Research Design Issues in Experimental Studies

Jonathan J. Koehler
Northwestern University
School of Law
jay.koehler@northwestern.edu

Nick Schweitzer
Arizona State University
New College of Interdisciplinary Arts and Sciences
njs@asu.edu

Part I:
Experiments…

Jay's Outline

- 1. What is a controlled experiment?
- 2. What are some of the issues to watch out for?
- 3. What do studies and researchers say?
- 4. What do courts say?
- 5. What should you do?

What is a Controlled Experiment?

- A scientific method for testing causal relationships among variables
- Randomly assign (relevant) participants to groups
  - experimental or control
- Manipulate independent variables
- Measure dependent variables

Subtle Confounds

- Manipulations may contain hidden biases
- Sex bias study

Don Smith

5321 Anywhere Street., Apt. 942
San Francisco, CA 94321
(415) 864-6692
anyemail@hotmail.com

EDUCATION

Golden Gate University School of Law — San Francisco, CA
Doctor of Jurisprudence Candidate, May 2007
- Honors Lawyering Program, Participant
- Foundation of the State Bar of California, Merit Scholarship Recipient for 2005 and 2006
- Public Interest Law Foundation (PILF): Board Member, Fundraising Auction Co-Chair
- National Lawyers Guild: Member

University of Virginia — Charlottesville, VA
Bachelor of Arts, Biology, January 1997

LEGAL EXPERIENCE

San Francisco Superior Court — San Francisco, CA
Judicial Extern — August-December 2006
Worked as a law clerk to San Francisco Unified Family Court Judge Anne-Christine Massulo; prepared case summaries, researched legal issues, and prepared orders in court.

Family and Children's Law Center — San Rafael, CA
Law Clerk/Client Advocate — May-August 2006
Did intake and counseling sessions with clients for dissolution proceedings, child custody matters, property division, and domestic violence matters; represented clients in court for restraining order hearings; wrote and filed motions; drafted Marital Settlement Agreements and Trial Briefs.

Bay Area Legal Aid — San Francisco, CA
Law Clerk, Family Law and Domestic Violence Division — September 2005-May 2006
Researched and wrote motions, pleadings, memoranda, and legal briefs; completed and filed Judicial Counsel forms; conducted investigations and discovery; handled client correspondence.

W.O.M.A.N., Inc. (Women Organized to Make Abuse Nonexistent) — San Francisco, CA
Crisis Services Program and Manager — October 2002-September 2003
Counseled survivors of domestic abuse at a community-based domestic violence nonprofit organization. Coordinated the crisis line as program manager including scheduling, operation and 24-hour on-call support for other counselors. Managed 30+ crisis line volunteers including recruitment and training.

SKILLS AND CERTIFICATIONS

- Proficient in Legal Solutions, Essential Forms, Microsoft Office Suite, and various database software programs; advanced statute and case law research certification for Westlaw and Lexis.
- Certified by the State Bar to represent clients through the Practical Training of Law Students Program.
- Certified in the State of California as a Domestic Violence Crisis Counselor.
- Conversant in French.

Subtle Confounds

- Manipulations may contain hidden biases
- Sex bias study
- Kasof (1993, *Psych. Bull.*)

Donna Smith

5321 Anywhere Street., Apt. 942
San Francisco, CA 94321
(415) 864-6692
anyemail@hotmail.com

EDUCATION

Golden Gate University School of Law — San Francisco, CA
Doctor of Jurisprudence Candidate, May 2007
- Honors Lawyering Program, Participant
- Foundation of the State Bar of California, Merit Scholarship Recipient for 2005 and 2006
- Public Interest Law Foundation (PILF): Board Member, Fundraising Auction Co-Chair
- National Lawyers Guild: Member

University of Virginia — Charlottesville, VA
Bachelor of Arts, Biology, January 1997

LEGAL EXPERIENCE

San Francisco Superior Court — San Francisco, CA
Judicial Extern — August-December 2006
Worked as a law clerk to San Francisco Unified Family Court Judge Anne-Christine Massulo; prepared case summaries, researched legal issues, and prepared orders in court.

