Exhibit 6

IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF FLORIDA
TALLAHASSEE DIVISION

LEROY PERNELL et al., Plaintiffs, v. FLORIDA BOARD OF GOVERNORS OF THE STATE UNIVERSITY SYSTEM et al., Defendants.	Case No.: 4:22-cv-304

DECLARATION OF DR. RUSSELL G. ALMOND IN SUPPORT OF PLAINTIFFS' MOTION FOR A PRELIMINARY INJUNCTION

I, Dr. Russell Almond, hereby declare and state as follows:

A. Background

1. My name is Russell Almond. I am over 18 years of age and identify as a white man.
2. I have personal knowledge of the following facts and if called to testify could and would competently do so.
3. Since 2010, I have worked as an associate professor of Measurement and Statistics in the Department of Educational Psychology and Learning Systems at Florida State University in Tallahassee, Florida.

4. I received a Bachelor of Science degree in Mathematics from California Institute of Technology, a Master of Arts degree in Statistics from Harvard University, and a Doctor of Philosophy degree in Statistics from Harvard University.

5. After graduating from Harvard, I worked as a research scientist in the StatSci Division of MathSoft, Inc.; as a research associate for RRA, Inc.—a contractor at the Los Alamos National Laboratory; and as a research scientist and senior research scientist in various groups at the Educational Testing Service (ETS). I also served as a visiting lecturer in the Department of Statistics at the University of Pennsylvania's Wharton School of Business and as an acting assistant professor in the Department of Statistics at the University of Washington.

6. My research interests focus on the question of how to gather, track, and monitor evidence of student growth using both traditional sources (e.g., assessments and homework) and non-traditional sources (e.g., simulations and games), especially in situations where multiple dimensions of student proficiency are considered. I am known for my work in Evidence Centered Design, a methodology for building assessments that focus on how observations of student performance provide evidence of knowledge, skills, abilities, and other psychological traits and states.

7. I have earned numerous awards, honors, and accolades, including the National Council on Educational Measurement Award for Outstanding Scientific or Technical Contribution to Educational Measurement in 2000.

8. I currently serve as a director on the board of the Florida Educational Research Association. I am also a member of the National Council on Measurement in Education, Psychometric Society, Association for Computing Machinery, American Statistical Association, and American Educational Research Association.

9. In the upcoming school year, I will be teaching Basic Descriptive and Inferential Statistics Applications, Missing Data Analysis, Bayesian Data Analysis, and Educational Psychology Colloquium. I teach Basic Descriptive and Inferential Statistics Applications and lead the Educational Psychology Colloquium every semester, and in alternate years, I teach various seminars on Applied Statistics and Scale and Instrument Design. My courses are entirely geared towards students seeking doctorates and master's degrees—primarily in the College of Education, with some students from other colleges, usually in social science programs.

B. The Interaction Between House Bill 7 (H.B. 7)[1] and the Content of My Courses

10. On April 22, 2022, Governor Ron DeSantis signed H.B. 7 into law. This law in relevant part amended § 1000.05(4), Fla. Stat., to add a list of eight prohibited concepts that instructors are permitted to denounce but are not permitted to advance.

11. I understand H.B. 7 to limit my ability to instruct my students on theories arguing that past patterns of discrimination are contributing causes to current discrepancies in achievement between different racial groups, including critical race theory. I feel that persons pursuing careers in educational research must be aware of and understand these theories.

12. Additionally, I understand the eighth prohibited concept in H.B. 7 to substantially restrict my ability to discuss topics that are relevant to what I teach. This prohibited concept limits negative statements about merit, excellence, and hard work. Although these may make a strong difference in the achievement of individuals, they are not important factors when looking at the performance of groups, where differences in talent and effort should average out. However, it has often been insinuated (without supporting

[1] Ch. 2022-72, Laws of Fla.

evidence) that differences in "merit" or "effort" could account for observed racial differences. This echoes long-discredited theories of eugenics.

13. The eighth prohibited concept also restricts criticism of "colorblind" policies. But it is well documented that colorblind policies are not sufficient to achieve colorblind results.[2] Yet, H.B. 7 seems to limit my ability to articulate that viewpoint. Consequently, I interpret H.B. 7 as compelling me to not deviate from the state-mandated viewpoint on "colorblindness," despite it being contrary to what I believe the evidence supports.

14. With regard to my courses, H.B. 7 will directly impact three of them: Basic Descriptive and Inferential Statistics Applications (EDF 5400), Scale and Instrument Design (EDF 5448), and Educational Psychology Colloquium (EDF 5922).

15. The Basic Descriptive and Inferential Statistics Applications course is an introductory course providing students with the skills necessary to interpret observational and experimental data. In particular, students must learn to properly interpret models where race is included as either an explanatory variable or covariate. My other statistics courses explore more advanced

[2] I am defining *colorblind results* as racial distribution in a category of high-status positions being consistent with the racial distribution of the candidate pool (within the limits of sampling variability).

techniques for more complex studies, but the necessity of interpreting models that include race is critical to applying these models in the social science realm.

16. My Scale and Instrument Design course presents a different challenge. Here, the students are taught to build tests and psychological assessments to the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) joint *Standards for Educational and Psychological Assessment*.[3] This includes specific procedures to ensure that test takers are not advantaged or disadvantaged because of their race, as well as methods for testing this.

17. Finally, the Educational Psychology Colloquium is a series of seminars meant to provide professional training for psychometricians that does not fit into the normal coursework. In particular, this is knowledge these students would need to know before taking jobs in testing companies; national, state and local departments of education; and research positions involving educational research. As over half of our students are international, this often

[3] American Educational Research Association et al., *Standards for Educational and Psychological Testing* (2014), available at https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf.

means explaining some of the complexities of the history of race in the United States.

18. The purpose of the Educational Psychology Colloquium series is to provide the Measurement and Statistics students with many of the soft skills they need to function as professionals in the testing and educational policy industries. This includes instruction on how to talk about sensitive topic like race without arousing the prejudices and passions of a racially or politically mixed audience. Many of the students are international students, and thus they are mostly unaware of the cultural and political baggage that race has in the United States. As many of them will go on to jobs in the United States, they need to understand all of the political perspectives on race, not just those perspectives favored by the current administration.

19. Race is frequently used as an explanatory variable or covariate in social science (including education) models. In all of my statistics classes, starting from the introductory ones, it is important that my students be aware of the research surrounding race and its effect on educational and other economic outcomes. This includes frank discussions of how past patterns of discriminatory laws and practices affect observed differences between racial groups. This is documented in the text *Reading Educational Research: How*

to Avoid Getting Statistically Snookered by Gerald W. Bracey,[4] which I have required of my students. *See* Attached Exhibit A at 2. In particular, without a thorough understanding of the various mechanisms by which race affects educational and economic indicators, students can too easily arrive at glib explanations for racial gaps that echo long-discredited theories of eugenics. As some of these students will go on to work in policy research, this will lead to ineffective policies for addressing serious concerns.

