**Understanding Educational Testing**

**EDST 0213A**

**Syllabus: 20 August 2024 (subject to revision)**

# **I.** **Course/Instructor Information:**

Course Title: Understanding Educational Testing

Semester/Year: Fall 2024

Class meeting times: Tuesdays & Thursdays, 2:15 – 3:30

Class location: Axinn 105

Instructor: Steve Hoffman

Email: [hoffman@middlebury.edu](mailto:hoffman@middlebury.edu)

Office Hours [via Zoom](https://middlebury.zoom.us/j/5225642365?pwd=Tk00OWpmZ2dwbEFOVzNVMGs2K3Z0QT09):

* Tuesdays: 3:45 – 5:15 PM
* Thursdays: 9:15 – 10:45 PM
* By Appointment

# **II.** **Course Description:**

Achievement testing is now a cornerstone of education policy. It is also complex and routinely misunderstood by educators, policymakers, and the media. In this course students will use statistical methods to explore and address testing issues that arise in both policy and practice. We will examine the uses and abuses of educational assessment. We will examine and interrogate trends and group differences in achievement. And we will broaden our understanding of essential concepts of measurement, such as reliability, validity, and bias, by analyzing both large and small datasets. Prior experience with the statistical package “R” is not required, as learning this package will be part of the course. 3 hrs. lect.

# **III.** **Learning Outcomes:**

In this course, we will:

* critically examine the uses and abuses of educational assessment
* examine and interrogate trends and group differences in achievement
* broaden our understanding of essential concepts of measurement, such as reliability, validity, and bias
* gain facility employing statistical methods to explore and address testing issues that arise in both policy and practice
* use the statistical package “R” to analyze testing data sets
* communicate our understandings and findings clearly to the general public

# **IV.** **Course Materials and Required Texts:**

Koretz, D. (2008). *Measuring up: What educational testing really tells us.* Harvard University Press. (ebook OK)

Early on (on page 2), Professor Koretz noted that at the time he wrote this book there were many books on educational testing. Some were very pro-testing while others were anti-testing. Fifteen years later this is still true. Koretz intended to provide a balanced approach to explaining the field of educational testing to a wide audience, and this book is still required for courses on testing at other colleges and universities — both those for the “informed consumers of test scores” as well as for more technical courses. So, while this book is not new, and we will come across explanations that seem dated, it remains an excellent resource for us as we work through today’s issues around educational testing.

Popham, W. (2016). *The ABCs of Educational Testing*, Corwin Press. (ebook OK)

This book is an easier read than the Koretz text. Professor Popham insists at the start of this book that it’s about “the basics of educational testing” written so that readers can learn “a handful of foundational concepts and procedures linked to the testing that routinely goes on in our schools” (p. 1). We won’t follow the contents in the exact order he outlined, but we will use the whole text during the semester.

Wickham, H., Çetinkaya-Rundel, M. & Grolemund, G. (2022). *R for Data Science (2e)*. <https://r4ds.hadley.nz> (free to use)

This ebook is about how to “do” data analysis and data science using the “R” package, which you may have used in other courses at Middlebury. We will utilize only a portion of this book, though you may find it useful to bookmark this for work outside of this course. For those of you with no experience with R at all, don’t fret. I and many others here at Middlebury will help you learn the basic tools in R, allowing you to thrive in this course. Be sure to use the second edition (2e). This is a substantial rewrite from the first edition, and you’ll be learning some different coding techniques from the students in the previous iterations of EDST 0213.

# 

# **V.** **Course Structure:**

Lecture/discussion sessions are scheduled every Tuesday and Thursday from 2:15 – 3:30. Your active participation and engagement are essential to both *your* success and *your colleagues’* success in this course. Come prepared to participate each Tuesday and Thursday. The general format is a mixture of pre-recorded lectures, in-class activities, lectures, in-class discussion, asynchronous Canvas discussion, small-group work, partner projects, and opportunities for students to lead learning.

# **VI.** **Grading Information:**

Course grades will be based on a typical 100-point system. A = 93 – 100; A- = 90 – 92; B+ = 87 – 89; etc. Late work will likely be penalized.

