# PP449 Methods of Data Collection: Experiments, Quasi-Experiments, and Surveys

## Winter 2026

## Colm O'Muircheartaigh

**Mondays and Wednesdays 9:00-10:20am [***Except*** Week 3 Wednesday Jan 21 and Friday January 23 9-10:20am]**

**Office Hours**

*Colm O'Muircheartaigh*: Thursdays 2-3:30pm in Room 3021.

*Sushant Banjara*: TBA

**Course Description**

The pressure in many fields (notably medicine, health research, politics, and education) for evidence-based results has increased the importance of the design and analysis of social investigations in providing a basis for policy decisions. This course will address: (i) the design of experiments, quasi-experiments, and surveys; and (ii) the use of these social investigations to provide data for generalization. The course will explore the relative relevance of evidence from these different sources in formulating policy.

Randomized clinical trials in medicine, field experiments in economics, psychology and political science, tests of quasi-experimental interventions, and national sample surveys will be among the examples.

**General Notes on In-Person Classes**

Computer and phone use: Laptops, phones, and other electronic devices may NOT be used in class. If you must make (or take) a call or text, you must leave the room to do so.

Name cards: Bring your name card to all the classes (and display it)

**Attendance and Participation***: Attendance and participation count for 20% of the grade*

Attendance: Attendance will be taken at every class.

Participation: I believe that active participation improves classes for everyone. Questions from members of the class are always welcome. I also will ask questions frequently during class. Volunteering answers is excellent but, in order to make sure that everyone gets an opportunity to display interest and knowledge, I will also call regularly on those who do not volunteer. Attendance at office hours and written submissions will also count towards the participation score.

**Homework Assignments***: The assignments will account for 30% of the class grade. The homework grade will exclude the weakest score (late homework assignments will receive a zero score).*

There will be 7 or 8 homework assignments, mostly short. These will include practice exam questions from previous years. Due dates for assignments will be indicated in the schedule and announced in class.

**Final:** *The in-class 2-hour final will count for 50% of the grade*

**Course Materials:**

**Topic-specific materials will be posted on the Canvas site each week.**

**The following three books provide background reading:**

A S Gerber and D P Green: *Field Experiments: Design, Analysis, and Interpretation*. W W Norton and Co. 2012

W R Shadish, T D Cook, and D T Campbell: *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Houghton Mifflin Co. 2002

G Kalton: *Introduction to Survey Sampling* (Quantitative Applications in the Social Sciences). A Sage University Paper. 1983

**Class materials**

A draft version of the slides to be presented in class will be posted on Canvas in advance of each class.

Each assignment will be posted on Canvas a week ahead of the due date.

**Grading**

The grade will be based on three components: class attendance and participation (20%), homework assignments (30%) and a final (50%). Attendance at office hours and written questions will count towards participation.

**Syllabus:**

*The course is not linear; topics will be addressed in one context and revisited in others as we move from one type of investigation to another.*

*There are two fundamental challenges in (social) science): first, how to establish a plausible argument that a change in one factor leads to a change in another (causation); second, how to generalize what you have observed in one data set to units or circumstances that you have not observed (generalization). This course frames different types of social investigations in terms of how they address these challenges.*

*The first module introduces the concepts of causation and generalization and the challenges of designing a set of studies that provides a basis for both. Randomized experiments, quasi-experiments and social surveys will be located in a general framework for social investigations. Module 2 focuses on the formal meaning of causation and its implementation in social investigations. In module 3, the three foundational designs for randomized experiments will be presented. The range of experimental designs will be illustrated by examples from two very different contexts. In module 4, the requirements of designs in realistic social science settings will be discussed – experiments in hierarchical settings [organizational or social] -- and the Implications for interpretation of results and their standard errors. Modules 5 and 6 deal with quasi-experimental designs and other designs where full randomization is not possible. These designs are very helpful in understanding the challenges in designing and interpreting all experiments and are particularly relevant in the public policy sphere. The strengths and weaknesses of randomized and non-randomized experiments will be discussed. Modules 7 and 8 of the course deal with (scientific) sample surveys, the dominant form of non-interventionist social investigation in the public sphere. The main sample designs will be presented, along with their practical and technical strengths and weaknesses. Module 9 will bring together the three strands and re-visit the challenge of strategic design for social research.*

**Module 1: Introduction** [January 5]

1.1   Introduction to generalization: experiments, surveys, and observational studies

1.2   An early experiment: Fisher (1936)

Assignment 1 (due January 11, midnight CT): Inference from an early experiment; Darwin, Galton, and you.

**Module 2 Design of Experiments** [DE] [January 7 and 12]

2.0 DE 0:   History of experimental design

2.3 DE 1:   Stable Unit Treatment Value Assumption [SUTVA]

2.1 DE 2:    Causation; Average treatment effect [ATE]

2.2 DE 3:   Control through matching and blocking

2.4 DE 4:   Terminology

Discussion of assignment 1

Assignment 2 (due January 18): Evaluate NYT Editorial opinion on COVID clinical trials

**Module 3: Examples of experiments** [January 14 and 21]

3.0 Example: Jellinek (1946)

3.1 Example:  Current Population Survey redesign [applied test of new processes];

3.2 Example: The Very Study [cross-disciplinary context for results]

3.3 DE 5:   Three foundational designs: completely randomized, randomized block, and hierarchical designs

3.4 DE 6:   Use of covariates in analysis

Discussion of assignment 2

Assignment 3 (due January 25): Generalizing the results of an experiment examining racial bias in medicine

**Module 4: More complex experiments** [January 23 and 26]

4.1   DE 7:   Precision in clustered models

4.2   DE 8:   Relationships in hierarchical systems

4.3   DE 9:    Raudenbush generalization; IES guidelines for evidence in education studies

Discussion of Assignment 3 (January 28)

Assignment 4 (due February 1): Thinking More About Generalization

**Module 5: Quasi-experiments** [January 28 and February 2]

5.1            Definitions of evidence: IES guidelines; public health evidence on Covid

5.2 QE1        Challenges to quasi-experiments; internal validity

5.3 QE2        External validity

Discussion of Assignment 4 (February 4)

Assignment 5 (due February 8): Comparing competing claims

**Module 6: Quasi-experiments continued** [February 4 and 9]

6.1 QE3        Basic designs; case-control designs

6.2 QE4        More complex designs

                Experiments within surveys

Discussion of Assignment 5 (Feb 11/18)

Assignment 6 (due February 15): Practice exam questions

**Module 7: Survey Design** [February 11 and 16]

7.1 SS1:        Survey sampling: background and principles

7.2 SS2:        Design-based inference; random sampling, stratification, clustering

7.3 SS3:        Design effect, effective sample size, precision

Discussion of Assignment 6 (February 18)

Assignment 7 (due February 22): Computation of design effect and effective sample size +sample size + stratification allocation

**Module 8: Combining surveys and experiments** [February 18 and 23]

New 8.1 SS3A:        ABS + US National Sample Design

8.2 SS4:        Population inference and experiments

8.3 SS5        Strategic design [O'Muircheartaigh and Hedges, 2014]

Discussion of Assignment 7 (February 25)

Assignment 8 (due March 3): Final practice question

**Module 9: Causation and Generalization** [February 25 and March 2]

8.3 SS5 (cont) Strategic design

9.1 CG1  Generalization without sampling

9.2 CG2  Meta-analysis

**Module 10: Wrap-up** [March 4]

10.1: CG3  IES, Bradford Hill, 'Student'

10.2: CG4  From '*It works somewhere'* to '*It works here*'

Review, including discussion of Assignment 8