Practice with Effect Sizes

20 March 2020

Modern Research Methods

Logistics

- As a reminder, this session will be audio/video recorded for educational use by other students in this course.
- Office hours – same time, but over Zoom
- Must **sign-up** using spreadsheet
  - If you're unable to make those times, let us know and we will do our best to accommodate you
- No assignment this week

INSTRUCTOR

Dr. Molly Lewis

mollylewis@cmu.edu

ZOOM OFFICE

Office Hours: W 4:30-6:30pm

Signup: **signup sheet**

TA

Jaeah Kim

jaeahk@andrew.cmu.edu

ZOOM OFFICE

Office Hours: M 1:00-3:00pm

Signup: **signup sheet**

COURSE

MW (lecture); F (lab)

10:30-11:20am

Lecture/Lab: **ZOOM CLASSROOM**

Effect sizes

- P-values give you a yes/no answer – is the difference significant or not?
- Confidence intervals give you a range of plausible values for the means in the conditions.
- Effect sizes – how big is the effect and what direction is it in?

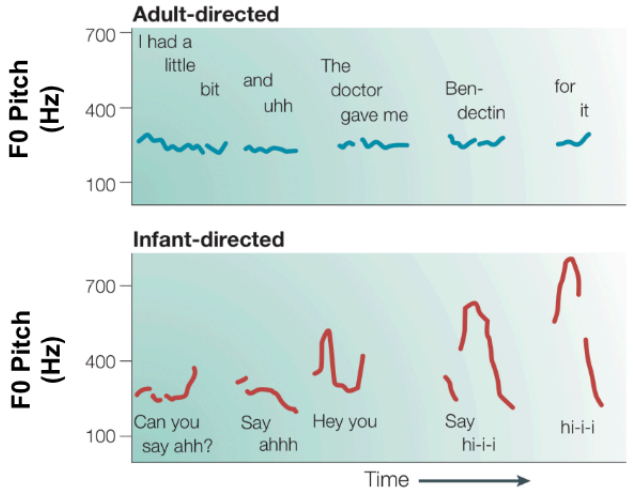
- “Statistical significance is the least interesting thing about the results. You should describe the results in terms of measures of magnitude – not just, does a treatment affect people, but how much does it affect them.” - Gene Glass

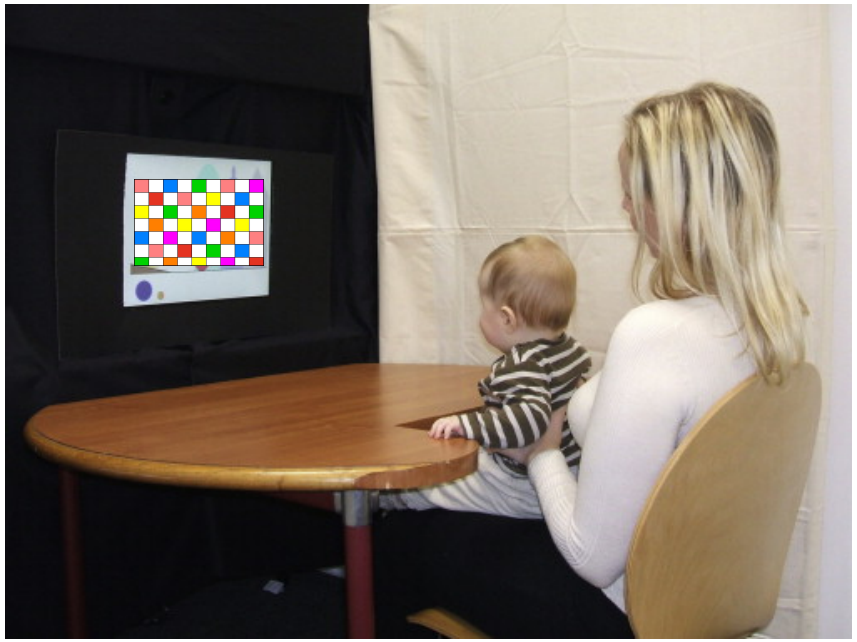
How to quantify the effect?