* Attendance and participation – 20%
* Quizzes – 20%
* Discussion posts – 20%
* Data-Analytic Memos – 20%
* Final project – 20%

**Attendance and participation**

Attendance is required. Come to class on Tuesday and Thursday afternoons prepared and ready to participate. Students who must miss a class should notify Hoffman before the class meeting and arrange for alternate ways to contribute to the class community. (Pro tip: Don’t write “let me know if I missed anything.”)

Participation grades are calculated based on attendance, punctuality, contributions during small-group activities, and thoughtful engagement in discussions.

Note: During the first two weeks of November, students will (in small groups) take turns leading class, providing appropriate classroom activities, and authoring assessments.

**Quizzes**

Short, low-stakes quizzes to check understanding of the course content will occur often throughout the term. Dates for scheduled quizzes are listed on the Course Calendar, but there will be additional low-stakes quizzes throughout the term. Sometimes quizzes will be administered during face-to-face class sessions. Sometimes quizzes will be administered asynchronously through Canvas. *These are not meant to be high-pressure events.*

**Discussion posts**

Students will write or record your reflections in response to Canvas discussion prompts about designated assigned readings. Sometimes these will be written posts of 400 words or fewer. Other times you will be asked to compose a short video (typically about one minute long). Organize your thinking about the ideas and arguments made by the authors. If you reference articles, papers, books, or other media beyond the course texts, cite them appropriately. Make an informed critique of ideas, rather than merely summarizing the readings. Finally, read or watch your colleagues’ posts, and write a brief, thoughtful response to at least three (3) of your classmates’ posts. Posts are tentatively due on September 18, October 2, October 16, October 30, November 13, and December 4.

**Data-Analytic Memos**

Designed to develop and extend your data-analytic skills and to help you learn to communicate your findings clearly to others, these are collaborative assignments to be completed with a partner. Please engage in a full, fair, and mutually agreeable collaboration with your partner. Do not simply divide up the work. Discuss and plan the analyses together, debate what you have found with each other, and collaborate on the writing of your memo. My objective is to provide a high-quality opportunity for you to learn, review, teach, and communicate the course material to others. DAMs are tentatively due September 25, October 9, October 23, November 6, and November 20.

**Final project**

Due December 16.

Assigned in November, with an opportunity to draft an outline of your project before Thanksgiving, this project will serve as our final exam for the course. (There will be no scheduled in-person test during finals week.)

# **VII.** **Relevant Policies:**

**Academic Integrity**: As an academic community devoted to the life of the mind, Middlebury requires that every student complete intellectual honesty in the preparation and submission of all academic work. Details of our Academic Honesty, Honor Code, and Related Disciplinary Policies are available in Middlebury’s handbook.

**Honor Code Pledge**: The Honor Code pledge reads as follows: "I have neither given nor received unauthorized aid on this assignment." It is the responsibility of the student to write out in full, adhere to, and sign the Honor Code pledge on all examinations, research papers, and laboratory reports. Faculty members reserve the right to require the signed Honor Code pledge on other kinds of academic work.

**Disability access/accommodation**: Students who have Letters of Accommodation in this class are encouraged to contact their professor as early in the semester as possible to ensure that such accommodations are implemented in a timely fashion.

For those without Letters of Accommodation, assistance is available to eligible students through the [Disability Resource Center](https://www.middlebury.edu/office/disability-resource-center) (formerly called Student Accessibility Services). All discussions will remain confidential.

Please contact one of the ADA Coordinators at [ada@middlebury.edu](mailto:ada@middlebury.edu) for more information.

# **VIII.** **Expectations of Students:**

**Attendance:** Attendance is required. Come to class on Tuesday and Thursday afternoons prepared and ready to participate. Students who must miss a class should notify Professor Hoffman before the class meeting and arrange for alternate ways to contribute to the class community. (Pro tip: Don’t write “let me know if I missed anything.”)

**Participation:** Your active participation during class sessions is expected. Participation grades are calculated based on attendance, punctuality, contributions during small-group activities, and thoughtful engagement in discussions.

