EDUC19:
EDUCATIONAL TESTING

Spring Term, 2020
12: Mon, Wed, Fri @ 12:50 pm – 1:55 pm
X-hr: Tuesday @ 1:20 pm – 2:10 pm
Kemeny 008

Professor: David J.M. Kraemer, PhD
Contact: *david.kraemer@dartmouth.edu*
Office Location: Raven House 212
Office Hours: Friday @ 2:00 pm – 4:30 pm

COURSE OVERVIEW

"The test" has become the ubiquitous measure of performance for evaluating educational achievement in the US and in many other nations. Academic test scores are used by teachers to evaluate what students have learned, by administrators to evaluate the performance of teachers, by the government to evaluate the performance of school districts, and by parents to decide where to raise a family. But how we should use tests is still a hotly debated question, with real consequences for the entire educational system.

This course is focused on practical and theoretical issues regarding how we evaluate academic performance and how various factors influence test scores. The term is divided into two major sections:

1. The first section focuses on achievement testing. What kinds of tests are used in school to measure learning? How do teachers, administrators, policy developers, and students use these tests? What are the key concepts in understanding what a test measures and how it should be interpreted? What can (and can't) a test tell us about the achievement of a student, of a teacher, of a district, and of a nation?

2. The second section of the course focuses on aptitude testing. What makes one student different from another? What factors determine success in a school or who should be admitted to a certain college? Which aspects of one's cognition, personality, and development are relevant to educational performance? How should we deal with these individual differences?

From this course, students will gain a broad understanding of the lay of the educational assessment landscape, and a deep understanding of the promises and perils of academic testing. Developing the skills of reading research and opinion articles, discussing and understanding different interpretations of data, and writing for a scientific audience are central to these goals and therefore are critical components of the course.

GRADING OVERVIEW

30% Midterm Exam
32% Final Exam
15% Class Debates and Position Paper
10% Group Research Paper (Social Impact Practicum)
10% Quizzes (average of all 6)
3% Class Participation and Attendance
.5% Research Participation *(optional extra credit)*

ASSESSMENTS and ASSIGNMENTS

Mid-term examination – ***Friday, May 3rd***

- Mix of short answer and fill-in-the-blank questions
- Covers all assigned material in the current unit (readings, slides, presentations)

Final examination – ***Sunday, June 2nd @ 3:00 pm***

- Mix of short answer and fill-in-the-blank questions
- Covers all material from the course
- Emphasis on material since second midterm

Pop Quizzes

Six quizzes throughout the term:

- Will consist of short answer and fill-in-the-blank questions
- These are intended to be low-stakes opportunities for both you and me to gauge your understanding of the material

Class Participation and Attendance

- Arrive on time for each class
- Prepare for all class discussions and be an active class member (e.g., read the assigned materials, stay awake during class, stay off internet, etc.)

ASSESSMENTS and ASSIGNMENTS *(continued)*

Position Papers

The issues surrounding educational testing – as with many aspects of education – are complicated, emotionally charged, and have real-world consequences for students across the country. Therefore, an important skill to develop is the ability to distill a large body of opinions and research on a given topic down to a concise set of practical actions that can be argued for and implemented. Accordingly, in this course you will practice doing just that, in the form of writing short position papers (~ 1,000 words) which present a clear case for a specific course of action regarding current topics in educational testing.

For the debate teams, a short position paper is due after each debate as the follows:

- **#1: *Should every state use the same standards-based tests?***
- **#2: *Should teachers be evaluated based on their students test scores?***
- **#3: *Should we place 6th graders on academic tracks?***

The position paper assignment is due after each debate **for the teams who are debating that topic**. The full description of the assignment for each paper can be found on Canvas. For each position paper I'm looking for a well-researched, well-reasoned essay relating to your take on each topic, of roughly 1,000 words in length (about 3 double-spaced pages). Additionally, each paper must reference at least 4 sources (peer-reviewed journal articles, news pieces, etc.) that were not assigned in class; and remember: always cite your sources when you assert a fact. Each position paper is an individual assignment (not a group project), and will be graded as such – *you are expected to do your own research and convey your own assessment and understanding of the material*

Group Research Paper

Education researchers have identified a need for an evidence-based general method for identifying practical and quantifiable indicators to measure improvement in specific courses at specific grade levels (e.g., 8th grade math, 5th grade reading, etc.). The goal of this project is to understand the process by which one develops high-leverage and high-impact measures in real time without waiting many years after a study comes out to assess progress. In order to achieve this goal, you and your group will perform a literature review on measures of learning and effective teaching, and write a summary report including implementable suggestions for teachers in the classroom today.