20. Additionally, a handout I created in August 2020 and have used since in my Basic Descriptive and Inferential Statistics Applications course violates H.B. 7's prohibited concepts. *See* Attached Exhibit B. This handout discusses the variable of race in the field of statistics. It ties race to colonialism and eugenics, offers a brief overview of the history of discrimination in the United States, and examines the appropriateness of analyzing race in statistics. I caution my students to not make causal inferences about racial disparities, to be selective in measuring race, and to consider the effects of past and ongoing discrimination. I also speak about my white privilege. I am concerned that I cannot use this handout because I speak of my "white privilege" and of

[4] Gerald W. Bracey, *Reading Educational Research: How to Avoid Getting Statistically Snookered* (1st ed. 2006).

systemic racism—in specific terms, the way past and ongoing racial discrimination are still causing achievement gaps today.

21. Because of the history of tests being used for racial discrimination, designers of high-stakes assessments must attend to strict standards of fairness, many of which are documented in the AERA/APA/NCME joint *Standards*. From my years at ETS, I know how seriously these standards are taken in the testing industry. I also know this from talking with my colleagues at Pearson, American Institutes for Research, the Institute for Educational Sciences (U.S. Department of Education), and the Florida Department of Education. All of these organizations have procedures in place to ensure that assessment instruments are unbiased measures of performance. A full discussion of this requires understanding how cultural differences among racial groups can have a racially differential impact, producing a biased instrument.

22. In particular, social scientists need to distinguish between *impact*—noting a difference between racial groups on some measure—and *bias*—a measure that gives different scores to members of different groups with the same underlying ability (in other words, unfairly advantaging one group over another). One important source of bias is *construct-irrelevant variance*—features of a task which contribute to its difficulty but are unrelated to the construct being measured. One example I have used to illustrate this is an old

literacy test from the state of Louisiana. *See* Attached Exhibit C. H.B. 7 would prohibit me from using this memorable teaching aid because of the discussion on discrimination and test design fairness that necessarily accompanies this example.

23. Another example of content H.B. 7 directly impacts is in the chapter on Fairness in Testing (Chapter 3) of the AERA/APA/NCME joint *Standards*. This teaching material discusses how potential inequities in school resources available to students from traditionally disadvantaged groups—including racial minorities—affect the quality of the education they receive, which naturally affects their test scores and the validity of conclusions stemming from analyses of those test scores.[5]

24. H.B. 7's text limits my ability to train my students on this aspect of the AERA/APA/NCME joint *Standards*. This is deeply problematic for a few reasons. My program's handbook requires graduates to understand and uphold the AERA/APA/NCME joint *Standards*, including fairness in testing. *See* Attached Exhibit D at 4-5. And as a member of two of those organizations, I am bound by the ethical principles embodied in those standards. Plus, almost all prospective employers of our students expect that our graduates will be

[5] American Educational Research Association et al., *supra* note 3 at 66-67.

familiar with the standards and will follow the principles laid out in them. In fact, the peer review process for state assessments instituted by the U.S. Department of Education evaluates the assessments based on those standards.

25. Absent the ability to fully and openly discuss the context behind why these inequities in school resources exist, I cannot convey the significance of how inequities in school resources affect fairness in testing. Nor can I help the students distinguish between a real disparity that is the result of past discrimination (impact) and an apparent disparity because the measure is not agnostic to the race (or other protected group status) of the examinee. This could result in poorly designed measurement systems that would be less effective at educating students and lead to policy choices that do not address the correct problem.

26. H.B. 7's sixth prohibited concept also restricts my ability to explain to my students why the AERA/APA/NCME joint *Standards* call for review of assessments by a diverse body of reviewers. Diversity on a large number of factors, including race, gender and religion, is important because members of one group may spot issues that are not readily apparent to members of other groups. I understand H.B. 7's sixth prohibited concept to forbid taking these steps, which are needed to achieve the kind of diversity that is important to

being able to create valid measures of psychological and educational constructs.

27. Basically, H.B. 7's restrictions put me in the position of being forced to choose between my professional obligation to fully teach the joint *Standards* to my students in order to prepare them for their careers and expose them to the full range of psychometric theories accepted in the field and my obligation as a citizen to comply with Florida law.

28. Legal scholar David Schum, in his 1994 book *The Evidential Foundations of Probabilistic Reasoning,* argues that the process of how truth is determined in science is similar to how truth is determined in a court of law. All sides are allowed to present their evidence, and the participant must then weigh that evidence and come to a conclusion. I try to take the same spirit in my classroom. The students are encouraged to bring different perspectives into the classroom and learn through discussion. The instructions on my classroom discussion forums read in part: "Some of these posts may touch on political issues. Please focus discussion on the statistical issues raised and discuss the political ones in other forums. Also remember that just because we disagree, that does not mean we need to be disagreeable." H.B. 7 violates this spirit of open scientific inquiry. It privileges one perspective on understanding the

effect of race on education, economics, and assessment instead of encouraging the open exploration of multiple perspectives.

29. A Florida statute should not require me to choose between following the law and complying with my ethical obligations. Yet, that is exactly what H.B. 7 does. In considering what content to include in my courses moving forward, I must now weigh the consequences of violating H.B. 7 against my duty to instruct my students in accordance with the standards of my profession. Basically, I have to choose between adhering to the state of Florida's favored viewpoint, which will result in my failing to provide my students the full education they need to succeed in their chosen field, or risk punishment for teaching about the significance of race as a variable and covariate in social science models.

C. How H.B. 7 Will Impact My Teaching and My Courses

30. I feel that the goal of H.B. 7 is to pressure me to self-censor when discussing topics that relate to or implicate a prohibited concept, particularly my viewpoint that we must be conscious of race when seeking to design unbiased measurement systems. I am afraid that I will be punished because a student complains about content in a course of mine. According to Florida Board of Governors proposed regulation 10.005, Florida State could discipline me, including terminating my professorship, in response to a student complaint.

31. I am concerned that H.B. 7 prevents me from teaching content that is critical to my students' success in their statistical fields, particularly educational measurements, and to my own professional obligations as a professor of Measurement and Statistics.

32. Because I created my own handout on race in statistics, I believe H.B. 7 does not allow me to continue to use this handout in my Basic Descriptive and Inferential Statistics Applications course. I am currently having to decide whether to continue including this handout in my future courses by weighing the importance of the concepts discussed against the potential penalties for violating H.B. 7.

33. I feel H.B. 7 forces me to choose between my professional obligation to educate my students based on the industry standards for the fields they will be entering and my duty to follow Florida law.