**Collaboration:** Collaborative assignments to be completed with one or more partners are required. Five Data-Analytic Memo assignments produced in partnership and a day leading class in November with a different group of colleagues require substantial work with your colleagues outside of the formal class sessions. Please engage in a full, fair, and mutually agreeable collaboration with your partners. Do not simply divide up the work.

**Late Work:** If you need an extension, please ask before the assignment is due and collaborate with me to come up with a reasonable plan for turning in late work. I generally assume that if you ask, you need an extension for a valid reason.

# **IX.** **Relevant Campus Resources:**

**Center for Teaching, Learning, and Research**: The CTLR provides academic support for students in many specific content areas and in writing across the curriculum through both professional tutors and peer tutors. The Center is also the place where students can find assistance in time- management and study skills. These services are free to all students. For more information on how to get the help you need, go to the [CTLR’s student resource page](http://www.middlebury.edu/academics/resources/ctlr/students).

[**Disability Resource Center**](https://www.middlebury.edu/office/disability-resource-center): The DRC provides support for students with disabilities and facilitates the accommodations process by helping students understand the resources and options available and by helping faculty understand how to increase access and full participation in courses. The DRC can also provide referrals for students who would like to undergo diagnostic testing. Students who are on financial aid and have never undergone diagnostic testing can apply to the CTLR for support to cover the cost of off-campus testing. DRC services are free to all students.

# **X.** **Course Calendar:**

**Dates for major assignments/projects (Quiz dates subject to change):**

* September 12: Syllabus Quiz
* September 18: Discussion Post
* September 19: Quiz on Coding in R4DS
* September 24: Quiz on Measurement
* September 25: DAM 1
* October 1: Quiz on Trends
* October 2: Discussion Post (Rosh Hashanah begins at sunset)
* October 9: DAM 2
* October 10: Quiz on Scales
* October 16: Discussion Post
* October 23: DAM 3
* October 24: Quiz on Error
* October 30: Discussion Post
* October 31: Quiz on Reliability and Validity
* November 5: Quiz (Student-led Assessment)
* November 6: DAM 4
* November 7: Quiz (Student-led Assessment)
* November 12: Quiz (Student-led Assessment)
* November 13: Discussion Post
* November 14: Quiz (Student-led Assessment)
* November 19: Quiz (Student-led Assessment)
* November 20: DAM 5
* November 21: Quiz (Student-led Assessment)
* December 4: Discussion Post
* December 16: Final Project Due

**Daily Schedule of readings/assignments (subject to minor changes):**

**September**

September 10: Introduction to the course

* *Measuring Up*, Prologue & Chapter 1
* *The ABCs of Educational Testing*, Preface & Chapter 1
* *R for Data Science 2e*, Welcome, Preface to 2nd edition

September 12: Fundamental Issues in Measurement

* *The ABCs of Educational Testing*, Chapter 2
* *Measuring Up*, Chapter 2
* *R for Data Science 2e*, Chapter 1
* Thorndike, Chapter 2, pp. 23 – 28, (Canvas file). This supplemental measurement textbook (*Measurement and Evaluation in Psychology and Education*, by Thorndike & Thorndike-Christ) includes a play data set for us to work with.

September 17: Formative Assessment & How we will learn coding

* *The ABCs of Educational Testing,* Chapters 8
* *R for Data Science 2e*, Chapters 2, 3 (ok to skim), 4, 5 (ok to skim), & 6

September 19: Coding, Projects, and Data-Analytic Memos

* *R for Data Science 2e*, Chapters 7 (ok to skim), 8, 20, 28, 29
* Thorndike, Chapter 2, pp. 28 – 38, (Canvas file). Just skim the instructions on how to perform various statistical procedures in SPSS and Excel; we will work in R.