- Depends on your design and what kind of variables you have.

Cooper & Aslin (1990)

Do infants prefer IDS to ADS?





Source: Moll & Tomasello, 2010

Dependent measure: Looking time to checkerboard

Independent variable: ADS vs. IDS played in pairs of trials within subjects

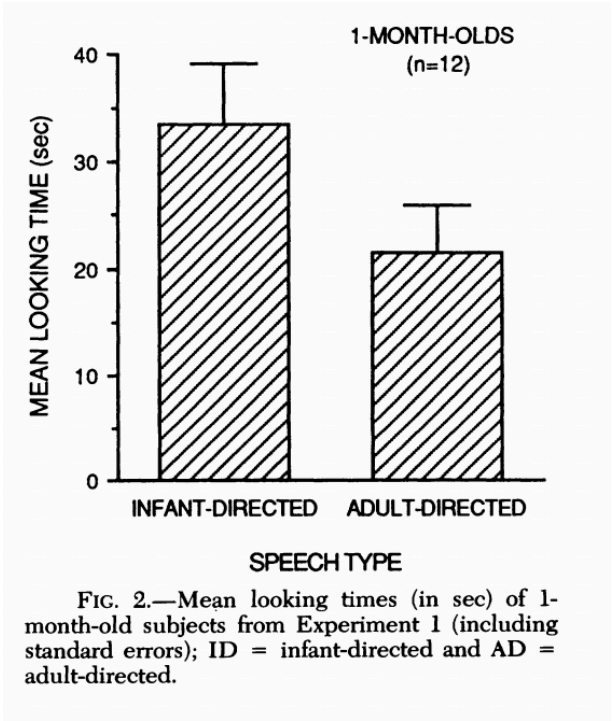
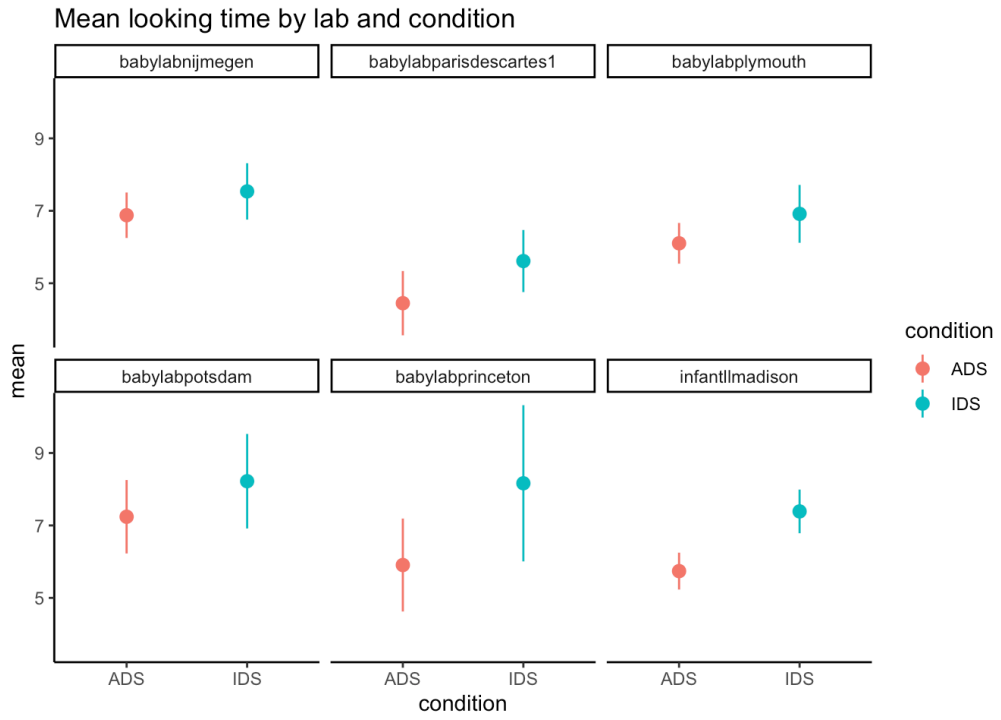