34. I am also worried that if I violate H.B. 7, whether intentionally or unintentionally, Florida State University could lose significant amounts of funding or have a lawsuit filed against the university for "discrimination" based on something I said in one of my courses.

35. If I omit content that touches upon or implicates H.B. 7's prohibited concepts, I am apprehensive about how my graduate-level students will perform in their

fields after graduation. I am particularly fearful that they could make errors in analysis that harm traditionally disadvantaged groups.

D. Exhibits

36. Attached as **Ex. A** is a true and correct copy of my Basic Descriptive and Inferential Statistics Applications (EDF 5400) syllabus.

37. Attached as **Ex. B** is a true and correct copy of a handout I use in my Basic Descriptive and Inferential Statistics Applications (EDF 5400) course titled "Some Thoughts about Race." I personally created this document in August 2020.

38. Attached as **Ex. C** is a true and correct copy of an excerpt of a PowerPoint Presentation that I have given since at least 2015 in my Scale and Instrument Design (EDF 5448) course. This part of the presentation covers construct-irrelevant variance and fairness in testing.

39. Attached as **Ex. D** is a true and correct excerpted copy of the Department of Educational Psychology and Learning Systems's Measurement and Statistics Program Handbook, as of December 2018.

Pursuant to 28 U.S.C. § 1746, I declare under penalty of perjury that the foregoing is true and accurate.

Executed on August 24, 2022.

Dr. Russell G. Almond

Ex. A

FLORIDA STATE UNIVERSITY

EDUCATIONAL PSYCHOLOGY AND LEARNING SYSTEMS, COLLEGE OF EDUCATION

Instruction that moves, leadership that inspires,
scholarship that makes a difference for the 21st century.

A. Course Prefix, Number, Title and Credits

EDF 5400.Basic Descriptive and Inferential Statistics Applications (4)

B. Prerequisites or Co-requisites

None.

C. Objectives/Description

This course prepares students to both read and write papers containing basic statistical analyses. Topics covered include descriptive statistics, basic plots and graphing, hypothesis testing, confidence intervals, correlational techniques, and introduction to the general linear model.

After completing this course, the student will be able to:

1. Differentiate between samples and populations and between parameters and sample statistics.
2. Distinguish representative and random sampling from other kinds of sampling and understand the statistical implications.
3. Classify variables as belonging to one of the four different types of measurement scales (Nominal, Ordinal, Interval, and Ratio) and understand the limitations of each type of scale.
4. Construct and interpret common kinds of graphs used in statistical analysis (including histograms, frequency polygons, scatterplots, boxplots, and bar charts).
5. Recognize terms used to describe distributional shapes (e.g., normal, bimodal, skewness, kurtosis) and properly identify distributions for which the term is appropriate.
6. Interpret expressions written in summation notation.
7. Calculate and interpret measures of location (mean, median, and mode) and understand their limitations.
8. Calculate and interpret measures of scale (range, variance, standard deviation, and interquartile range) and understand their limitations.
9. Calculate and interpret z-scores for data which come from a known distribution.
10. Convert between quantiles of a normal distribution (z-scores) and tail probabilities (p-values) using a table of the normal distribution.
11. Interpret Pearson's correlation coefficient and understand its limitations.

12. Apply the basic definition of probability and conditional probability to calculate the probability of events.
13. Recognize when two variables are probabilistically independent.
14. Calculate the standard error of the mean of a simple random sample and interpret the standard error of other statistics.
15. Recognize sources of bias especially in non-random samples.
16. Calculate and interpret confidence intervals for the sample mean, and interpret confidence intervals for other parameters.
17. Perform statistical inference using the null hypothesis framework in simple problems, including:
    a. Writing null hypothesis for common inference situations.
    b. Write alternative hypothesis for one-tailed and two-tailed tests.
    c. Calculate p-values from z-scores and critical values for a given level.
    d. Properly interpret Type I and Type II errors.
    e. Calculate and interpret effect sizes for the difference between two means, and interpreting effect sizes for regressions.
    f. Calculate (using software) and interpret the power for a given experiment.
18. Run and interpret the output of common statistical software for each of the following tests:
    a. One-sample and paired-sample *t*-tests
    b. Simple linear regression with one variable, including regression ANOVA and t-test for the slopes.
    c. Chi-squared tests of goodness of fit and independence
19. Recognize and distinguish situations in which the tests listed in 18 are appropriate.
20. Recognize the simple linear regression model as a basic example of statistical models and perform model checking.
21. Recognize the limitations of the statistical procedures in 18.
22. Write up the results for the statistical techniques covered in 18 following APA conventions.

D. Required Texts, Readings, and/or other Resources

Required Text:

Coladarci, T. & Cobb, C. D. (2014). *Fundamentals of statistical reasoning in education (4th ed.).* Hoboken, NJ: John Wiley & Sons.

Bracey, G. W. (2006). *Reading educational research: How to avoid getting statistically snookered.* Portsmouth, NH: Heinemann.

Recommended Text:

Glass, G., & Hopkins, K. (1996 or any other year). *Statistical methods in education and psychology.* Boston, MA: Allyn and Bacon.

Software for Data Analysis:

The labs and some of the homework for this class will involve analyzing large data sets, made easier by the use of statistical software on a computer. To this purpose, the class will include demonstration of how to do common data-analysis operations with SPSS. Access to the full version of SPSS for Windows is available for FSU students at no cost in the COE Learning Resource Center (LRC) in 1301 Stone building (M-Th 8AM-10PM, F 8AM-6PM), and other locations on campus (facilities and schedules available at http://us.fsu.edu/index_labs.html). The graduate computers in Strozier also have SPSS installed. SPSS is also available through the FSU Virtual computer lab.

The use of SPSS is not required. If students are already familiar with and have access to another statistics package, they are free to use that instead. However, course instructors and graders will only be able to provide minimal support for other packages.

E. Topical Course Outline

Lesson	Topics	Reading
1	Common Sense Statistics, Making a Dataset	Ch 1†+Bracey
2	Variables	Ch 2,3
3	Distributional shapes, Graphs	Ch 3
4	Measuring Location & Center, and Measuring Variability & Scale	Ch 4, 5
5	Normal Distribution and *z*-scores	Ch 6
LRC1	**Exploratory Data Analysis***	Ch 7
6	Correlation	Ch 7, 8
7	Correlation and Regression	Ch 7, 8
8	Regression and Prediction	Ch 8
LRC2	**LRC 2: Regression***	Ch 9
9	Probability	Ch 9
10	Conditional Probability	Ch 10
11	Central Limit Theorem	
12	Inverse Probability (Confidence Intervals)	Ch 12
13	Hypothesis Testing (*z* tests)	Ch 11
14	Type I and Type II Errors	Ch 13
15	Student's *t* test for One Mean, CI for the Mean	Ch 13
16	Student's *t* test for Two Independent Samples	Ch 12-14
LRC3	**LRC 3: Student's *t* test***	Ch 14, 19
17	Student's *t* test for Two Dependent Samples	Ch 15

18	Testing for Correlation and Regression	Ch 17
19	Categorical Data	Ch 18
LRC4	**LRC 4: Chi-square tests***	

†Unless otherwise noted, Chapter numbers refer to Coladarci et al. (i.e., first week's reading is the first chapter of Coladarci et al, plus all of Bracy.