September 24: Measurement concepts

* *Measuring Up*, Chapter 3
* Optional reading: Lindquist, E.F. (1951) (Canvas file)

September 26: Standardized Testing in the USA – PLUS coding with AI

* *The ABCs of Educational Testing*, Chapter 3
* [*Standards for Educational and Psychological Testing*](https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf) 2014, pp. 1 – 4
* *Measuring Up*, Chapter 4
* GitHub Copilot <https://docs.posit.co/ide/user/ide/guide/tools/copilot.html>
* Optional reading: Happy and GitHub for the useR <https://happygitwithr.com>

**OCTOBER**

October 1: Trends in standardized achievement tests

* *Measuring Up*, Chapter 5
* [NAEP Mathematics Assessment Highlights](https://www.nationsreportcard.gov/highlights/mathematics/2022/)
* [NAEP Reading Assessment Highlights](https://www.nationsreportcard.gov/highlights/reading/2022/)
* [NAEP Long-Term Trend Assessment Results](https://www.nationsreportcard.gov/ltt/)
* The Educational Opportunity Project <https://edopportunity.org>
* EOProject Methods <https://edopportunity.org/methods/>
* Thorndike, Chapter 2, pages 39 – 57, (Canvas file). Skim the instructions on how to perform various statistical procedures in Excel and SPSS.

October 3 (Rosh Hashanah): Achievement patterns and COVID

* *Measuring Up*, Chapter 6
* Kuhfeld, Megan, James Soland, and Karyn Lewis. (2022). [Test Score Patterns Across Three COVID-19-impacted School Years.](https://www.edworkingpapers.com/sites/default/files/ai22-521.pdf) (EdWorkingPaper: 22-521).

October 8: Performance standards

* Measuring Up, Chapter 8 (pages 179 – 200)
* Pages 1 – 2 & Pages 57 - 68 [New Meridian Technical Report 2020-2021](https://www.isbe.net/Documents/New-Meridian-Tech-Report-2020-21.pdf) (formerly PARCC)
* Optional reading: Hansche L., et al. (1999). Handbook for the Development of Performance Standards: Meeting the Requirements of Title I. Washington, D.C.: Council of Chief State School Officers. Chapters 1 (pages 3 – 6), 3 (pages 11 – 16 only), and 10 (pages 87 – 103). (Canvas file)

October 10: Scales

* *The ABCs of Educational Testing*, Chapter 7
* *Measuring Up*, Chapter 8 (pages 200 – 214)
* Thorndike, pp. 72 – 95 (Canvas). Skim the instructions on how to perform various statistical procedures in Excel and SPSS.

**FALL BREAK (October 11)**

October 15: NAEP – The Nation’s Report Card

* Familiarize yourself with the [NAEP website](https://nces.ed.gov/nationsreportcard/):
* Read from the NAEP Primer: [A Technical History of NAEP](https://nces.ed.gov/nationsreportcard/about/newnaephistory.aspx)
* NAEP Report Card: [2022 NAEP Mathematics Assessment](https://www.nationsreportcard.gov/mathematics/?grade=4)
* NAEP Report Card: [2022 NAEP Reading Assessment](https://www.nationsreportcard.gov/reading/?grade=4)
* For Reading and Math in both 4th and 8th grade, please study national average scores and national score gaps (Asian-White, Black-White, Hispanic-White, female-male, and Eligible-Not Eligible for National School Lunch Program).

October 17: Frameworks for NAEP

* [Mathematics Framework for the 2026 National Assessment of Educational Progress](https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/mathematics/2026-math-frameowork/NAEP-2026-Mathematics-Framework-Combined.pdf), Ch. 2, Mathematics Content, pp. 14 – 48. This chapter includes detailed tables of standards, but you don’t need to keep track of all of this. The goal is to get a sense of the breadth of the standards as a whole and the nature of the ‘chunks’ into which mathematics is broken into for the purpose of building the NAEP test. (Note: If you are curious, look at the archived copies of the math frameworks for 2017 or 2019. Most of the content of these frameworks remain in all versions.)
* [Reading Framework for the 2022 and 2024 NAEP](https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/reading/2022-nagb-reading-framework-508.pdf). 2, Content and Design of NAEP in Reading, pp. 17 – 46. (The details of the standards are less important than the big picture, but it’s important for you to understand the content breakdown and the assessment issues involved.)

