Data Collection: Validity & Ethics

Data Integrity
- Expectancy effects & their control
  - Experimenter expectancy effects
  - Participant Expectancy Effects
  - Single- and Double-blind designs
- Researcher and Participant Bias & their control
  - Reactivity & Response Bias – Problems with participants
  - Observer Bias & Interviewer Bias – Problems w/ researchers
  - Effects of attrition on initial equivalence

- Ethical Considerations
  - Informed Consent
  - Researcher Honesty
  - Levels of Disclosure

Experimenter Expectancy Effects
A kind of “self-fulfilling prophesy” during which researchers unintentionally “produce the results they want”. Two kinds…

Modifying Participants’ Behavior
- Subtle differences in treatment of participants in different conditions can change their behavior…
- Inadvertently conveying response expectancies/research hypotheses
- Difference in performance due to differential quality of instruction or friendliness of the interaction

Data Collection Bias (much like observer bias)
- Many types of observational and self-report data need to be “coded” or “interpreted” before they can be analyzed
- Subjectivity and error can creep into these interpretations – usually leading to data that are biased toward expectations

Participant Expectancy Effects
A kind of “demand characteristic” during which participants modify their behavior to respond/conform to “how they should act”. Two kinds…

Social Desirability
- When participants intentionally or unintentionally modify their behavior to match “how they are expected to behave”
- Well-known social psychological phenomenon that usually happens between individual’s and their “peer group”
- Can also happen between researcher and participants

Acquiescence/Rejection Response
- If participant thinks they know the research hypothesis or know the behavior that is expected of them they can “try to play along” (acquiescence) or “try to mess things up” (rejection response)
- Particularly important during within-groups designs – if participants think study is “trying to change their behavior”
Single & Double-blind Procedures

One way to limit or minimize the various biasing effects we’ve discussed is to limit the information everybody involved has.

In Single Blind Procedures the participant doesn’t know the hypotheses, the other conditions in the study, and ideally, the particular condition they are in (i.e., we don’t tell how the task or manipulation is designed to change their behavior).

In Double-blind Procedures neither the participant nor the data collector/data coder knows the hypotheses or other information that could bias the interaction/reporting/coding of the researcher or the responses of the participants.

Sometimes this simply can’t be done (especially the researcher-blind part) because of the nature of the variables or the hypotheses involved (e.g., hard to hide the gender of a participant from the researcher who is coding the video tape).

Reactivity & Response Bias

• Both of these refer to getting “less than accurate” data from the participants.

Reactivity is the term commonly used when talking about observational data collection.

– the participant may behave “not naturally” if they know they are being observed or are part of a study.
– Naturalistic & disguised participant observation methods are intended to avoid this.
– Habituation and desensitization help when using undisguised participant observation.

Response Bias is the term commonly used when talking about self-report data collection and describes a situation in which the participant responds how they think they “should”.

– The response might be a reaction to cues the researcher provides.
– Social Desirability is when participants describe their character, opinions or behavior as they think they “should” or to present a certain impression of themselves.
– Protecting participants’ anonymity and participant-researcher rapport are intended to increase the honesty of participant responses.

Observer Bias & Interviewer Bias

Both of these are versions of “seeing what you want to see.”

Observer Bias is the term commonly used when talking about observational data collection.

– Both observational data collection and data coding need to be done objectively and accurately.
– Automation & instrumentation help – so does using multiple observers/coders and looking for consistency.

• Interviewer Bias is the term commonly used when talking about self-report data collection.

– How questions are asked by interviewers or the interviewers’ reactions to answers can drive response bias.
– More of a challenge with face-to-face interviews.
– Computerized and paper-based procedures help limit this.

