syllabus.

Date/Time	Topic Area
Week 1: Mon., Jan. 9	**Introduction to language testing, each other, and this course** Required Reading: • Chapter 1, Assessment concepts and issues (Brown & Abeywickrama, 2019) • Chang-Bacon, C. K., & Colomer, S. E. (2022). Biliteracy as property: Promises and perils of the Seal of Biliteracy. *Journal of Literacy Research 54*(2), 111-215. https://doi.org/10.1177/1086296X221096676 • Davin, K. J. (2021). Critical language testing: Factors influencing students' decisions to (not) pursue the Seal of Biliteracy. *Harvard Education Review, 91*(2), 179-203. https://doi.org/10.17763/1943-5045-91.2.179 Optional Reading (highly recommended for Ph.D. students, or MA students doing a thesis): • Harding, L., & Kremmel, B. (2021). Chapter 6: SLA researcher assessment literacy. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 54-65). Routledge.
Week 2	No class. Martin Luther King Jr. Day
Week 3: Mon., Jan. 23	**Reliability, validity, and test washback** Required Reading: • Chapter 2, Principles of language assessment (Brown & Abeywickrama, 2019) • Winke, P. (2011). Investigating the reliability of the civics component of the U.S. Naturalization Test. *Language Assessment Quarterly, 8*(4), 317-341. https://doi.org/10.1080/15434303.2011.614031 **Guest presenter: Caitlin Cornell, 10 to 11 AM (Survey trialing & design)** Read ONE of these three (whichever one interests you most): • MacKay, T., & Plonsky, L. (2021). Chapter 44: Reliability analysis: Estimating error. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 468-482). Routledge. • Phakiti, A. (2021). Chapter 10: Likert-type scale construction. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 102-114). Routledge. • Green, R. (2021). Chapter 11: Pilot testing: Why and how we trial. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 115-124). Routledge.

	Turn in your LLBA search assignment to the Teams channel "LLBA Review" before the start of class. In class, you will be asked to use the chalk board to outline for the class what you found in your LLBA search. ✓ Assignment turned in and discussion of it = 10% of grade
Week 4: Mon., Jan. 30	**Washback continued. / Construct definition, target-language use domain, the test design cycle, and test specifications** Required Reading: • *Tsang, C. L., & Isaacs, T. (2022). Hong Kong secondary students' perspectives on selecting test difficulty level and learner washback: Effects of a graded approach to assessment. *Language Testing, 39*(2), 212-238. https://doi.org/10.1177/02655322211050600 • Chapter 3, Designing classroom language tests (Brown & Abeywickrama, 2019) • Green, A., & Fulcher, G. (2021). Test design cycle. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 69–77). Routledge. ***Schedule individual meetings with me during this week <u>or next week</u> to discuss preliminary final project plans. These can be solo, in pairs, or in groups. We can meet on Zoom in the evenings or on the weekend if needed. Just let me know.*** ✓ Meeting = 10% of grade
Week 5: Mon., Feb. 6	**Assessing listening** Required Reading: • Chapter 6, Assessing listening (Brown & Abeywickrama, 2019) • Lam, D. (2021). Don't turn a deaf ear: A case for assessing interactive listening. *Applied Linguistics*, *42*(4), 740–764. https://doi.org/10.1093/applin/amaa064 • *Rukthong, A., & Brunfaut, T. (2020). Is anybody listening? The nature of second language listening in integrated listening-to-summarize tasks. *Language Testing, 37*(1), 31–53. https://doi.org/10.1177/0265532219871470 • Nishizawa, H. (2023). Construct validity and fairness of an operational listening test with World Englishes. *Language Testing*. Advance online publication. https://doi.org/10.1177/02655322221137869
Week 6: Mon., Feb. 13	**Assessing speaking** ***Guest Instructor: Dylan Burton*** ***Readings*** • Chapter 7, Assessing speaking (Brown & Abeywickrama, 2019) • de Jong, N. H. (2023). Assessing second language speaking proficiency. *Annual Review of Linguistics, 9*. https://doi.org/10.1146/annurev-linguistics-030521-052114 • *Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing, 20*(1), 1–25. https://doi.org/10.1191/0265532203lt242oa

