S-061A1: Statistical and Psychometric Methods for Educational Measurement
Harvard Graduate School of Education
August 31 – October 17, 2016

Class meets Mondays and Wednesdays, 10:10-12:00, in Longfellow 229
Course website: https://canvas.harvard.edu/courses/18969/
Instructor: Andrew Ho
455 Gutman Library
Andrew_Ho@gse.harvard.edu
TF: Maria (Masha) Bertling
mbertling@g.harvard.edu

Instructor Office Hours: After class or email Wendy_Angus@gse.harvard.edu for appointments.

Description

This is the first of two sequential modules on quantitative methods for educational measurement. Students will learn and apply techniques essential for the design and analysis of educational and psychological assessments, including reliability, generalizability theory, validation, differential item functioning, item response theory, scaling, linking, standard setting, and adjustments for measurement error. Contexts of assessments include small-scale educational and psychological assessments for targeted research studies as well as large-scale district, state, and national assessments for formative, summative, and evaluative purposes. In this first module, the emphasis will be on learning and applying methods in class and through completion of data analytic assignments. In the second module, S-061A2, which students are required to enroll in subsequently, methods training will continue, with greater emphasis on reading and critiquing recent research in educational measurement and the development of a research proposal that has promise for advancing the field.

Prerequisite: S-052 or at least two semesters of applied statistics that includes estimation and interpretation of logistic regression coefficients. Enrollment in S-061A2 in the same semester is required. This course complements S-043 and S-090, and students may enroll in these courses in any order. Students who do not meet the prerequisite may enroll instead in S-011, which provides a nontechnical introduction to educational measurement.

Grading

The requirements of the course include regular attendance (5%), regular participation in class and in out-of-class Google Doc discussions (15%), satisfactory completion of 3 assignments (collectively 40% of the unadjusted course grade), and a take-home final exam (40% of the unadjusted course grade). These weights are approximate—the final course grade may factor in improvement over time and exemplary performance on one or more dimensions. Students are required to complete assignments in pairs. The final exam must be completed individually, without assistance from any “animate” resources.

This course is letter-grade-only; students may not take this course on a Satisfactory/No Credit basis. Registered students must submit a course evaluation form at the end of the semester in order to fulfill the requirements of the course. In-person auditing of this course is not allowed—all attendees must be registered students.
Support
Our TF, Masha Bertling, will hold weekly office hours by appointment. Occasionally, Masha may also offer optional discussion sections as she sees necessary; these may be most common around assignment due dates. Participation is strongly recommended. Office hours with the instructor are typically available after class meetings and are also available by appointment. Contact my assistant, at Wendy_Angus@gse.harvard.edu to schedule appointments. Occasional one-on-one check-ins with the instructor are strongly recommended.

CALENDAR

<table>
<thead>
<tr>
<th></th>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>September</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31 – Class 1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2016</td>
<td>4</td>
<td>5 – Holiday</td>
<td>6</td>
<td>7 – Class 2</td>
<td>8 – Asgn 1 Released</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>12 – Class 3</td>
<td>13</td>
<td>14 – Class 4</td>
<td>15</td>
<td>16 – Asgn 1 Due 5PM</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>19 – Class 5</td>
<td>20</td>
<td>21 – Class 6</td>
<td>22 – Asgn 2 Released</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>26 – Class 7</td>
<td>27</td>
<td>28 – Class 8</td>
<td>29</td>
<td>30 – Asgn 2 Due 5PM</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>October 20</td>
<td>2</td>
<td>3 – Class 9</td>
<td>4</td>
<td>5 – Class 10</td>
<td>6 – Asgn 3 Released</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>10 – Holiday</td>
<td>11</td>
<td>12 – Class 11</td>
<td>13</td>
<td>14 – Asgn 3 Due 5PM</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>17 – Class 12</td>
<td>18</td>
<td>19 – S-061A2 Begins… Class 13</td>
<td>20 – Take home final out: 12PM</td>
<td>21 – Take home final due: 5PM</td>
<td>22</td>
</tr>
</tbody>
</table>
I require students to respond to online, Google Doc, discussion questions to central readings by 10PM the night before each class. A good response will demonstrate that the student has read and carefully considered the central reading and spent time considering what the discussion question is asking. Central readings that are very technical need not be mastered in detail, but do pay attention to notation and the underlying motivation of derivations. I intend noncentral readings as additional context and citations for future reference. Links will work on campus and, if you are off campus, if you have a VPN connection. Other readings are available via the iPa© tab on the Canvas website.