Effects of participant-research gender, race, age, personality, etc. match/mismatch have been shown to influence the behavior of both !!!
### Data Collection Biases & Inaccuracies -- Summary

<table>
<thead>
<tr>
<th>Type of Data Collection</th>
<th>Observational</th>
<th>Self-report</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observer Bias</strong></td>
<td>“inaccurate data recording/coding”</td>
<td>Interviewer Bias “coaching” or “inaccurate recording/coding”</td>
</tr>
<tr>
<td><strong>Reactivity</strong></td>
<td>“reacting” to being observed</td>
<td>Response Bias “dishonest” responding</td>
</tr>
<tr>
<td><strong>Participant</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attrition</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Attrition
- Also known as drop-out, data loss, response refusal, & experimental mortality
- Random assignment is intended to produce initial equivalence of subject variables, so the groups (IV conditions) have equivalent means on all subject variables (e.g., age, gender, motivation, prior experience, intelligence, topical knowledge, etc.).
- Attrition can disrupt the initial equivalence – producing inequalities
- “Differential attrition” – related to IV condition differences – is particularly likely to produce inequalities
- E.g., If one condition is “harder” and so more participants drop out of that condition, there is likely to be a “motivation” difference between the participants remaining in the two conditions (i.e., those remaining in the harder condition are more motivated).

### So, “Attrition” works much like “Self Assignment” to trash initial equivalence

Both involve a non-random determination of who provides data for what condition of the study!

Imagine a study that involves a “standard treatment” and an “experimental treatment”…

- Random assignment would be used to ensure that the participants in the two groups are equivalent
- Self-assignment is likely to produce non-equivalence (different “kinds” of folks likely to elect the different treatments)
- Attrition (i.e., rejecting the randomly assigned condition) is similarly likely to produce non-equivalence (different “kinds” of folks likely to remain in the different treatments)
How to combat attrition…

• educate participants about the important role of random assignment to the validity of the research

• if there is differential “value” of the different treatments or conditions (especially in a “treatment vs no-treatment” comparison) – offer folks an opportunity to participate in the preferred condition after data collection

• replacement of participants who drop out of the study

• If there is a more aversive condition, then ask the participants before assignment if they would still participate even if they were in the aversive condition. Then, only allow the people who say yes to be in the study (note implications for external validity)

• collect data about possible confounding variables for statistical comparison later

• replication & convergence of the study

Ethical Considerations -- participation

Research ethics are summarized in the “risk-benefit” trade-off model.

What do participants risk when participating?
– social (embarrassment), psychological (learning uncomfortable things about themselves), or even physical risk
– risk might be from manipulation, task, data collection or being “associated” with the research

trades off with …
What are the benefits of the research?
– to society (knowledge gained) or the participant (remuneration – pay or research credit or direct benefit of the treatment)

Each university has an Institutional Review Board (IRB) that must approve the manipulations, procedures, data collection and data storage of all research involving human participants, under the review of the federal government. Individuals or universities that violate the relevant guidelines can/have been denied research support (grants), research/data collection privileges and are legally responsible to participants!

Voluntary Informed Consent without Deception

• Before participating each participant must read and sign a document that describes his/her participation (including random assignment) and all related activities, as well as the possible social, psychological or physical risks involved in that participation.

• No information may be withheld from the participant the possession of which might alter her/his decision to give informed consent

• “Deception” is withholding information from the participant that might possibly alter their decision whether or not to participate
  – Sometimes, the IRB approves research with deceptive elements as long as the risk/benefit ratio is ultimately positive

• The participant is free to withdraw informed consent and stop participating in the research at any time with no consequences
  This guarantee is the cornerstone of Ethical Research !!!
Ethical Considerations – reporting research

• When proposing and reporting research, researchers must be completely forthcoming concerning the procedures and resulting data.
• APA format & style is designed to ensure that all the features of the study are adequately represented as part of the research report.
• Remember, the central idea behind “scientific empiricism” is that science is a “public enterprise”
  • not only the results but how those results are obtained is made public
• any properly trained individual should be able to understand procedures and results and (if properly outfitted) to replicate the findings
• this level of procedural disclosure helps replication and convergence efforts

Levels of disclosure

Assuring participants that their responses are “safe” is important when requesting their participation.

• Privacy
  – no one knows the info; Participant does not disclose it (not collected as data)
• Anonymity
  – no direct connection between info and identity
• Confidentiality
  – researcher has connection between info and identity, but doesn’t disclose connection
• Group Disclosure
  – info about “group” is released
  – must avoid “indirect disclosure” for small groups
• “Masked” Individual Disclosure
  – identity is hidden by pseudonym
• Individual Disclosure
  – requires explicit informed consent