* Classes called "LRC X" noted in bold face will be held in the LRC's computer lab.

F. Teaching Strategies

The course is offered in a "flipped" format with lectures pre-recorded. Class sessions will contain only short reviews meant to provide students opportunities to seek assistance and ask questions as well as in-class activities to reinforce learning. Online activities include viewing lectures and online quizzes and tests. In-class activities include worksheets, problem sets, case studies and discussions.

G. Field/Clinical Activities

N/A

H. Expectations/Attendance

University Attendance Policy

Excused absences include documented illness, deaths in the immediate family and other documented crises, call to active military duty or jury duty, religious holy days, and official University activities. These absences will be accommodated in a way that does not arbitrarily penalize students who have a valid excuse. Consideration will also be given to students whose dependent children experience serious illness.

Netiquette Statement

As this is a blended class, some activities will take place online. Considering online classes will take place in a variety of settings, it is important to have a reference point for successful participation in this online environment.

Be mindful of the Core Rules of Netiquette taken from Virginia Shea's Book and Website - "http://www.albion.com/netiquette/corerules.html"

Rule 1: Remember the Human.
Rule 2: Adhere to the same standards of behavior online that you follow in real life.
Rule 3: Know where you are in cyberspace.
Rule 4: Respect other people's time and bandwidth.
Rule 5: Make yourself look good online.
Rule 6: Share expert knowledge.
Rule 7: Help keep flame wars under control.

Rule 8: Respect other people's privacy.

Rule 9: Don't abuse your power.

Rule 10: Be forgiving of other people's mistakes.

Sexual Harassment Policy

Sexual harassment is a form of discrimination based on a person's gender. Sexual harassment is contrary to the University's values and moral standards, which recognize the dignity and worth of each person, as well as a violation of federal and state laws and University rules and policies. Sexual harassment cannot and will not be tolerated by the Florida State University, whether by faculty, students, or staff; or by others while on property owned by or under the control of the University.

Course Expectations

Class attendance policy

In addition to the valid excuses in the University Policy, the class has the following expectations for class attendance:

- Aside from the first day of class, attendance will not be taken in class.
- *However, in-class participation points (case studies and practice problems) are assigned only to students who are present on that day.* Students with a valid excuse will be given an alternative assignment.
- Changes to assignment content and/or due dates are routinely announced at the beginning of class, students who miss or who are late for class are responsible for catching up with those announcements.
- Students are responsible for material covered in missed classes.

Communication Expectations

Clear and regular communication is essential to success in this class. Students need to check their email regularly, be active in class, and be proactive about asking questions. The instructor will respond to emails and posts to the Q&A board within two business days, usually sooner.

Policy for Late Work

The instructor will track the number of assignments submitted after the posted due date. Students are allowed one free late assignment. After that, the grade for that assignment will be lowered by 5% for each previously late assignment, unless the student has instructor approval. Work submitted for problem sets after the answer key has been posted will be assigned a grade of zero unless the student has received instructor approval.

I. Grading/Evaluation

Final grades will be based on the following combination of evidence:

1.	Participation	25%	(556 Participation Points)
2.	Problem Sets	25%	(500 Problem set points)
3.	Labs	25%	(30 Lab points)
4.	Exams (2)	25%	(200 Exam points).

We will use the following means of assessment:

1. *In class and online participation (556 Participation Points; 25% of grade).* The primary purpose of these activities is for formative assessment: if instructors do not know how well the students understand the material, the instructors cannot adapt teaching to meet the students' needs. Participation grades will be based on participation – *not on whether or not the student provides the right answer* – but on the students' honest expression of what they do and do not know. Although instructors will not take attendance, participation points are assigned in class, so students don't show up (excluding excused absences) will have lower participation scores. Participation grades will be based on the following kinds of activities:

   A. *Self-checks (391 participation points; 17.4% of total grade)* – Short, self-graded quizzes are delivered by the learning management system (Canvas) after each lecture segment as well as some other times during the course. Students may repeat these quizzes any number of times, with the highest score contributing to the grade. These are meant as formative assessments for students to assess their comprehension of the class material, and questions about them should be raised in class or on the discussion boards. These are assigned due dates to help students with pacing of the material, but students may revise their answers without penalty up until the last day of class (Friday before Final's week).
   B. *Practice Problems (7 problems, 10 participation points each; all 7 are 3% of total grade)* – This is a worksheet of questions which the students complete normally in class (although some may be moved online for scheduling reasons). Students may discuss the answers while they work. After completion, the students will be supplied the answer (often as a class discussion) and will correct their own work. Students then complete an online self-assessment questionnaire about the topic of the worksheet. Points are given for completing the self-assessment with all answers worth maximum points.
   C. *Case Studies (10 case studies, 10 participation points each; all 10 are 4.4% of total grade)* – A case study consists of a set of materials (usually the description of a research study and SPSS analysis of data from that study, but occasionally a short paper) and a set of questions. Students work in groups to answer the questions. These may be administered face-to-face or through online discussion forums. If through the forums, then points will be awarded based on contributions to the discussion, if face-to-face, then though an online self-assessment questionnaire.
   D. *Miscellaneous Participation (5 participation points; 0.2% of total grade)* – One participation point is awarded for making a self-introduction post in the introductory forum. One participation point is awarded for turning in drafts of Labs 1, 2 and 3 (see Lab Assignments below). One additional point is awarded for being the first author (the one who distributes the feedback to other students) on at least one of the labs.

2. *Problem Sets (5 problem sets, together they are 25% of total grade)* – A problem set is a group of problems related to a specific topic. Team work is allowed with up to 3 students per team. Students will need to turn the assignment in through the course website for credit (assignments emailed to the instructor may or may not be accepted). The answer key for problem sets will be posted on the course web site after grading. Late submissions will not be accepted after the answer key is posted without permission of the instructor.

3. *Laboratory exercises (3 labs, together they are 25% of total grade)* – The labs consist of a statement of a problem and a set of data. Students produce a short report (essentially the results section of a scientific paper) using the data to address the problem. Students turn in a draft report which is returned with feedback but no grade (one participation point is awarded for turning in the draft). The students may then revise the report and turn it in for a grade; usually one week is available between when the feedback is released and the final report is due. Group work is allowed on labs with up to 3 students per group. *Note: The labs are a critical part of 5400 and are weighted equally with the exams exam.* Complete lab instructions along with information on the grading rubrics will be posted on the course website.

