Statistics Review

• Most research designs intended to provide evidence that one variable caused another
  – In a true experiment, does mean score in one experimental group differ from another group?

• “Statistical significance” assesses the probability that results could be due to chance rather than the hypothesized cause
  – E.g., could difference between 2 means be as large as it is by chance? (cf. Ruth Rosenholtz lecture about t-test)
  – Could the outcomes be as large as it is by chance alone?
Internal Validity

• **Internal validity**: cause-effect relationship between the independent and dependent variable
• **Watch out for “confounding effect”** (the correlate of “controls”)
  
  • Example of confounded variables which may or not have an impact on the data:
  • Time of the day, year (group 1 testing morning, group 2 afternoon)
  • Gender, Age, education
  • Familiarity with the task
  • Subject's mood - (did you run one condition right after a holiday?)
  • Subjects' hobbies - video game players in one condition, gardeners in another?
  • Different experimental machines?
  • Familiarity with experimenter? (Were all your friends in one task?)

Construct Validity

• Construct validity: extend to which the results support the theory behind the research:
• **Ask the question**: Would another theory predict the same experimental results?
• You can never ensure construct validity, but you can plan your research so that it is more plausible
  
  • Examples (p. 172)
  • In internal validity, you strive to rule out alternative variables
  • In construct validity, you rule out other possible explanations
  • In most cases, you have to run another experiment to rule out threats of validity
  • For project 1 presentation, one of the exercise in class will be to think about internal and construct validity while your colleagues will present.
Statistical Validity

- Extend to which data are shown to be the result of cause-effect relationship rather than accident (chance alone)
- Cf. Ruth Rosenholtz lecture on t-test
- Did you have enough subjects? Enough stimuli? Was the variance between your groups comparable?
- Threats to validity: textbook pages 173-179

The t Test

\[ t = \frac{\text{Difference between groups (means)}}{\text{Normal variability within group(s)}} \]

- If \( t \) is large, the difference between groups is much bigger than the normal variability within groups.
  - Therefore, two groups are significantly different from each other

- If \( t \) is small, the difference between groups is much smaller than the normal variability within groups.
  - Therefore, two groups are not significantly different from each other
Comparing 2 Means

- **Null hypothesis** ($H_0$): Population means are equal. Any differences between sample means are due to chance (random error).
- A good example: The collected data for CogLab 1: spatial cuing.
- **Research hypothesis** ($H_1$): Population means are not equal.
- **T-test**: Test statistic associated with a probability of obtaining sample means that differ by observed amount if population means were equal.

Type I and Type II Errors

- **Type I error**: incorrectly rejecting the null hypothesis when it is in fact correct.
- **Type II error**: incorrectly accepting the null hypothesis when it is in fact false.
- Alpha is the p-value at which we decide to reject the null hypothesis
  - As alpha gets larger, the probability of a type I error increases and the probability of a type II error decreases.
- As $n$ (sample size) increases, probability of a type II error decreases.