Family and Children's Law Center — San Rafael, CA
Law Clerk/Client Advocate — May-August 2006
Did intake and counseling sessions with clients for dissolution proceedings, child custody matters, property division, and domestic violence matters; represented clients in court for restraining order hearings; wrote and filed motions; drafted Marital Settlement Agreements and Trial Briefs.

Bay Area Legal Aid — San Francisco, CA
Law Clerk, Family Law and Domestic Violence Division — September 2005-May 2006
Researched and wrote motions, pleadings, memoranda, and legal briefs; completed and filed Judicial Counsel forms; conducted investigations and discovery; handled client correspondence.

W.O.M.A.N., Inc. (Women Organized to Make Abuse Nonexistent) — San Francisco, CA
Crisis Services Program and Manager — October 2002-September 2003
Counseled survivors of domestic abuse at a community-based domestic violence nonprofit organization. Coordinated the crisis line as program manager including scheduling, operation and 24-hour on-call support for other counselors. Managed 30+ crisis line volunteers including recruitment and training.

SKILLS AND CERTIFICATIONS

- Proficient in Legal Solutions, Essential Forms, Microsoft Office Suite, and various database software programs; advanced statute and case law research certification for Westlaw and Lexis.
- Certified by the State Bar to represent clients through the Practical Training of Law Students Program.
- Certified in the State of California as a Domestic Violence Crisis Counselor.
- Conversant in French.

Random Assignment

Special Problems With Experiments

- Researchers want results to apply in real legal settings.
- Researchers want to persuade judges (or other legal actors) that their results apply

1. Who are the Participants?

- Empanelled jurors
- People called for jury duty
- National samples screened for jury-eligibility
- Community members
- Students

2. What's at Stake?

- No consequences

3. What Materials are Used?

- Written
  - Simple summaries (1 page)
  - Detailed (judicial welcome, opening arguments, witness statements, closing arguments, judicial instructions, deliberation)
- Audio
- Video

4. Did Jurors Deliberate?

- Most mock jury studies don't involve juries. Research is "done on the juror level and then extrapolated to the jury"

  - Nunez et al., 2011, *Behav. Sci & Law*

Responses From the Academy

- Participants
- Stakes
- Materials
- Deliberation

1. Participants? Not (too) Worried

" [T]he overwhelming majority of studies that have directly compared different mock juror samples have failed to find consistent differences…[There is] strong evidence that factors at trial affect students and nonstudents in the same way."

- Bornstein, 1999, *Law & Hum. Behav.*

Participants?
Not (too) Worried

- Meta-analysis of 23 studies: No differences across type of participant.

- "[F]ears regarding comparability of student and community data may be unwarranted…. [W]e failed to identify any significant effects of sample type on performance for the 16 items included in this review."

  - Desmarais & Read, 2011, *Law & Hum. Behav.*

Participants? Worried

- "Jury venires should remain the preferred source for maximizing both face and external validity"

- Diamond, 1997, *Law & Hum. Behav.*

see also Weiten & Diamond, 1979, *Law & Hum. Behav.*

- Use samples that are "closer to the population of ultimate interest"
  - Wiener et al. 2011, *Behav. Sci. & Law*

2. Stakes? Might Matter

- "Findings from the lab based on college students making low-stakes or hypothetical decisions in limited interactions … may provide little guidance to the behaviors likely to be observed under the conditions of the real workplace involved in the case."

  - Monahan, 2011, *Emory L. J.*

Incentives Data

- Review of 74 experiments with no, low, or high performance-based financial incentives.

  - Camerer & Hogarth, 1999, *J. Risk & Uncertainty*

- Incentives generally don't affect performance quality (though they often decrease variance).

- Incentives most effective for judgment tasks that are responsive to greater effort.

3. Materials?

"Studies that have directly compared presentation media [e.g., written summaries, transcripts, audiotape, videotape] – for either a whole trial or a portion of testimony – fail to offer consistent findings. . . . [R]esearch on the trial medium tends not to find many differences."

- Bornstein, 1999, *Law & Hum. Behav.*

4. Deliberation?

- Intuition: Deliberation corrects misunderstandings, reduces error.