4. *Exams (Midterm and Final, together they are 25% of total grade)* – Both the midterm and final will be given online (through the learning management system). Exams will be open book and students will be allowed to use online resources (including class notes, and online probability calculators). Students must not consult with others (aside from the course instructors) during the exam. Exams with students having 3 hours to complete the exam.

GRADE SCALE

In computing the final grade each of the four categories will be scaled separately. That is the 556 participation points together will consist of 25% of the grade, the 500 homework points will be 25% of the grade, the 200 exam points will be 25% of the grade, and the 30 lab points will be 25% of the grade.

A = 93-100%
A- = 90-92.9%
B+ = 87-89.9%
B = 83-86.9%
B- = 80-82.9%
C+ = 77-79.9%
C = 73-76.9%
C- = 70-72.9%
D+ = 67-69.9%
D = 63-66.9%
D- = 60-62.9%
F = below 60%

Collaboration Policies:

- Students are also encouraged to use the discussion boards to discuss other course related content. Instructors and graders will monitor the discussion and will try to answer questions (chances are good that if one student has a question, at least one other student has a similar question). This is also a good way for students to provide feedback to instructors about which material to review in class.

- Students may work together in groups of 2 or 3 on homework and lab reports. However, *all* team members must participate for this to work well. Students help their friends much more by offering them a hint towards the solution than by simply giving them the answer. Using the class discussion boards to discuss the homework is encouraged.

- Students working in teams will sign a statement saying that all partners collaborated fairly on each assignment (students who feel that team members are not properly contributing should contact the instructor). Teams of students should turn in single copies of the group homework assignment or lab report with all names listed.

- The first author of each lab report will turn in the assignment on the course web site and distribute feedback to the rest of the team. Students working in teams on the labs will be expected to take the first author role for at least one lab assignment.

- Lab assignments must be submitted through Turnitin within the course website. This will flag possible plagiarism – copying from other students or other sources. Plagiarism is theft, and it is a serious problem!! Do not do it at risk of failing the assignment! Links to online resources about plagiarism will be posted on the course web site.

Free Tutoring from FSU

On-campus tutoring and writing assistance is available for many courses at Florida State University. For more information, visit the Academic Center for Excellence (ACE) Tutoring Services' comprehensive list of on-campus tutoring options - see http://ace.fsu.edu/tutoring or contact tutor@fsu.edu. High-quality tutoring for fundamental concepts in math, statistics, science and additional subject area tutoring is available by appointment and on a walk-in basis. These services are offered by tutors trained to encourage the highest level of individual academic success while upholding personal academic integrity.

J. Honor Code

The Florida State University Academic Honor Policy outlines the University's expectations for the integrity of students' academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process. Students are responsible for reading the Academic Honor Policy and for living up to their pledge to ". . . be honest and truthful and . . . [to] strive for personal and institutional integrity at Florida State University." Florida State University Academic Honor Policy, found at http://fda.fsu.edu/Academics/Academic-Honor-Policy.

Please be aware that using social media to collaborate on and share course exams or assignments with other students that are not identified by the course instructor as group work is a violation of the FSU Academic Honor Policy.

Any violation of the collaboration policies, including exchanging answers electronically, plagiarism or cheating on tests will be considered violations of the Academic Honor Code of FSU.

F. Americans with Disabilities Act

Americans With Disabilities Act:
Students with disabilities needing academic accommodation should:
(1) register with and provide documentation to the Student Disability Resource Center; and
(2) bring a letter to the instructor indicating the need for accommodation and what type.

Please note that instructors are not allowed to provide classroom accommodations to a student until appropriate verification from the Student Disability Resource Center has been provided.

This syllabus and other class materials are available in alternative format upon request.

For more information about services available to FSU students with disabilities, contact the:

Student Disability Resource Center
874 Traditions Way
108 Student Services Building
Florida State University
Tallahassee, FL 32306-4167
(850) 644-9566 (voice)
(850) 644-8504 (TDD)
sdrc@admin.fsu.edu
http://www.disabilitycenter.fsu.edu

K. Syllabus Change Policy

Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice

Ex. B

Some Thoughts about Race

Race is a commonly used categorical variable in statistics. Almost every application or questionnaire seems to have a box which asks about *race* or *ethnicity*. National statistics are commonly disaggregated by race, and there are numerous papers comparing "whites" to "others".

Two deaths in 2020 have gotten me thinking more about the construct of *race* and how we use it in statistics. One is the unfortunate death of George Floyd at the hands of law enforcement. The other is the death of Congressman John Lewis, a hero of the civil rights movement. In fact, the graphic novel *March* (Lewis, Aydin & Powell, 2016) has been part of my summer reading.

Race is an interesting construct in that it has a long an complex history, much of it fraught with problems. In particular, it has its roots in colonialism and slavery and its misapplication has resulted in genocide. The worlds of statistics and psychometrics have not be innocent bystanders here. Many prominent statisticians were active in the eugenics movement.

This screed, is my attempt to struggle with what this construct is, what it has been, and how we should tackle it in the 21st century. I am well aware that as an old white dude that I only have one perspective on this, and a privileged one at that. I'm hoping that my musings are helpful in your own thinking about this construct.

Race is not really a genetic or cultural construct

Although skin color, which is a genetic trait, is commonly used to classify a person into a racical group, racial groups are not really natural genetic sub-populations. For example, "Asian" is a racial group, but Chinese, Koreans, Japanese, and Vietnamese people have distinct genetic pools and cultures. (In fact, Chinese is probably too big a category, too, as there are a large number of different groups within China, e.g., Han, Wu, Uighur, &c). Indians are also called "Asian", but India is quite different from China (even though they share a border). Hutu and Tutsi people are both African, but that did not stop them from fighting a civil war in Rwanda.

Racial groups are also not unified culturally. Latinx refers to a large number of groups who speak Spanish. However, Mexican, Puerto Rican, Central American, South American, Filipino, and European Spanish are all very different. Similar the group "Native Americans" includes a large number of different peoples who have all been lumped together. Even though Hopis and Apaches both live in the American southwest, they have different traditions and religions.

Race arises from colonialism

In 1599, Lope de Vega wrote a play *El neuvo mundo descubierto por Cristóbal Colón* (The New World Discovered by Christopher Columbus). In this play, Columbus is put on trial by Fate, with the Devil as the prosecution and the Christianity as the defense. At issue is whether or not Columbus is justified in colonizing the new world. The verdict that de Vega, a devout Catholic, reaches is that Columbus is justified as he will bring Christianity to the savages. Writing in the 16th century, de Vega is not that far removed from the Spanish Inquisition, which justified torture in this world to save a person's immortal soul.