October 22: International Comparisons

* [PISA 2018 Results: What students know and can do, Volume 1](https://www.oecd-ilibrary.org/docserver/5f07c754-en.pdf?expires=1659981276&id=id&accname=guest&checksum=F26222E03A67065F17AE708CF6262156), pages 15 – 30 & 41 – 84
* PISA 2022 Results: The State of Learning and Equity in Education, Executive Summary (pages 26 – 37) & What is PISA? (pages 38 – 42) & Variation in performance within countries (pages 60 – 65)
* <https://www.oecd.org/en/publications/pisa-2022-results-volume-i_53f23881-en/full-report.html>
* [PISA 2012 Released Mathematics Items](https://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf) (Please read the Forward and then scan through some of the items)
* Peruse the TIMSS & PIRLS website <https://timssandpirls.bc.edu>

October 24: Error

* *Measuring Up*, Chapter 7

October 29: Reliability

* *The ABCs of Educational Testing*, Chapter 5
* Read this to understand the general concepts and how they are conveyed: New Meridian Tech Report 2020-21 (Illinois state test), Pages 99 – 116 <https://www.isbe.net/Documents/New-Meridian-Tech-Report-2020-21.pdf>

October 31: Validity

* *Measuring Up*, Chapter 9
* *The ABCs of Educational Testing,* Chapter 4
* Skim (for about 20 minutes) the Validity chapter in the New Meridian Tech Report 2020-21 (Illinois state test), Pages 117 – 152 <https://www.isbe.net/Documents/New-Meridian-Tech-Report-2020-21.pdf>

**November**

November 5: Score Inflation

* *Measuring Up*, Chapter 10
* Holcomb, R., Jennings, J. L., & Koretz (2013). The roots of score inflation: An examination of opportunities in two states’ tests. In G. Sunderman (Ed.), Charting Reform, Achieving Equity in a Diverse Nation. Information Age Publishing. (Canvas file)

November 7: Adverse impact and bias

* *The ABCs of Educational Testing,* Chapter 6
* *Measuring Up*, Chapter 11
* Rosales, J. & Walker, T. (2021). The racist beginnings of standardized tests. *NEA News.* <https://www.nea.org/advocating-for-change/new-from-nea/racist-beginnings-standardized-testing>
* Papay, J., Mantil, A., & Murnane, R. (2021). [On the threshold: Impacts of barely passing high-school exit exams on post-secondary enrollment and completion.](https://www.edworkingpapers.com/sites/default/files/ai22-627.pdf) (EdWorkingPaper: 22-627).

November 12: College Admissions *(Student led, readings may be revised)*

* *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College* <https://www.supremecourt.gov/opinions/22pdf/20-1199_hgdj.pdf>
* Lewis Jr., N. (2023). Are Standardized Tests Racist, or Are They Anti-racist? *The Atlantic.* <https://www.theatlantic.com/science/archive/2023/01/should-college-admissions-use-standardized-test-scores/672816/?utm_source=copy-link&utm_medium=social&utm_campaign=share>
* Koretz, D., Yu, C., Mbekeani, P., Langi, M., Dhaliwal, T., and Braslow, D. (2016). Predicting freshman grade-point average from college-admissions and state high-school test scores. AERA Open, 2(4), 1-13.
* Westrick, P. A., et al. (2019). [National SAT Validity Study.](https://satsuite.collegeboard.org/media/pdf/national-sat-validity-study.pdf) The College Board.
* Marini, J. P., at al. (2019). [Validity of SAT Essay Scores for Predicting First-Year Grades](https://research.collegeboard.org/media/pdf/validity-sat-essay-scores-predicting-first-year-grades.pdf%20The%20College%20Board.).
* [Report of the UC Academic Council Standardized Testing Task Force](https://senate.universityofcalifornia.edu/_files/underreview/sttf-report.pdf) (STTF). (2020). Executive Summary, pp. 3 – 7