You will work with your group to review the literature on evaluation of learning in your specific focus area (e.g., elementary math, secondary English Language Arts, or on of the WCSU goals for graduates), then you will generate suggestions for the educators to help develop measures to test for achievement in that domain.

Each group will submit a ~15-20 page paper that includes a well-researched description on the research about topic-specific teaching practices and useful suggestions for teachers to measure their own effectiveness in class. During the last week of the term, the groups will share their work with the class

ASSESSMENTS and ASSIGNMENTS *(continued)*

In-Class Debates

Frequently, educators, administrators, and policy-makers must deal with tricky issues and decide on the best path forward based on the available evidence. That evidence will never be perfect, and on many topics reasonable people will disagree. Therefore, it is important to develop the ability to synthesize information and carry out an informed discussion in which you attempt to convince others of your position and listen to their perspective and opinions as well. Accordingly, the purpose of this assignment is to research a given topic relating to a current issue in education, and to participate in an in-class debate in which you argue one side of that issue convincingly.

In three separate class periods, two groups will debate a given motion – one group arguing for the motion and one group arguing against the motion:

- **The first debate (Friday, April 12th) will focus on the motion: *"All states should adopt a national standards-based test for math and English Language Arts."***

- **The second debate (Friday, April 26th) will focus on the motion: *"Teachers should be evaluated based on their students test scores."***

- **The third debate (Monday, May 20th) will focus on the motion: *"Students would benefit from secondary-level academic tracking (beginning in 6th grade)."***

- Each group will research the assigned topic and, as a group, decide on the best evidentiary and logical support for their designated position (for or against)
- Each group will divide the relevant aspects of their argument amongst their members such that *all team members except one* will make an opening statement reflecting a different aspect of their team's overall argument
- Following the opening statements, class members who are not debating (i.e., the audience), will each ask a question directed at one or both teams
- Finally, the team member on each team who *did not* make an opening statement will make a closing statement summarizing his or her team's position and incorporating points that were raised during the debate
- At the start of the debate, I will poll the audience to determine how many students agree with the motion prior to the debate. I will poll again afterwards to see if anyone was persuaded by the debate to change his or her mind. The team who persuades the most audience members to choose their side wins the debate. In other words, the goal is to use facts and reasoned argument to change minds.
- Prior to the debate, each team will post 1 reading to Canvas that they deem the most pertinent for the audience to read in advance. These readings are to be posted on Canvas no later than 1 week preceding the debate.
- *A position paper related to the motion will be due for each team following the debate*

ASSESSMENTS and ASSIGNMENTS *(continued)*

Research Participation (optional extra credit)

The goal of this optional assignment is to provide you with an insightful perspective on how research studies are conducted.

Volunteer to participate in lab research in the Education Department (one study)

Volunteer to participate in one of the research studies conducted by any research lab in the Education Department. Specific instructions on how to sign up for a research study are posted on Canvas. To get credit, you must hand in a signed *EDUC51 Participant Confirmation Form* (one signed copy for each study in which you participate) to me in class anytime before ***Wednesday, May 29th.*** A copy of the form is posted on Canvas.