The association of race and skin color probably dates back to the enslavement of African by early American settlers. As skin color was a convenient marker to separate the enslaved from the free, it became codified. I'm sure that much of this was rationalizing the cruelty of slavery. If people believed in the superiority of the white race the could justify the cruelty of slavery and genocide.

Similarly, the Native Americans were put into a "non-white" savages category. This was used to justify stealing their lands and programs of genocide. They weren't using the land optimally, so we will take it from the "savages" and give it to the "civilized" settlers. Once again, the belief in the superiority of the white race was used to rationalize genocide.

Once the concept of race existed, it hung on socially, even after the end of slavery. In particular, there are a lot of whites who considered race a part of their identity. In many ways, race is defined by whites who don't want to share their sense of superiority with others. Witness the recent question I was asked about race in on a loan application, where there is only one box for "white" but lots of boxes for various categories of "other."

The eugenics movement needs special mention because it is tied so intimately with the history of statistics and psychometrics. Names like Galton and Fisher are responsible for much of the theoretical foundations we use today, but they were also involved with the eugenics movement. When Simon Binet's education test was first brought to the U.S., the purpose was not to as Binet wanted, to seek out people in need ot help, but rather to separate the strong from the "feebleminded." (Frank, 2018, Intelligence Testing and the Beginning of Eugenics, *Owlcation* https://owlcation.com/social-sciences/Intelligence-Testing-and-the-Beginning-of-Eugenics)

Although early researcher showed racial difference on "intelligence" tests, intelligence is an ill-defined concept. In particular, much of what we think about as intelligence is cultural acclimatization. Consider a person from a remote region of the Amazon. If that person was taken to a big city and walked through a red light, the label "stupid" might come to mind. However, if I was to go to deep into the Amazon and step on a poisonous snake, the Amazon native would think I was stupid. Intelligence (which has more to do with the capacity to learn quickly) is different than specific situational knowledge.

Fortunately, we've come a long way in our understanding of measurement of both intelligence and achievement. In particular, psychometricians pay a lot of attention to *differential item functioning*—whether people from different racial groups with similar abilities respond differently to test items. We also make sure that tests are reviewed by people from diverse ethnic and racial backgrounds so that language that is either specific to one subgroup or has negative associations for some subgroup is not inadvertantly used. Once psychometricians learned to take some of the bias out of the instruments, many of the observed racial differences went away. Others remain, but quite possibly as a function of discrimination.

Any two groups are different, but is that different important?

Another issue is that if you go to enough decimal points (and implicitly have a big enough sample size) any two groups will be different. Lets use left-handed and right-handed people as a pair of groups that do not have a long history of discrimination. [Okay, I hear you lefties about how mice, scissors and all kinds of other tools are designed by and for righties.] If we looked at all of the left-handed and right-handed people in the United States, then one of the two groups would have a higher average height. But that difference is likely to be a small fraction of an inch: it just wouldn't be important.

On the other hand, the standard deviation for the population of heights is around 3 inches. So the *effect size*, the size of the difference divided by the population standard deviation, is likely to be small. So if you were picking people for your basketball team, you would likely not favor left-handers or right handers. (Actually, the fact that lefties and righties shoot from different angles is probably more important than the height difference, and you really want a mix of lefties and righties to give your team balance.)

This is one of the weaknesses of statistical hypothesis tests, something we will address later in the class. A *significant* difference is merely one for which the sample was big enough to be sure that the different between the two groups is probably not just sampling error. If the sample is huge (thousands or millions) even very tiny differences can be significant. The very next question you should ask every time you learn that a difference is significant is "Is the difference really big enough to be important?"

The second problem is that finding a difference between two groups, especially two naturally occurring groups, is an observation, not a causal analysis. Suppose that we learn that the focal group (which is usually a group that has been historically discriminated against) does not do as well on a certain kind of math item as the reference group (which is the higher status group, usually white males). Does that mean there is some genetic difference between the groups, or is there some environmental or social factor that causes the difference.

There is a history of Discrimination

"Your zip code is a bigger predictor of health and economic outcomes than your genetic code." (Guest David R Williams on the *Ezra Klein Show*, https://podcasts.apple.com/us/podcast/why-the-coronavirus-is-so-deadly-for-black-america/id1081584611?i=1000474196638, approx 59 min into podcast. [Much of this podcast talks about racial differences before the Coronavirus outbreak, and some about why those same forces result in different outcomes for the virus.])

Before the civil war, Black Americans were mostly enslaved. Although they were given some property after the emancipation and the civil war, this was mostly taken away during the Jim Crow era. There were actually legal barriers to where people of various races could live. Although the Civil Rights movement in the 1960s eliminated some of the legal barriers, cultural ones remained. The past pattern of discrimination has impacted the wealth formation.

Let me give you an example. In 1968, my parents bought a home for around $40,000. They still live in this home, it is worth around $400,000 now. They have paid off their mortgage, so this represents a substantial hunk of savings for my family. Now 1968 is an interesting year. This was the year in which red-lining, a practice which would exclude people of certain races (i.e.,Blacks) from certain neighborhoods became illegal. Even though the law was changed, real estate agents still continued to show Black families home in Black neighborhoods. Not only that, but very few Black families in 1968 (or even today) would have the money needed for the down payment; and they would be considered a credit risk resulting in higher interest payments. Nor did many Black students before 1968 have had the opportunity to go to college and earn a PhD, giving the steady stream of income needed to get an inexpensive home. Although neither my parents, nor my parents are racist, we have certainly benefited (or at least not been disadvantaged) by racist policies.

In education, segregation has had a huge impact. In the United States, education is largely paid for through local taxes, particularly property taxes. This means that the rich get access to better education than the poor. Given that there were active policies (and to this day covert practices) to keep non-whites in poor neighborhoods, that means that they do not get the same schooling resources. Hence, for most of the 20th and the first part of the 21st century, our education system has been separate and unequal.

In Florida, school districts correspond to counties. So the biggest problems are poor rural counties. In New Jersey, the school districts correspond to townships. So Princeton Township, Lawrence Township, and the City of Trenton are all in Mercer County, but have different levels of resources, and hence different qualities of schools. Unsurprisingly, Trenton has the least resources and the largest Black and Latinx populations. Note that the problem occurs more subtly even within Leon County. My wife volunteered at Deer Lake Middle school doing office work

(e.g., making copies); however, Deer Lake also served a population that was wealthy enough to have families that could get by on one income, leaving one parent to provide volunteer labor.