- One Possibility: Deliberation increases confidence but doesn't increase accuracy (much).


  - Kaye, Hans, et al. (2007, *J. Emp. Legal Stud*

  - When most jurors are confused "deliberations may simply reinforce the inaccuracies of the majority" (Diamond, 1997).

Responses From the Courts

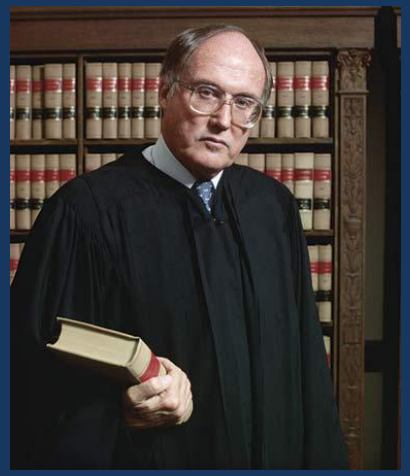
- “This study is flawed”
- “This study is inapplicable”

Participants & Stakes

- Participant concerns

- "Serious doubts" about studies that are "based on the responses of individuals randomly selected from some segment of the population, but who were **not actual jurors** sworn under oath to apply the law to the facts of an actual case **involving the fate of an actual capital defendant**."

Lockhart v. McCree (S. Ct. 1986)



Materials & Deliberation

- "[The study] must be discounted because the people interviewed for the study … were given hypothetical facts that were different than the facts in this case, and they did not hear the testimony of witnesses, observe physical evidence or deliberate with eleven other jurors."

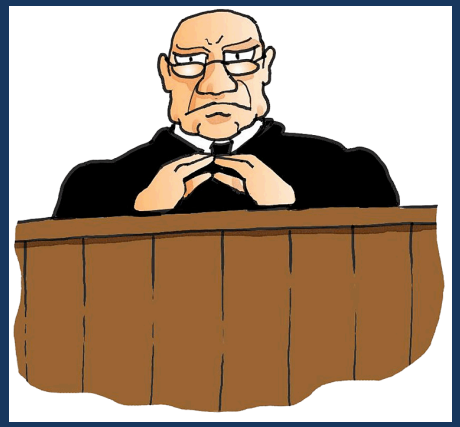
  - *State v. Deck* (Missouri S. Ct. 1999)

- Criticizes studies that "did not even attempt to simulate the process of jury deliberation"

  - *Lockhart v. McCree* (S. Ct. 1986)

Recommendations

- Clean manipulations
- Variation & replication
  - Participant populations, incentives, materials, deliberation



- Field studies

Part II:
Experimental Design

Outline

- Experimental design as the pursuit of causal description
- The problem of confounds
- Basic types of experimental designs

Experiments and Causation

- Experimental design is meant to fulfill necessary criteria for causal inference:
- Temporal Precedence
  - Assured by control over administration of "cause" and observation of "effect."
- Covariation
  - Demonstrated by appropriately chosen and correctly interpreted statistical analysis
- Non-spuriousness
  - Most difficult to ensure

Fulfilling the non-spuriousness criteria: Dealing with unwanted variables

- Two varieties of variables that may influence findings
- Disturbance variables:
  - Suspected of influencing the dependent measure
  - N**ot** related to the independent variable.
  - Individual Level: Variables that differ from participant to participant. Effects: Increased error variability.
  - Experiment Level: Variables that are specific to the experimental environment. Effects: Decreased generalizability.
  - Dealing with disturbance variables: Standardize experiment procedure; speculate as to potential disturbance variables and then measure and remove w/ ANCOVA or similar.

Fulfilling the non-spuriousness criteria: Dealing with unwanted variables

- Confounding Variables
  - Suspected of influencing the dependent measure
  - Are related to the independent variable.
    - Perfect Confound (1.00 correlation with IV)
    - Imperfect Confound (<1.00 correlation with IV)
  - Effects: Ruins your study. Sometimes not obvious, and may be introduced in various ways.
  - Imperfect confounds can potentially be dealt with statistically.