Week 7: Mon., Feb. 20	**Consequences of standards-based assessment and standardized testing** *Readings* • Chapter 4, Standards based assessment (Brown & Abeywickrama, 2019) • Sleeter, C. E., & Flores Carmona, J. (2017) Chapter 1: Standards, multicultural education, and central curriculum questions. In *Un-standardizing the curriculum: Multicultural teaching in the standards-based classroom,* 2nd ed. (pp. 5-23). Teachers College Press. • *Macqueen, S., Knoch, U., Wigglesworth, G., Nordlinger, R., Singer, R., McNamara, T., & Brickle, R. (2019). The impact of national standardized literacy and numeracy testing on children and teaching staff in remote Australian Indigenous communities. *Language Testing, 36*(2), 265-287. https://doi.org/10.1177/0265532218775758 *Due before the start of class: Submit to "Proposal Outline" file folder in Teams a one-page proposal outlining the final project plan.* ✓ Proposal outline = 10% of grade *Participants will orally share, in class, ideas (time limit = 8 to 10-minutes) that they have concerning their final project plans to gather peer feedback and to organize peer-feedback pods.* ✓ Presentation = 10% of grade
Week 8: Mon., Feb. 27	**Standardized or "general" proficiency testing** *Readings* • Chapter 5, Standardized testing (Brown & Abeywickrama, 2019) • Deygers, B. (2022). Chapter 34: Standards in language proficiency measurement. In G. Fulcher & L. Harding (Eds.), *The Routledge handbook of language testing* (2nd ed.) (pp. 578-596). Routledge. • *Tremblay, A. Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition, 33*(3), 339-372. https://www.jstor.org/stable/44486009
Mon., March 6	MSU Spring Break (Monday, 3/6 - Friday, 3/10)
Week 9: Mon., March 13	**Assessing reading, and multimodality in language assessment** *Readings* • Chapter 8, Assessing reading (Brown & Abeywickrama, 2019) • *Serafini, F., Moses, L., Kachorsky, D., & Rylak, D. (2020). Incorporating multimodal literacies into classroom-based reading assessment. *The Reading Teacher, 74*(3), 285– 296. https://doi.org/10.1002/trtr.1948 • *Košak-Babuder, M., Kormos, J., Ratajczak, M., & Pižorn, K. (2019). The effect of read-aloud assistance on the text comprehension of dyslexic and non-dyslexic English language learners. *Language Testing, 36*(1), 51-75. https://doi.org/10.1177/0265532218756946 • Plakans, L. (2022). Chapter 21: Writing integrated tasks. In G. Fulcher & L. Harding (Eds.), *The Routledge handbook of language testing* (2nd ed.) (pp. 357-371). Routledge.
Week 10: Mon., March 20	**When standardized testing goes bad, it can go very bad.**