Some of the problems could be solved if we were willing to send money from the wealthier school districts to the poorer ones. New Jersey actually tried that experiment. In a lawsuit called *Abbott versus the Board of Education*, parents from several of the poorest districts (mainly urban centers like Camden, Trenton, New Brunswick, Patterson and Newark) sued the state over the constitutional requirement to provide a quality education. As a result, the "Abbott districts" received extra funds from the State. As a result the racial gap closed somewhat. In general, however, such transfers have not proved politically popular.

Ongoing segregation and wealth gaps which are at least partially the result of past discrimination are not the only factor that could be causing achievement gaps. Williams (in the same *Ezra Klein Show* podcast) notes that ongoing discrimination (even if only occasional) causes increases in stress which are particularly bad for health outcomes, but also probably affects other things like education.

The problem mostly arises when we look for simple causal stories. What is happening today with race is actually quite complex. There are a lot of different pathways by which what was originally a social construct to support colonialism can continue to affect outcomes. I have heard this called *causal density*—a situation in which there are so many potential causes and causal pathways that it is difficult to trace the effects of any single cause.

When to include the Race box

Over the years, I've noticed a certain laziness towards the inclusion of race in scientific studies. In particular, I think people put boxes for *race*, as well as *gender* and *age* on questionnaires, particularly background questionnaires, without thinking. How does *race* relate to the research questions being addressed?

After all, people have a right to privacy. I didn't ask about race in the introduction questionnaire for this class, because it is none of my business. (My business is teaching statistics, and my experience has led me to believe that race has no relationship to the ability to learn statistics.) Constantly asking about race is constantly forcing people to identify with smaller groups and remind them of their exclusion. Why should I do that instead of remind people of what we share, like the fact that we are all associated with FSU?

Its okay to ask about race if race is of legitimate scientific interest. Race could could be interesting for either direct, or indirect reasons. The direct reason is that we think that there might be an interaction with the between the treatment we are exploring and the complex social construct that is race. For example, do people from different racial groups have different reactions from educational

games? Or maybe race is directly what we are studying. That is to say, we are trying to untangle the Gordian knot that is race and discrimination and how it affects education, public health, economic welfare, voting patterns, or something else.

The second case where it is okay to ask the race question is if you are concerned about the representativeness of the sample. One way to look at that is to ask, is the racial composition of my sample something can can reasonably arise from random sampling, or do I need to do something to make sure I have enough of particular minority groups in my sample.

One example of this is looking at the racial breakdown of trial juries. Looking at the proportions, it is clear that whites are overrepresented on juries and other groups, particularly, Blacks and Latinx, are overrepresented. Trying to understand why this is happening is difficulty. One explanation is that lawyers are more likely to use challenges to get rid of non-white jurors. Another explanation is that jurors are often selected from voting registration lists, and whites are more likely to be registered to vote.

One more thought: sample size is an issue. In particular, don't do racial comparisons if there are very few (e.g., fewer than 5–10) people from one group in the sample. If you start crossing a bunch of categories together, you can get cells in your data table that only have one or two people in them. There is only one person who fits the category male Latino doctoral student in the Measurement and Statistics program. (Trudes is a great guy, who helped a lot with the design of the 5400 web site, but it would be completely wrong to generalize from him to all Latino males.) If you are looking for information about a minority group with a small or localized population (e.g., Native Americans or Somali refugees) you may need to do something more complex than simple random sampling.

Closing thoughts

One of the advantages of white privilege is that I got away with living most of my life without the need to think deeply about what race is and how and why we should measure it. Actually, I've put some thought about race here and there over time, but my model of what this construct really is has gotten more complex over time.

Two pieces of advise, which apply to other socially loaded constructs (gender, disability, religion, &c), as well as to race. First, don't just ask the question without thinking about why you are asking the question. What are you going to do with the data? What would you do if you found an imbalance? Asking questions on questionnaires without thinking about why you are asking them is a recipe for sloppy science.

Second, don't leap to causal conclusions from observational studies. There are

people who are putting in the hard work to untangle all of the causal pathways that lead from race to performance. The wrong way to do this is to look at a single observed difference and decide it is just a function of genetics, or just of discrimination. Both of those are lazy answers. Going from observed differences to causality always takes careful work and usually synthesizing many sources of evidence.

Finally, as an old white dude, I should definitely not conclude that my ramblings are the final word on race. After all, one aspect of white privilege is that I don't experience racial discrimination routinely. I have learned a lot from listening to other people's opinions, particularly those of Black intellectuals and Black and African students. In that spirit, I would welcome your comments, feedback and stories. Reading about your thoughts and experiences will help me shape my own beliefs, and I hope that reading my ramblings will help you shape your own.

Ex. C

Cognitive and Non-cognitive Items

"Non-cognitive" constructs and playtesting

Construct-Irrelevant Variance

- We want to maximize variation in response for reasons related to the construct being measured (signal)
- We want to minimize variation in response for reasons *unrelated* to the construct being measured (noise)
- Good item writing identifies sources of construct-irrelevant variance and minimizes them (or averages over on the test form)

Negatively worded items

- Avoid negative wording in cognitive items
  - Missing *not* is construct irrelevant mistake
- Negative wordings are used in non-cognitive assessments
  - Agreeability
  - Laziness

Items arousing strong emotion

- Reading passage about a dying pet
  - Could be very distressing to a person who has suffered a recent loss
- Sometimes impossible to avoid in the context of a history test
  - Slavery
  - Genocide
  - Standards call of introducing such topics only when necessary
- Cultural Sensitivity

Alternative Explanations

- Symptoms that might have a lot of causes
  - Fatigue is a symptom of depression, but also of a lot of other things
- Non-construct related background knowledge required to answer question
  - Dinosaur knowledge helps with reading comprehension

Fairness

- Actual and appearance are both important
- Face-validity: Important that items appear to be fair
- Impact: Different groups will have different mean scores
- DIF: Given level of construct response is independent of group membership
  - Subgroups might have access to different background knowledge
  - Subgroups might bring different context interpretations to the question

Question Phrasing

- Is this at appropriate reading level?
- Want to aim several grade levels below target (unless we are testing reading)
- 8th grade level for U.S. Population
- ESL

Example

Sherlock examined the handle of the fire alarm. "Are any of your suspects art students?" he asked.

What did Sherlock see on the fire alarm handle that caused him to suspect an art student?

A Truly Bad Test

The State of Louisiana

Literacy Test (This test is to be given to anyone who cannot prove a fifth grade education.)

Do what you are told to do in each statement, nothing more, nothing less. Be careful as one wrong answer denotes failure of the test. You have 10 minutes to complete the test.

1. Draw a line around the number or letter of this sentence.
2. Draw a line under the last word in this line.
3. Cross out the longest word in this line.
4. Draw a line around the shortest word in this line.
5. Circle the first, first letter of the alphabet in this line.
6. In the space below draw three circles, one inside (engulfed by) the other.
7. Above the letter X make a small cross.
8. Draw a line through the letter below that comes earliest in the alphabet.