Sources of Confounds

- Experimenter Error:
  - Use of improper design or incomplete factorial designs
    - Example: Mecklenberg study
  - Manipulations that are too broad (& manipulate more than one IV).
  - Experimenter bias

Sources of Confounds

- Non-random assignment
  - Non-manipulatable variables
    - Lawyer vs. Layperson judgments of rule violations. Gender effect discovered.
  - Selective informed consent
    - Mock-trial paradigm; rape case vs. burglary; capital case vs. non-capital (death qualification)
  - *Side question: Why is IV random assignment essential to experimentation?*
  - *Answer: Average correlation of X and random sequence is 0 – thus, reduces chance of unmeasured variable being correlated with IV; turns potential confounds into disturbance variables.*

Sources of Confounds

- Inadvertent / Environmental variables
  - Different experimenters, noise, different rooms, difficulty with the protocol, etc.
  - More problematic when participants are run in groups than individually (e.g., random assignment by session / jury)

Sources of Confounds

- Dealing with confounds:
  - Imperfectly correlated confounds (e.g., gender & lawyer from above) can be addressed statistically (ANCOVA, factorial designs).
  - Perfectly correlated confounds cannot be addressed without collecting additional data.
  - Prevention and measurement are key.

Basic Experimental Designs

- Experimental condition vs. control condition
- Carefully chosen control group(s) help eliminate confounds
- Example: What influence does neuroimagery have on factfinders' judgments of defendants?
  - Expert evidence w / Neuroimage versus what?
    - No expert evidence
    - General psychological expert evidence
    - Neurologically-framed expert evidence
    - Neurologically-framed and graphically depicted evidence?

Basic Experimental Designs

- **Between Subjects Design**
  [Tx -> O] vs. [Control -> O]
  - **Example:** Perceived bias in judicial decisions by manipulating affiliations of judge
  - **Benefits:** Simple, short, difficult for participants to be sensitized to manipulation, lacks potential confounds found in other designs.
  - **Challenges:** Each group of experiment requires sufficiently large sample. Despite use of random assignment, groups may still have confounding differences. Not as "powerful" as other designs.

Basic Experimental Designs

- **Repeated Measures Design [O -> Tx -> O]**
  - **Example:** Measuring juror attitudes pre/post deliberation or after each witness
  - **Benefits:** Participants compared to self—eliminates individual differences. Powerful design, easier to find effects. Fewer participants needed.
  - **Challenges:** Introducing time as a confound; must be able to ensure time does not bring about other changes to DV (maturation, attrition, history). Pre-test measures may sensitize participants to IV (relative effects) or change post-test scores (order effects, testing). Experiments take longer and may require multiple sessions (produces attrition).

Experimental Designs

- Mixed Designs: Pre-Post w/ Control:
- **[O -> Tx -> O] vs. [O -> Control -> O]**
  - **Example:** Perceived likelihood of mental defect before & after expert evidence involving neuroscience or not
  - **Benefits:** Use of pre-test eliminates individual differences & increases power. Use of separate condition w/ control group allows to assess any effects of time/sensitization.
  - **Challenges:** Takes as long as the repeated measures design, but can require about the same number of participants as a between-subjects design. Effect manifests as interaction, can be more difficult to find statistically.

Lab vs. Field Experiments

- Field experiments involve natural settings, high ecological validity.
- Some in semi-controlled environments (e.g., Arizona jury innovations)
- Some in less structured environments (e.g., Minnesota et al. domestic violence experiments)
  - Idiosyncrasies of situations can still limit external validity.
  - Difficult to maintain random assignment.
  - Ethical concerns with consent.

Suggested Reading

- Bornstein (1999) The ecological validity of jury simulations: Is the jury still out? Law & Human Behavior.
- Diamond (1997) Illuminations and shadows from jury simulations. Law & Human Behavior.
- Lawless, Robbennolt, & Ulen (2010) Empirical Methods in Law.
- Shadish, Cook, & Campbell (2002) Experimental and Quasi-Experimental Designs for Generalized Causal Inference.
- Wiener, Krauss, & Lieberman (2011) Mock jury research: Where do we go from here? Behavioral Sciences & the Law.

Thank you.