GENERAL POLICIES

1. **Read all materials and prepare for class.** You are expected to read the materials posted on Canvas *before* each class. Be prepared to discuss that material *in class*. Everyone is expected to come to every class and to arrive on time. You are also expected to contribute to class discussion. You will learn the material better and others will learn from you. The success of this course depends on everyone coming to class prepared and ready to discuss the material. Both attendance (on-time) and preparation for class will determine a portion of your grade (see "Assignments and Assessments" below).
2. **Before you turn in your papers…** make sure that you use 12-point Times New Roman font, that you double-space the whole document, that your print margins are 1-inch on all sides (not the default in *Word)*, that all your pages are numbered, and that your document is stapled together (if printed). For citations in all papers, you must use APA Style formatting (refer to the APA Style Manual or online guides, such as: **http://owl.english.purdue.edu/owl/resource/560/01/**)
3. **Tell me sooner rather than later.** If you know ahead of time that you will be missing a class, e.g., for sports, please let me know in advance in order to avoid losing participation credit. Some students may wish to take part in religious observances that occur during this academic term. If you have a religious observance that conflicts with your participation in the course, please meet with me before the end of the second week of the term to discuss appropriate accommodations.
4. ***ASSUME THAT I WILL NOT ACCEPT LATE ASSIGNMENTS.***
5. **Cell phones are not to be used in class.** If an emergency arises that requires the use of a phone, please quietly excuse yourself from the room to respond.
6. **Accommodations.** Students with learning, physical, or psychiatric disabilities enrolled in this course who may need disability-related classroom accommodations are encouraged to make an office appointment to see me early in the semester (i.e., within the first two weeks). If you have not done so already, students requiring disability-related accommodations should register with the Student Accessibility Services office (301 Collis Student Center).
   Dartmouth's policies and resources: **http://www.dartmouth.edu/~accessibility**
   Contact info: 646-9900, **Student.Accessibility.Services@Dartmouth.edu**
7. **Plagiarism is unacceptable.** All work submitted as your own must be written by you and not previously submitted for any other class. It is important to attribute material that is the work of others to the original source. If you are unsure how to properly cite a source, please consult with me prior to handing in an assignment (and see: **http://www.dartmouth.edu/~writing/sources/**). You should be familiar with Dartmouth's Honor Principle, which applies to all courses at Dartmouth (available here: **www.dartmouth.edu/~uja/honor/**). I do not expect any violations of this code, but if any concerns do arise I will forward all related materials to Dartmouth's Committee on Standards.

SCHEDULE of TOPICS
(ASSIGNED READINGS ARE POSTED ON CANVAS)

1. Wednesday, March 27

Introduction and course overview; getting to know you

SECTION I: MEASURING ACHIEVEMENT

2. Friday, March 29

Evolution of testing in the US: *how did we get here?*

Readings:

- The Evolution of American Testing (pp. 46-73), *in* Koretz, D. M. (2009). *Measuring up: What educational testing really tells us*. Harvard University Press.
- Lee, J. (2008). Is Test-Driven External Accountability Effective? Synthesizing the Evidence From Cross-State Causal-Comparative and Correlational Studies. *Review of Educational Research*, *78*(3), 608–644. doi: 10.3102/0034654308324427

3. Monday, April 1

Current policies and issues in standardized testing: *why are there so many tests?*

Readings:

- Excerpt from "The Test: Why our schools are obsessed with standardized testing but you don't have to be" by Anya Kamenetz
- No Child Left Behind Has Finally Been Left Behind. (2015) *US News and World Report.*
- Wiliam, D. (2010). Standardized Testing and School Accountability. *Educational Psychologist*, *45*(2), 107–122. doi: 10.1080/00461521003703060

4. Wednesday, April 3

Test construction and validation: *what is a test score anyway?*

Readings:

- What *Is* a Test?; and What We Measure (pp. 16-45), *in* Koretz, D. M. (2009). *Measuring up: What educational testing really tells us*. Harvard University Press.
- Giving Meaning to Scores, *excerpt* (pp. 66-108), *in* Thorndike, R. M. (2010). *Measurement and evaluation in psychology and education (8th ed.).* Boston: Prentice Hall.
- Error and Reliability (pp. 143-179), *in* Koretz, D. M. (2009).

5. Friday, April 5

Side-effects of test-based accountability: *what's the deal with grade inflation?*

Readings:

- What the Ivies can learn from Wellesley. (2014). *The Economist.*

- Koretz, D. (2005). *Alignment, high stakes, and the inflation of test scores*. National Center for Research on Evaluation, Standards, and Student Testing, Center for the Study of Evaluation, Graduate School of Education & Information.
- Carey, T., & Carifio, J. (2012). The Minimum Grading Controversy: Results of a Quantitative Study of Seven Years of Grading Data From an Urban High School. *Educational Researcher*, *41*(6), 201–208. https://doi.org/10.3102/0013189X12453309

6. Monday, April 8

Portfolios: *a better assessment and the solution to grade inflation?*

Readings:

- Koretz, D. (1998). Large-scale Portfolio Assessments in the US: evidence pertaining to the quality of measurement. *Assessment in Education: Principles, Policy & Practice*, *5*(3), 309–334. doi: 10.1080/0969595980050302
- Burner, T. (2014). The potential formative benefits of portfolio assessment in second and foreign language writing contexts: A review of the literature. *Studies in Educational Evaluation*, *43*, 139–149. https://doi.org/10.1016/j.stueduc.2014.03.002
- Chang, C.-C., Liang, C., & Chen, Y.-H. (2013). Is learner self-assessment reliable and valid in a Web-based portfolio environment for high school students? *Computers & Education*, *60*(1), 325–334. https://doi.org/10.1016/j.compedu.2012.05.012