There are two required textbooks that are available for purchase:


August 31 – Class 1: Validation. SAT/MCAS example (read 1 and 3, skim 5, 7, 10, 11, glance at 2, 4, 6, 7, 8, 9)

1) Koretz (2008), Chapter 2, pp. 16-34. (required text)
2) Koretz (2008), Chapter 9, pp. 215-234. (required text)
3) AERA/APA/NCME Standards, Chapter 1, pp. 11-31. (required text)
11) 2014 MCAS Technical Report Section 3.9: [link]
September 7 – Class 2: Reliability and Classical Test Theory… an MCAS example (read 1 and 5, skim 3, glance at 2 and 4)

1) AERA/APA/NCME Standards, Chapter 2, pp. 33-47. (required text)
2) Koretz (2008), Chapter 7, pp. 143-178. (required text)
5) 2014 MCAS Technical Report Section 3.7: [link](#)

September 12 – Class 3: Generalizability Theory (read 1, skim 3, glance at 2)


September 14 – Class 4: Generalizability Theory (read 2 skim 1)

1) Ho, A.D. & Kane, T.J. (2013). *The reliability of classroom observations by school personnel*. Bill & Melinda Gates Foundation. [link](#)

September 19 – Class 5: Scaling and Item Response Theory (read 2-5, skim 1)

2) AERA/APA/NCME Standards, Chapter 5, pp. 95-109. (required text)
4) 2014 MCAS Technical Report Section 3.6: [link](#)

September 21 – Class 6: Item Response Theory (read 1, skim 2)


**September 26 – Class 7: Item Response Theory – Polytomous Items. Linking (read 1 skim 2)**


**September 28 – Class 8: Bias, Differential Item Functioning, and Accommodations (read 1 and 2, skim 3 and 4)**

1) AERA/APA/NCME Standards, Chapter 3, pp. 49-72. (required text)

**October 3 – Class 9: Standard Setting and Criterion-Referenced Reporting (read 3, skim 1 and 2)**


**October 5 – Class 10: Growth and Vertical Scaling (read 1 and 3, skim 2 and 4)**


**October 12 – Class 11: Test-Based Policy Metrics (read 1 and 2, skim 3)**

1) AERA/APA/NCME Standards, Chapter 12 and 13, pp. 183-213. (required text)
2) Koretz (2008), Chapter 10, pp. 235-259. (required text)

**October 17 – Class 12: Psychometric Frontiers (read 1 and 3, skim 2)**


### Additional Optional Texts

All students should consider #1, below, for their reference library. Application-oriented students should consider #3 and #6 for Generalizability Theory and IRT, respectively. More technically oriented students should consider #4 and #5. Students interested in practical methods for large-scale testing should consider #2.


3) Generalizability Theory: A Primer. ISBN: 978-0803937451

4) Generalizability Theory. ISBN: 978-1441929389


6) Applications of Item Response Theory to Practical Testing Problems. ISBN: 978-0898590067
Statistical and psychometric computing

Statistical computing is an integral part of S-061. I will be using Stata this year, and you will require Stata 14 to use Item Response Theory methods. I assume that everyone is comfortable using a computer to perform basic statistical analysis, although I don’t necessarily assume that you’ve used Stata.

I do not teach programming during class time, although code is threaded through the lecture slides. We provide resources to help you learn how to program on your own at your own pace. Masha may also cover coding issues in their sections.

There are two ways you can access Stata. The least expensive option is to use one of the networked workstations available in the Learning Technology Center (LTC) on the 3rd floor of Gutman Library and elsewhere on the HGSE campus (e.g., on the 2nd and 4th floors of Gutman Library). For students who would like to use Stata on their own PCs, you may purchase Stata following this link: http://www.stata.com/order/new/edu/gradplans/student-pricing/. Stata/IC, which will be sufficient for this course, is available for $75 for a 6-month license and $198 for a perpetual license.

Collaboration and study groups

Many people learn best when working in a group, and I encourage collaborative learning. To mimic statistical and psychometric work in the real world and to provide a chance for you to use this language actively, I mandate completion of assignments in pairs throughout the course, excepting only the final exam.

We mandate collaboration for at least three reasons. First, learning statistical and psychometric methods is like learning a language. To learn it, one must “speak” it actively and in a genuine context with other individuals. Second, collaborative quantitative analysis is the norm and individual work is the exception in the world of practice. Third, my experience has been that, on average, students who work in pairs and groups both perform better and enjoy themselves more than students who work individually. Statistical and psychometric collaboration is a case where the whole is greater than the sum of its parts.

Beyond pairs, study groups can be helpful to you as you prepare to do the assignments, both in terms of how to approach the work (including how to use the computer effectively) and in terms of how to think about important concepts. However, students must turn in work as pairs or individuals where specified above, not group work. Papers should be written in the pair’s own words—your text should reflect your own understanding of the material.

Each group will undoubtedly develop its own structure; nevertheless, here are a few suggestions:

- Groups with six or more members become less useful and may be harder to organize because finding common meeting times becomes increasingly problematic.

- Plan at least one session of 1½ to 2 hours (early enough so that there is sufficient time if an additional session is necessary). After 2 hours of statistics, everyone’s eyes will be glazing over.

- Schedule the meetings so that you have sufficient time afterwards to write in pairs or individually. When we read your assignments, we focus on what you say and how you say it. The assignments have been devised to require not only computation and programming skills, but skills in analyzing and reporting the material.