	Due to this week being AAAL, will have asynchronous class work. The work will be scored on participation only. You will be presented with a Google Slide Show with links to articles and videos on language testing scandals. Pick the one you are most interested in, and read all you can about the specific scandal through the links provided. Then, fill in the Google Slide with your name on it to respond to the scandal you read about. More info to be given out March 13. *If you are attending AAAL, you can optionally choose to help me develop the "Scandal" Google slides before March 13 (which should be a ton of fun), and then you will not have to do the assignment during the week of AAAL.*
Week 11: Mon., March 27	**Assessing writing, creating writing rubrics, and discussions on the construct of writing** *Readings* • Chapter 9, Assessing writing (Brown & Abeywickrama, 2019) • Knoch, U. (2009). Chapter 3: Rating scales. *Diagnostic writing assessment.* Peter Lang. • *Norris, J. M., & East, M. (2021). Task-based language assessment. In M. Ahmadian & M. Long (Eds.), *The Cambridge handbook of task-based language teaching* (pp. 505–528). Cambridge University Press. • Winke, P., & Lim, H. (2015). ESL essay raters' cognitive processes in applying the Jacobs et al. rubric: An eye-movement study. *Assessing Writing, 25*, 38-54. https://doi.org/10.1016/j.asw.2015.05.002
Week 12: Mon., April 3	**Complexity, forms, and vocabulary testing** *Readings* • Chapter 10, Assessing grammar and vocabulary (Brown & Abeywickrama, 2019) • *Lan, G., Liu, Q., & Staples, S. (2019). Grammatical complexity: 'What does it mean' and 'So what' for L2 writing classrooms? J*ournal of Second Language Writing, 46.* https://doi.org/10.1016/j.jslw.2019.100673 • *Schmitt, N., Nation, P., & Kremmel, B. (2020). Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation. *Language Teaching, 53*(1), 109-120. https://doi.org/10.1017/S0261444819000326
Week 13: Mon., April 10	**More on rubric design, score reporting, and on scoring open-ended responses** *Readings* • Chapter 11, Grading and student evaluation (Brown & Abeywickrama, 2019) • Knoch, U., Fairbairn, J., & Jin, Y. (2021). Chapter 1: Understanding raters and ratings. *Scoring second language spoken and written performance*. Equinox. • *Khabbazbashi, N., & Galaczi, E. D. (2020). A comparison of holistic, analytic, and part marking models in speaking assessment. *Language Testing, 37*(3), 333–360. https://doi.org/10.1177/0265532219898635 • Zenisky, A. L., & Hambleton, R. K. (2016). A model and good practices for score reporting. In S. Lane, M. R. Raymond, T. M. Haladanya (Eds.), *The handbook of test development* (pp. 585–602). Routledge.

Week 14: Mon., April 17	**Self-assessment / A discussion on the contributions of background to assessment results** ***Readings*** • Chapter 12, Beyond letter grading (Brown & Abeywickrama, 2019) • Winke, P., Zhang, X., & Pierce, S. J. (2022). A closer look at a marginalized test method: Self-assessment as a measure of speaking proficiency. *Studies in Second Language Acquisition.* Advance online publication. https://doi.org/10.1017/S0272263122000079 • *Gujourd, A.-K. H. (2022). Who succeeds and who fails? Exploring the role of background variables in explaining the outcomes of L2 language tests. *Language Testing.* Advance online publication. https://doi.org/10.1177/02655322221100115 ***Participants will orally share, in class, their final projects (time limit = 10 to 15-minutes) to gather peer feedback and get tips for any potential revisions.*** ✓ Final project presentations = 10% of grade
Week 15: Mon., April 24	Key tips on test design for SLA research • *Suzuki, Y., & Koizumi, R. (2021). Using equivalent test forms in SLA pre-test, post-test design research. In P. Winke & T. Brunfaut (Eds.), The Routledge handbook of second language acquisition and language testing (pp. 115-124). Routledge. • Lee, H., & Winke, P. (2013). The differences among three-, four-, and five-option multiple-choice item formats in the context of a high-stakes English-language listening test. *Language Testing, 30*(1), 99-123. https://doi.org/10.1177/0265532212451235 • Winke, P., & Gass, S. (2016) Using free recall and idea units for evaluating second language comprehension: Methodological choices and issues. *Applied Linguistics Forum*. TESOL International Association. http://newsmanager.commpartners.com/tesolalis/issues/2016-11-04/5.html We can also revisit (from Week 3): • MacKay, T., & Plonsky, L. (2021). Chapter 44: Reliability analysis: Estimating error. In P. Winke & T. Brunfaut (Eds.), The Routledge handbook of second language acquisition and language testing (pp. 468-482). Routledge. ***Participants will orally share, in class, their final projects (time limit = 10 to 15-minutes) to gather peer feedback and get tips for any potential revisions.*** ✓ Final project presentations = 10% of grade

Mon., May 1, 2023: Submission of final project by 12 Noon to LLT 808 Teams.
✓ Final exam project/deliverable = 35% of grade

GRADING

*Articles or chapters marked with an asterisk are ones that you or a pair or group can decide to have as the base of a **task, activity, debate, or discussion that you design or lead in class**. This is worth 15% of your grade. It is very wide open as to what you can do. I encourage you to talk to me ahead of time to discuss ideas of what you might want to do. We will talk in class too about various ideas about this part of the course and student-led, interactive work.