November 14: Cheating *(Student led, readings may be revised)*

* Aviv, R. (2014). Wrong answer: In an era of high-stakes testing, a struggling school made a shocking choice. *The New Yorker*. July 21, 2014. (Canvas)
* Merrow, J. (2013). [Michelle Rhee’s reign of error](https://themerrowreport.com/2013/04/11/michelle-rhees-reign-of-error/). *The Merrow Report*.
* Flanagan, C. (April 4, 2019). [They had it coming.](https://www.theatlantic.com/ideas/archive/2019/04/what-college-admissions-scandal-reveals/586468/) *The Atlantic*.
* Janowski, E. (May 15, 2021). [Cheating investigation embroils Geisel in controversy](https://www.thedartmouth.com/article/2021/05/cheating-investigation-embroils-geisel-in-controversy). *The Dartmouth*.
* Morey, A. (June 10, 2021). [Dartmouth drops cheating charges against med students, apologizes for flawed investigation.](https://www.thefire.org/dartmouth-drops-cheating-charges-against-med-students-apologizes-for-flawed-investigation/) *The Fire*.
* Levenson, M. (2022). [Hoping to Identify Cheaters, A Professor Sues His Own Students](https://www.nytimes.com/2022/03/17/us/chapman-law-cheating-professor.html?referringSource=articleShare). *The New York Times*.
* Goldstein, M. (2022). [Ernst & Young to Pay $100 Million Fine After Auditors Cheated on Ethics Exams](https://www.nytimes.com/2022/06/28/business/ernst-young-sec-cheating.html?referringSource=articleShare). *The New York Times*.

November 19: Testing Special Populations

* *Measuring Up*, Chapters 12 & 13

November 21: Performance Assessment

* [Sample performance assessment items](https://sampleitems.smarterbalanced.org/) from Smarter Balanced
* SBAC “Penny Argumentative Performance Task” (Grade 8) (Canvas)
* [Performance Task Writing Rubric](https://portal.smarterbalanced.org/library/en/performance-task-writing-rubric-explanatory.pdf) (Grades 6-11)

**THANKSGIVING BREAK**

**December**

December 3: Assorted school-based topics

* *The ABCs of Educational Testing,* Chapters 9, & 10

December 5: Course Review and Final Project workshop

**Additional Resources:**

* Alvero, AJ., Giebel, S., Gebre-Medhin, B., Antonio, A.L., Stevens, M.L., & Domingue, B.W. (2021). Essay Content is Strongly Related to Household Income and SAT Scores: Evidence from 60,000 Undergraduate Applications. (CEPA Working Paper No.21-03). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp21-03
* Bitler, M. Corcoran, S. Domina, T. & Penner, E. (2019). Teacher effects on student achievement and height: a cautionary tale. *NBER Working Paper* 26480.
* Chetty, R., Deming, D. J., Friedman, J. N. (2023). Diversifying society’s leaders? The Determinants and causal effects of admission to highly selective private colleges. <https://opportunityinsights.org/wp-content/uploads/2023/07/CollegeAdmissions_Paper.pdf>
* Classic Learning Test. 2018 Technical Report: <https://www.cltexam.com/wp-content/uploads/2022/04/CLT_Technical_Report.pdf>
* Geiser, S. (2015). The growing correlation between race and SAT scores: New findings from California. <https://escholarship.org/uc/item/9gs5v3pv>
* Haertel, E. H. (1999). Performance assessment and education reform. *Phi Delta Kappan, 80*(9), 662-666. (Canvas)
* Harden, K. P. (2022). The SAT isn’t what’s unfair. *The Atlantic.* <https://www.theatlantic.com/ideas/archive/2022/04/mit-admissions-reinstates-sat-act-tests/629455/?utm_source=copy-link&utm_medium=social&utm_campaign=share>
* Hansche L., et al. (1999). Handbook for the Development of Performance Standards: Meeting the Requirements of Title I. Washington, D.C.: Council of Chief State School Officers. Chapters 1 (pages 3 – 6), 3 (pages 11 – 16 only), and 10 (pages 87 – 103). (Canvas file)
* Jacob, B. & Rothstein, J. (2016). [The measurement of student ability in modern assessment systems.](https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.30.3.85) *Journal of Economic Perspectives, 30*(6), 85-108
* Joint Committee on the Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education <https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf>
* National Center for Education Statistics (2023). Number, percentage distribution, and SAT mean scores of high school seniors taking the SAT, by sex, race/ethnicity, first language learned, and highest level of parental education: Selected years, 2017 through 2022. <https://nces.ed.gov/programs/digest/d22/tables/dt22_226.10.asp>
* SAT Suite of Assessments Technical Manual (2017) <https://satsuite.collegeboard.org/media/pdf/sat-suite-assessments-technical-manual.pdf>