Notes for 5.2 (in order of appearance in the text)

Observational study pg 266
Observes individuals and measures variables of interest but does not attempt to influence responses

Experiment pg 266
Deliberately imposes some treatment on individuals in order to observe their response.
If we want to understand cause and effect, experiments are the only source of convincing data.
The distinction between explanatory and response variables is essential if we are to see how one variable changes in response to changes in another variable.
Design of an experiment: *Describes the response variable(s)*
The factors (explanatory variables)
The specific treatments

Experimental units
Individuals on which the experiment is done. (When these are people we call them subjects.)

Treatment
A specific experimental condition applied to the units

Factors
The explanatory variables in an experiment. There may be more than one factor—some experiments study the joint effects of several factors. Each treatment is formed by combining a specific value or level of each factor.

Placebo pg 269
A treatment that has no physical effect—a dummy treatment.

Placebo effect
Some subjects respond favorably to any treatment even a placebo.

Control group pg 270
The group of subjects who receive the placebo treatment (or no treatment). This controls the effects of lurking variables on the outcome.

Randomization pg 271
The use of chance to divide experimental units into treatment groups. Randomization means that each member of the population must have an equal chance of being chosen. (Random does not mean "haphazard." Surveying the first 10 people you meet in the hallway is not a random sample.)

Completely randomized design pg 273
All experimental units are allocated at random among all the treatments. Matching if trying to pair subjects or groups in the experimental group with similar subjects or groups in the control group. This is difficult (impossible?) to do.

The logic of experimental design: pg 275

- **Randomization produces groups of experimental units that should be similar in all respects before the treatment is applied.**
- **Comparative design ensures that influences other than the experimental treatments operate equally on all groups.**
- **Therefore, differences in the response variable must be due to the effects of the treatments. That is, the treatments not only are associated with the observed differences in the response, but must actually be the cause of those differences.**

We know that an association (either positive or negative), even if it is a strong association, does not mean that the change in the explanatory variable CAUSES the change in the response variable. A strong association from a well-designed experiment DOES imply causation.

Statistically significant pg276
When we observe differences or change in the experimental units, some of that may be due to just random chance. We want to know whether the observed difference is too large to have resulted from chance alone. An observed effect too large to attribute plausibly t o chance alone. Experiments with many subjects are better able to detect differences among the effects of the treatments than similar experiments with fewer subjects.

Principles of experimental design:

1st Control *of the effects of lurking variables on the response (most simply by comparing several treatments).*

Principles of experimental design:

2nd Randomization-- *the use of impersonal chance to assign subjects top treatments.*

Principles of experimental design:

3rd Replication *of the experiment on many subjects to reduce variation in the results.*

Hidden bias pg 277
Happens when subjects are not treated in exactly the same way. Unequal conditions produce bias. Be careful to treat all subjects alike.

Double-blind *experiment* pg 278
Neither the subjects nor the people who interact with them know which treatment a subject is receiving. (Obviously, the person in charge of the experiment must know who is receiving which treatment!)

Matched pairs design
A type of block design where we compare two treatments. Each block consists of just two units as closely matched as possible. The individual unit is assigned to a treatment randomly (flip a coin, odd/even number, etc.)

Lack of realism

Subjects or treatments or settings of an experiment may not realistically duplicate the conditions we really want to study. The true scope of a new finding must usually be explored by a number of experiments in various settings.

Block pg 281

A group of experimental units or subjects that are similar in ways that are expected to affect the response to the treatments.

Blocking is used to control the factors you CAN see; randomization helps balance out the ones you CANNOT see. Randomization reduces bias and blocking tends to reduce unwanted variation.

Block design

The random assignment of units to treatments is carried out separately within each block.

Your concern is that groups are as alike as possible except for the treatment. Smaller blocks give less variability. In blocked design, the subjects of the experiment are divided according to some characteristic that is believed to create variation in the response. The reduction of variability because of blocking is greatest when there are as many blocks as possible. The ultimate goal is to have the number of subjects in each block equal to the number of treatments in the experiment. The purpose of randomization is to reduce the variation due to unanticipated factors. Randomization will make it equally likely for a treatment group to get an individual who will have a particularly good or a particularly bad response, so that at least on average these good and bad responses are evenly distributed between two treatments. Thus a completely randomized design, which relies only on randomization to form the treatment groups, may have more variation than a randomized block design

Stratification in sampling is similar to blocking in experiments. When you have groups in the population that are very different from other groups but the units within the group are similar to each other, it is good to stratify--to encourage our sample to resemble our population. If you have strata that are different in a way important to your study, choosing a stratified sample will have less variation than a simple random sample of the same size. Stratifying can reduce the number of people you have to contact.