Problems and Diagnostics

1. Common Implementation Problems
   1. Types of problems:
      1. Compliance problems:
         1. Incomplete control over treatment delivery. Use intent-to-treat or placebo design to solve this.
         2. Accidental delivery to control group (spillovers)
      2. Randomization doesn’t work as intended
      3. Treatment not delivered or not received (e.g., letters not read or emails not opened)
      4. Unintended effects of treatment.
      5. Hawthorne Effect: the effect of change or novelty. Social experiment conducted at Hawthorne factory. Every variable seemed to affect productivity but effects were only temporary.
      6. Demand Effect: subjects who are aware of being studied often give different answers. May give answers they think the researcher wants to hear. Treatment can alert subjects to what researchers are looking for. Researchers try to conceal the connection between survey and treatment.
      7. Unexpected implementation results: abstract concepts may result in missed real-world contexts (e.g., field workers refuse to randomize or think they are helping)
   2. Blind Trials: eliminates fishing expedition problem; biased analysis may cause study to be halted prematurely.
      1. Single-blind trial: subject doesn’t know
      2. Double-blind trial: experimenter doesn’t know
      3. Triple-blind trial: analyst doesn’t know group assignments.
   3. Detecting Errors
      1. Pilot studies: discover unanticipated problems. Always run.
         1. Unanticipated reactions (e.g., fake emails to state legislators intended to study responses based on geographic origin of messages but resulted in spillover between legislators)
         2. Flawed power calculations: likelihood of effect of different sizes, baseline, response rates, variance of y
         3. Training staff for correct implementation
         4. Determining if systems really work
         5. New ideas, potential for improvement
      2. Placebo Tests: if experiment/experimentation system works, there should be no difference n a variable.
         1. A/A test: tests a treatment against itself to detect any difference in outcome
         2. Traditional placebo test: checks other outcomes that treatment shouldn’t affect. To prove bias, look for differences that shouldn’t exist if assumption is right.
            1. Gerber and Green Placebo Treatment: blood drive non-compliance. Can’t compare compliers to everyone in control group. Compliers tend to be systematically different. Compare entire treatment group to entire control group or treatment compliers to control compliers.
      3. Manipulation checks: was treatment successfully delivered? If experiment worked, there should be differences in this variable.
      4. Covariate balance checks: checks for problems in implementation of experiments. Experiments guarantee balance in observable and unobservable characteristics. Check balance on observable covariates to ensure random assignment was done correctly.
         1. Especially important when randomization scheme is complex. Blocking clustering, different probabilities. Complex systems between researchers and subjects.
2. Advocating Experimentation
   1. Increased perceived benefits
      1. Stimulate curiosity: intellectual interest and get people excited.
      2. Vivid examples of current data leading to bad decisions.
         1. Conduct placebo test and show practices failing
         2. Present potential conclusions experiment could produce and ways this could change practices
         3. Tell story of why causal inference from observational data might be wrong.
      3. “Investment in information for future decisions”: short-term costs pay off in future. Down payment.
      4. Build rapport: personal connections lead to willingness to run experiments
      5. Do small studies as proofs of concept: helps secure greater cooperation.
   2. Decrease perceived costs
      1. Administration, unfairness, giving up potential gains from treatment (e.g., ads)
      2. Delay for some units: e.g., randomize order of mailings, campaign contributions (spread out donations to allow experimentation)
      3. Limited resources
      4. Experimentation as investment in information
   3. The Long View
      1. Building knowledge over time.
         1. Try pilot studies.
         2. System should produce covariate balance.
         3. Parameters should look as expected (e.g., 80/20 treatment/control split)
         4. Understand parameters for power analysis
         5. When creating a system, studies should be useful.
      2. Pooling results: meta-analysis. Pool across studies, take estimates and then take average of them.
         1. Precision-weighted average: . Double the standard error gets ¼ the weight. ¼ the sample size means ¼ the information. If variance is the same, weight by number of subjects. Standard error used to represent adjustment of overall view of treatment. Meta-analysis summarizes data.