Each student will meet individually with me to plan for their final exam project in Week 4 or 5. These final projects may be produced by an individual, a pair, or a group of students. Auditors are encouraged/asked to join or form groups to participate in the final projects; however, they will not receive credit at MSU for doing so. Auditors who are joining groups whose members are taking the class for credit must receive permission from the credit-seeking students before joining their group.

Grading Criteria (and methods used to determine final course grades)
The Grade Book is in Teams. A person's final course grade is an average (weighted based on the assignment percentages) of the assignments for this course. Projects in the course are not compared for grading purposes.

It is assumed that all students will attend and participate in all aspects of class. It is assumed you will participate in discussions and ask questions. If you are not participating on par with your peers, I will encourage you to do so in multiple ways. I will call on you. I will email you asking you questions. I will ask you to present. I will put you in a group of super talkers and tell them to flood you with questions. I will ask you to write things on the board. I used to give a participation grade, but I don't anymore, as it was unfair, but you should engage in the community of learners, as it helps you learn.

✓ Wk. 3: LLBA review	= 10% of grade
✓ Wk. 4 or 5: Meeting with Dr. Winke	= 10% of grade
✓ Wk. 7: "Proposal Outline" submission	= 10% of grade
✓ Wk. 7: Presentation	= 10% of grade
✓ Various weeks: *Task/activity lead (solo or pair/group)	= 15% of grade
✓ Wk. 14-15: Final project presentations	= 10% of grade
✓ December 14: Final exam project/deliverable due	= 35% of grade
	= 100% highest grade possible

Grades: 93-100% = A (4.0): Superior performance and/or effort
84-92% = B+ (3.5): Good performance and/or effort
76-83% = B (3.0): Adequate performance and/or effort
68-75% = C+ (2.5): Below adequate performance and/or effort
60-67% = C (2.0): Insufficient performance and/or effort
52-59% = D+ (1.5) Insufficient performance and/or effort
44-51% = D (1.0) Insufficient performance and/or effort
36-43% = D- (0.5) Insufficient performance and/or effort
Below 36% = F (0.0) Insufficient performance and/or effort

When special or unusual circumstances occur, I may postpone assignment of the student's final grade by use of an I-Incomplete, a DF-Deferred, or an ET-Extension marker. Normally, in graduate classes, when a student is obtaining a grade below a 3.5 for any reason, I and the student begin discussing whether an I, DF, or ET should be used, or if assignments need to be adjusted or customized to better align with their particular situation. Please see the MSU Academic Programs Catalog to learn more information about an I, DF, or ET and how they can be used and applied:
https://reg.msu.edu/academicprograms/Print.aspx?Section=528 Immediately talk to me if your grade falls below a 3.5. Please also note that under the Americans with Disabilities Act and Section 504 of the Rehabilitation Act of 1973, individuals with disabilities are guaranteed certain protections and rights of equal access to programs and services, including this course. The attainment of reasonable accommodations for this course and the assignments in this course can be assisted through a Verified Individualized Services and Accommodations (VISA) certificate. Please see MSU's Resource Center for Persons with Disabilities (RCPD) (https://www.rcpd.msu.edu) to receive aid in obtaining documentation of a disability or to obtain a needs assessment that will help you obtain an outline of reasonable accommodations.

Attendance Policy: Attendance is required, but students may opt to review class materials online and/or visit with me during office hours to discuss what was missed if class must be missed for health, wellness, or family reasons.

Please tell me in advance if you cannot attend or why you missed if you could not tell me ahead of time. I understand life is still very unpredictable these days, with illnesses and disruptions being quite common because of the (hopefully wanning or post) COVID-19 pandemic.

Required and supplemental course material: The textbook is "Language Assessment: Principles and Classroom Practices," 3rd Edition (2019) by H.D. Brown and P. Abeywickrama (Pearson). PDFs of the required readings will be provided via the course's Teams site. Supplemental readings are also provided via Teams.

Proctoring: There are no required proctoring arrangements to which students must adhere.