7. Wednesday, April 10

Formative assessments: *how should we use homework and quizzes?*

Readings:

- Wimmer, H., Powell, L., Kilgus, L., & Force, C. (2017). Improving Course Assessment via Web-based Homework. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, *7*(2), 1–19. https://doi.org/10.4018/IJOPCD.2017040101
- McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, *103*(2), 399–414. https://doi.org/10.1037/a0021782
- Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, *1*(3), 181–210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

8. Friday, April 12

Debate #1: *All states should adopt a national standards-based test for math and English Language Arts.*

Readings:

- TBD

9. Monday, April 15

High school graduation assessments: *what should be required to graduate?*

Readings:

- Jones, B. (2017). In controversial move, LAUSD's Deasy wants to raise high-school graduation requirements. *Los Angeles Daily News*.
- Dounay, J. (2008). Dispelling the Myths About the Negative Effects of Raising High School Graduation Requirements. *Education Commission of the States Policy Brief*.
- Palardy, G. J. (2013). High School Socioeconomic Segregation and Student Attainment. *American Educational Research Journal*, 50(4), 714–754. https://doi.org/10.3102/0002831213481240
- Schiller, K. S., & Muller, C. (2003). Raising the bar and equity? Effects of state high school graduation requirements and accountability policies on students' mathematics course taking. *Educational Evaluation and Policy Analysis*, *25*(3), 299–318.

10. Wednesday, April 17

Extending the model: *how does your community influence your test scores?*

Readings:

- Helping Children Thrive. (2016, June 11). *The New York Times*.
- Oreopoulos, P., Brown, R. S., & Lavecchia, A. M. (2014). *Pathways to education: An integrated approach to helping at-risk high school students*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w20430
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*(6), 645–662. https://doi.org/10.1016/j.appdev.2003.09.002

11. Friday, April 19 and 12. Monday, April 22

Assessing the assessors: The teacher evaluation controversy -
how are teachers evaluated on the job and is there a better way?

Readings:

- Controversial teacher evaluation method is on trial — literally — and the judge is not amused. (2015) *The Washington Post*.
- Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-Added Modeling of Teacher Effectiveness: An Exploration of Stability across Models and Contexts. *Education Policy Analysis Archives*, *18*(23), n23.
- Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation*. National Bureau of Economic Research.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating Teacher Evaluation. *Phi Delta Kappan*, *93*(6), 8–15.
- Ho, A. D., & Kane, T. J. (2013). The Reliability of Classroom Observations by School Personnel. Research Paper. MET Project. *Bill & Melinda Gates Foundation*. Retrieved from http://eric.ed.gov/?id=ED540957

- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood*. National Bureau of Economic Research.
- Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. Research Paper. MET Project. *Bill & Melinda Gates Foundation*. Retrieved from http://eric.ed.gov/?id=ED540959

13. Wednesday, April 24

TIMMS and PISA: *how does the US rate on international tests?*

Readings:

- PISA's Worldwide Lessons in Equity and Excellence. (2016). *Alliance for Education.*
- Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher Quality, Opportunity Gap, and National Achievement in 46 Countries. *Educational Researcher*, *36*(7), 369–387. https://doi.org/10.3102/0013189X07308739
- Peterson, P. E., Woessmann, L., Hanushek, E. A., & Lastra-Anadón, C. X. (2011). Globally Challenged: Are US Students Ready to Compete? The Latest on Each State's International Standing in Math and Reading. PEPG 11-03. *Program on Education Policy and Governance, Harvard University*. Retrieved from http://eric.ed.gov/?id=ED526954

14. Friday, April 26

Debate #2: *Teachers should be evaluated based on their students test scores.*

Readings:

- TBD

SECTION II: MEASURING APTITUDES, ABILITIES, AND TRAITS

15. Monday, April 29

SAT, ACT, and other college entrance exams: *what do they actually predict?*

Readings:

- Maruyama, G. (2012). Assessing College Readiness: Should We Be Satisfied With ACT or Other Threshold Scores? *Educational Researcher*, *41*(7), 252–261.
- Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2012). The Role of Socioeconomic Status in SAT-Grade Relationships and in College Admissions Decisions. *Psychological Science*, *23*(9), 1000–1007. https://doi.org/10.1177/0956797612438732 3
- Scott-Clayton, J. (2012). *Do High-Stakes Placement Exams Predict College Success? CCRC Working Paper No. 41*. Community College Research Center.

