### Experimental design

#### Allocation of Units to Groups

<table>
<thead>
<tr>
<th>Selection of Units</th>
<th>By Randomization</th>
<th>Not by Randomization</th>
<th>Inferences to populations can be drawn</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>At Random</strong></td>
<td>A random sample is selected from one population. Units are then randomly assigned to different treatment groups.</td>
<td>Random samples are selected from existing distinct populations.</td>
<td></td>
</tr>
<tr>
<td><strong>Not at Random</strong></td>
<td>A group of study units is found. Units are then randomly assigned to treatment groups.</td>
<td>Collections of available units from distinct groups are examined.</td>
<td></td>
</tr>
</tbody>
</table>

Causal inferences can be drawn

**Examples:**

- **Upper left:** A random sample is selected from the population of eighth grade students in Wisconsin. The children are randomly divided into two groups. One group reads a passage on a subject and the other watches a video on it. The response is the score on an exam.

- **Upper right:** A random sample of WI students and a random sample of WY students watch a video on a subject. The response is the score on an exam.

- **Lower left:** A collection of eighth grade students in Wisconsin is available for study. The children are randomly divided into two groups. One group reads a passage on a subject and the other watches a video on it. The response is the score on an exam.

- **Lower right:** Collections of eighth grade students in Wisconsin and Wyoming are available for study. The children read the passage. The response is the score on an exam.

This table can be used for units that are trials too. Clyde wants to investigate the relationship between location and outcome of his three point basket attempts. The locations are front and left corner.

- **Upper left:** Clyde randomly assigns 50 shots to each location and assumes that his shots are Bernoulli trials.

- **Upper right:** Clyde shoots 50 shots from the front on Tuesday and 50 shots from the left corner on Wednesday. He assumes that his shots are Bernoulli trials.

- **Lower left:** Clyde randomly assigns 50 shots to each location but makes no assumption about Bernoulli trials.
• Lower right: Clyde shoots 50 shots from the front on Tuesday and 50 shots from the left corner on Wednesday. He does not assume that his shots are Bernoulli trials.

In the top row, use inference formulas for independent random samples. For lower left, you should use randomization-based inference, but nearly everybody “cheats” and uses the inference formulas for independent random samples. For lower right you should not do inference, but people often cheat.

Paired data. How? By matching or subdividing units.

1. Subdividing.
   • Units are teenagers. Study factor is acne medicine with two levels. Response is number (surface area?) of zits. Each unit is divided into a left face and right face. Medicine is assigned to side of face by randomization (toss a fair coin for each child).
   • Units are teenagers. Study factor is time of school year with levels “day before finals start” and “day after finals end.” The response is systolic blood pressure. Notice that randomization is not possible; the child necessarily experiences ‘before’ before ‘after.’
   • The earlier table applies.

2. Matching.
   • Units are 40 teenagers. Each student is given an IQ test. The IQ scores are used to form pairs (the two highest form a pair, then the next two highest, and so on). Within each pair, one student is randomly assigned to each of two learning methods. The response is the score on an exam. The first column of the earlier table applies.
   • Units are 20 teenagers in WI and 20 in WY. Each student is given an IQ test. The IQ scores are used to form pairs (the highest from each state form a pair, then the next highest, and so on). Each child is assigned to the same learning method and the response is the score on an exam. Paired inference is invalid! Do not form pairs by matching unless you can randomize! This fact is largely misunderstood (see books on psychology stats).