- Use the groups to ask questions, try out interpretations, and so on—you each represent each others’ resources. Often one person can explain something that makes you see something in a new way—
or the other way around. Different people have different insights and strengths—some are good programmers, some ask good questions, others value contextual analysis—and you can learn from listening to what others in a group have to offer.

• **Be careful about sitting in groups at laptops or computers and simultaneously composing text.**
  You and your partner must write your own paper, on your own, using your own language. **Your papers should be written in your own words, not those of your study group.**

• Be sensitive to the distinction between collaboration to plan for and interpret the assignment and collaboration to write up the assignment. The former is encouraged; the latter is forbidden beyond, when applicable, your partner. If the distinction begins to feel murky, refocus your group's work on lecture content and course materials. And see me if you have any questions at all.

**Accommodations**

Students needing accommodations in instruction or testing must notify the instructor early in the semester, and HGSE’s policies must be followed. Late requests for accommodations will not be honored unless there is a pressing reason, such as a recent injury.

**A Note on Plagiarism**

Please read the School’s policy on plagiarism in the HGSE Student Handbook, which includes the statement, "Students who submit work either not their own or without clear attribution to the original source, for whatever reason, ordinarily will be dismissed from the Harvard Graduate School of Education." Attention to this policy is particularly important in a course like S-061, in which collaboration with other students is often required and generally encouraged. If you work closely with other students or partnerships—a process that I encourage and fully support—recognize the other students' contributions explicitly in your written account (a footnote is fine for this purpose). This helps avoid the natural questions that arise when similarities are detected at grading. **If you have any questions about what constitutes appropriate collaboration, or how to define what constitutes your own work, please see me or a Teaching Fellow.**

*I cannot overemphasize the need for all students to monitor their own behavior. Assignments are structured such that you can receive feedback on your and your partner’s understanding of the material. The consequences for plagiarism are appropriately severe.*

**Other Writing Resources**

- HGSE Academic Writing Services: Gutman Library
- APA Online Tutorial: [http://isites.harvard.edu/icb/icb.do?keyword=apa_exposed](http://isites.harvard.edu/icb/icb.do?keyword=apa_exposed)
- Writing Resources (including Writing Like an Educator Course and Reference Materials): [http://isites.harvard.edu/icb/icb.do?keyword=awrs&pageid=icb.page48297](http://isites.harvard.edu/icb/icb.do?keyword=awrs&pageid=icb.page48297)
- Sign-up for Individual Sessions at the Writing Center: [http://www.appointmentquest.com/provider/2030159020](http://www.appointmentquest.com/provider/2030159020)
Finally, some tips from last year’s S-061 student cohort follow. These are the responses to my favorite question (#9) from our course evaluations. I have not edited or omitted any responses.

9. What advice would you give to students who are thinking of taking this course (about its level, the amount of work required, any prior training needed, ways to get the most out of the course, etc.)?

Ask questions often, and especially feel free to ask for clarifications. The material is sometimes quite dense, and just one explanation might not cut it. But an understanding of the content is invaluable.

This is a fantastic course, but be prepared to work hard and maintain a fast pace.

Make sure you have the time required!

This is a very natural continuation of S-052 (even though S-052 is not necessarily required) - expectations are about the same, the workload is manageable, the amount of readings is light (yet relevant), there is very little pressure on you

Budget a significant amount of time outside class to work through class lectures and examples, and homework

This is intensive course that requires much of background in statistics and material. Be prepared to invest your time in learning the material.

This is a great introductory psychometrics class that gives a high level overview of the major topics within the field of educational measurement. The instructor and the quality of delivered materials are outstanding. Professor Ho is a very skilled and experienced teacher. I especially liked the assignments and class discussions that built heavily upon the most recent and relevant practical issues in the field, as well as the key texts and articles by leading researchers instead of a “one size fits all” textbook. Some potential drawbacks of the class are its module structure and the coverage of a very large number of topics that provide a limited opportunity to dive deep into any specific issue. Additionally, the time investment and the workload is very high. One needs to spend at least 20 hours on a weekly basis and the number will increase during the weeks when assignments are due. I think it would be great to have such a class as a year long sequence to ensure deeper coverage of the content. For some topics, such as IRT, it would be very beneficial if HGSE considered to offer a separate class, which (without doubts) would have a high demand. The need for advanced and applied quantitative methods within HGSE is evident.

It’s best to have a well developed idea about the project early on. The first part of the course provides a nice time to think about different papers, and this module should not be spent thinking about what to do. It’s a short 6 weeks.

Read as much as you can and be prepared for a tough and high value course

Start looking for a project topic early, and check in with the teaching team often about your ideas. They are fantastic at helping you to decide what’s the best course of action given some data or some interest.

This course may help you get your paper prepared for publication.

This is a fantastic course for people interested in learning advanced psychometric statistical techniques such as IRT and Grth theory. Expectations from the instructor are appropriately high, and there is a lot of work required for the course, but it is well worth it.

Come with a good paper idea and with your own data-set.