Z V S B D M K I T P H C

9. Draw a line through the two letters below that come last in the alphabet.

14. Draw a line under the first letter after "h" and draw a line through the second letter after "j".

a b c d e f g h i j k l m n o p q

15. In the space below, write the word "noise" backwards and place a dot over what would be its second letter should it have been written forward.

16. Draw a triangle with a blackened circle that overlaps only its left corner.

17. Look at the line of numbers below, and place on the blank, the number that should come next.

2 4 8 16 ____

18. Look at the line of numbers below, and place on the blank, the number that should come next.

3 6 9 ____ 15

19. Draw in the space below, a square with a triangle in it, and within that same triangle draw a circle with a black dot in it.

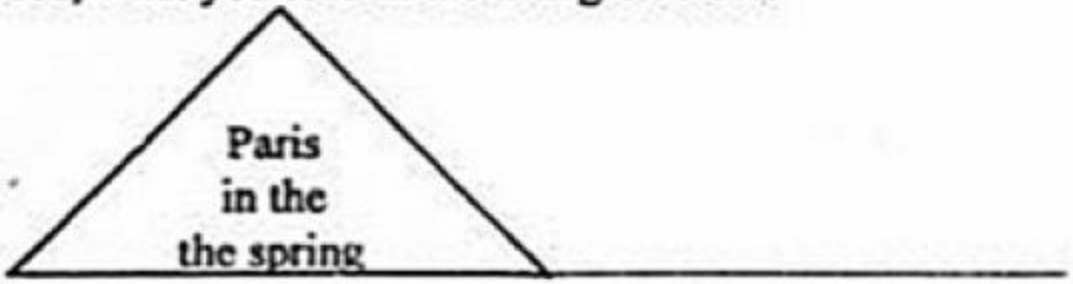
20. Spell backwards, forwards.

21. Print the word vote upside down, but in the correct order.

22. Place a cross over the tenth letter in this line, a line under the first space in this sentence, and circle around the last the in the second line of this sentence.

24. Print a word that looks the same whether it is printed frontwards or backwards.

25. Write down on the line provided, what you read in the triangle below:



26. In the third square below, write the second letter of the fourth word.

27. Write right from the left to the right as you see it spelled here.

28. Divide a vertical line in two equal parts by bisecting it with a curved horizontal line that is only straight at its spot bisection of the vertical

How is this wrong? Let me count the ways

1. Purpose is not to test literacy, but to discriminate
2. Eligibility criterion not fairly applied
3. Questions ambiguously worded
4. Scoring rubric ambiguous
    1. Raters are inconsistent (& biased)
5. Extreme Speededness (30 questions in 10 min)
6. Passing criteria extreme (all questions must be correct)
7. No face validity (what do these questions have to do with being able to understand and mark a ballot)
8. No validity evidence, but strong evidence of unfair impact!

http://www.slate.com/blogs/the_vault/2013/06/28/voting_rights_and_the_supreme_court_the_impossible_literacy_test_louisiana.html (Hat tip to Boingboing.net)

Usability Testing & Review

- Expert Review Panel
  - Diversity is important to assess cultural sensitivity
- Pilot test in front of a few potential subjects
  - Think aloud
  - Videotaping
  - Eyetracking
- Capture sample work products for testing evidence rules

Ex. D

Program Handbook

Measurement and Statistics

Department of Educational Psychology & Learning Systems

Florida State University

Update: December 2018

Table of Contents

Chapter 1: General Information

Welcome to the Measurement and Statistics (M&S) Program at Florida State University. This handbook describes the basics of our program, which offers two degrees: Master of Science (M.S.) and Doctor in Philosophy (Ph.D.). M&S is housed within the program area known as Educational Psychology, in the Department of Educational Psychology & Learning Systems in the College of Education.

The policies identified in this document are to be construed in light of existing University policies and with deference to the requirements imposed on graduate education by the University, the Board of Trustees of Florida State University, and the Governing Board of the State University System of Florida. The information outlined is subject to change and students should be alert to announced revisions required by the faculty of the program, the department, College, and University.

Mission of the M&S Program

The Measurement & Statistics (M&S) program prepares leaders in educational research to serve in various types of professional positions related to collecting, analyzing and interpreting educational statistics. Our mission is to prepare future professors of measurement and statistics at colleges and universities; psychometricians who may work in commercial testing firms; educational-measurement and/or educational-statistics specialists employed by test publishers, regional educational laboratories, and governmental licensing, certification, or assessment units; and directors of measurement activities for schools and school systems.

We also serve the College of Education and Florida State University by teaching education researchers and social-science researchers from other domains to use quantitative approaches appropriately in their areas of application.

Core Values

Graduates from our program are expected to understand and uphold the core professional values of the key professional organizations to which they will belong, including the American Educational Research Association (AERA), National Council on Measurement in Education (NCME), Psychometric Society, the American Statistical Association (AmStat), and the American Psychological Association (APA). In particular we expect our students and graduates to:

- Design test and assessments that are fair and reliable and are valid for the purposes to which they will be put.
- Analyze data in a way appropriate to the purpose of the analysis and respectful of the limitations of how the data were collected.
- Honestly report limitations and bounds on inferences, including standard errors and biases, using estimates drawn from data.

- Store, prepare, analyze and report on data about human participants in a way that respects their privacy and dignity.
- Subject to the above constraints, meeting the needs of all stakeholders in a decision process.

Faculty in the Measurement & Statistics Program

- Russell G. Almond, PhD
- Betsy J. Becker, PhD
- Insu Paek, PhD
- Yanyun Yang, PhD
- Qian Zhang, PhD
- Salih Binici, PhD (adjunct faculty)
- William Yeaton, PhD (adjunct faculty)

Program Staff

Mary Kate McKee
Academic Support Assistant
(850) 644-8792
MMcKee@campus.fsu.edu
3205D Stone Building

Bryan Richards
Administrative Specialist
(850) 645-7976
brichards@admin.fsu.edu
3210A Stone Building

OASIS, FSU Graduate School, and M&S Canvas Site

The COE Office of Academic Services and Intern Support (OASIS) is very helpful. Important guidelines and all of the forms described in this handbook are available on the OASIS website: **https://education.fsu.edu/student-resources/student-academic-services-oasis.** The FSU Graduate School also has many important resources and forms at **http://gradschool.fsu.edu/**. You can find additional information and announcements on the M&S Canvas site, which can be accessed by logging into your account (**https://canvas.fsu.edu**).

It is always important for graduate students to consult with OASIS and the Graduate School for the latest requirements, deadlines, and forms. Things often change and deadlines are often earlier than you think!