16. Wednesday, May 1

HS entrance exams and magnet schools: *do they promote or reduce equal access?*

- Figlio, D. N., & Page, M. E. (2002). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics*, *51*(3), 497–514.
- Dobbie, W., & Fryer, J. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, 6 (3), 58-75.
- Abdulkadiroglu, A., Angrist, J. D., & Pathak, P. A. (2011). *The Elite Illusion: Achievement Effects at Boston and New York Exam Schools* (Working Paper No. 17264). National Bureau of Economic Research. doi: 10.3386/w17264

17. Friday, May 3

*** **MIDTERM EXAM** ***

18. Monday, May 6

Group Project Day

- Discuss Social Impact Practicum

19. Wednesday, May 8

Placing students onto different academic "tracks": *does grouping by ability work?*

Readings:

- Thiemann, K. (2017). Ability tracking or comprehensive schooling? A theory on peer effects in competitive and non-competitive cultures. *Journal of Economic Behavior & Organization*, *137*, 214–231. https://doi.org/10.1016/j.jebo.2017.03.005
- Gamoran, A. (1992). The Variable Effects of High School Tracking. *American Sociological Review*, *57*(6), 812–828.
- Cortes, K. E., & Goodman, J. S. (2014). Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement.
- Clark, D., & Del Bono, E. (2016). The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom. *American Economic Journal: Applied Economics*, *8*(1), 150–76. https://doi.org/10.1257/app.20130505

20. Friday, May 10

Group Project Day

- Guest speaker to discuss Social Impact Practicum

21. Monday, May 13

Vocational tracking: *should we steer students towards certain career paths?*

Readings:

- Bol, T., & Van de Werfhorst, H. G. (2013). The measurement of tracking, vocational orientation, and standardization of educational systems: A comparative approach. *Gini Discussion Paper 81*.

- Eichhorst, W., Rodríguez-Planas, N., Schmidl, R., & Zimmermann, K. F. (2015). A Road Map to Vocational Education and Training in Industrialized Countries. *ILR Review*, *68*(2), 314–337. https://doi.org/10.1177/0019793914564963
- Wiswall, M., Stiefel, L., Schwartz, A. E., & Boccardo, J. (2014). Does attending a STEM high school improve student performance? Evidence from New York City. *Economics of Education Review*, *40*, 93–105.

22. Wednesday, May 15

Personality and Intelligence testing: *does is it tell us about educational aptitude?*

Readings:

- Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*(3), 441–451. doi: 10.1016/j.paid.2006.08.001
- Pittenger, D. J. (1993). The utility of the Myers-Briggs type indicator. *Review of Educational Research*, *63*(4), 467–488.
- Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21. doi: 10.1016/j.intell.2006.02.001
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*(1), 83–92.

Friday, May 17

Class not scheduled for today

23. Monday, May 20

DEBATE #3: *"Students would benefit from secondary-level academic tracking (beginning in 6th grade)."*

Readings:

- TBD

24. Tuesday, May 21 — x-hour

More about intelligence: *alternative models and "non-cognitive" factors*

Readings:

- Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: predicting retention in the military, the workplace, school and marriage. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00036
- Sternberg, R. J. (1999). Successful intelligence: Finding a balance. *Trends in Cognitive Sciences*, *3*(11), 436–442.

- Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond g: Putting multiple intelligences theory to the test. *Intelligence*, *34*(5), 487–502. https://doi.org/10.1016/j.intell.2006.02.004

25. Wednesday, May 22

Academic anxiety: *how do emotions and anxiety influence our test scores?*

Readings:

- Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, *11*(5), 181–185.
- Jamieson, J. P., Peters, B. J., Greenwood, E. J., & Altose, A. J. (2016). Reappraising Stress Arousal Improves Performance and Reduces Evaluation Anxiety in Classroom Exam Situations. *Social Psychological and Personality Science*, 1948550616644656. https://doi.org/10.1177/1948550616644656
- Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science*, *331*(6014), 211–213.

26. Friday, May 24

Group presentations: day 1

27. Monday, May 27

Memorial Day

- No class held today

28. Wednesday, May 29th

Group presentations: day 2

Sunday, June 2 @ 3pm: * FINAL EXAM *****