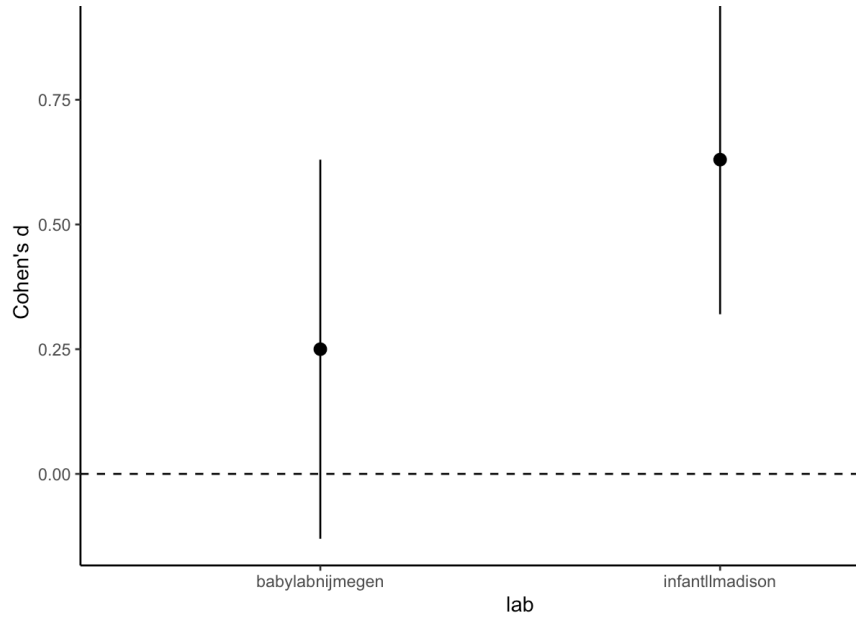
FIG. 2.—Mean looking times (in sec) of 1-month-old subjects from Experiment 1 (including standard errors); ID = infant-directed and AD = adult-directed.

How to quantify the effect?

- Depends on your design and what kind of variables you have.
- What is the design here?



Midterm 9b



Cohen's *d*

Standardized measure of the size of an effect when you have a categorical IV and a continuous DV.

Cohen's *d*:

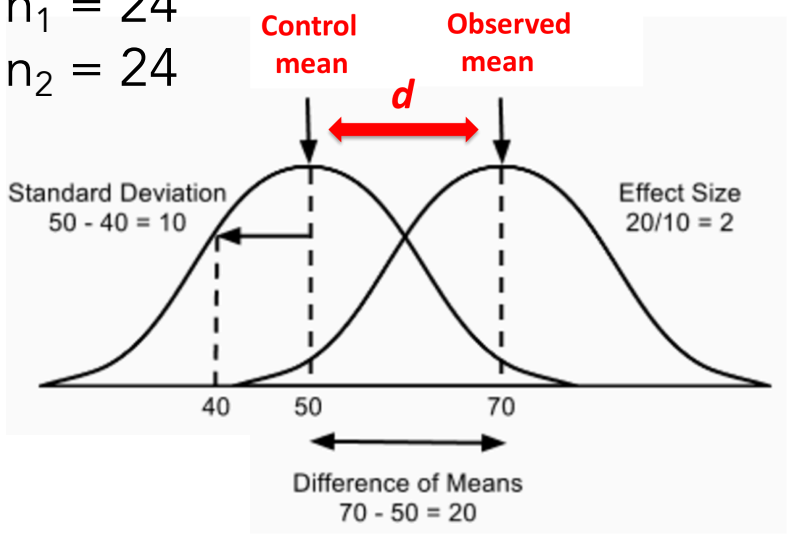
Effect Size = $\frac{\text{diff. between means}}{\text{standard dev.}}$

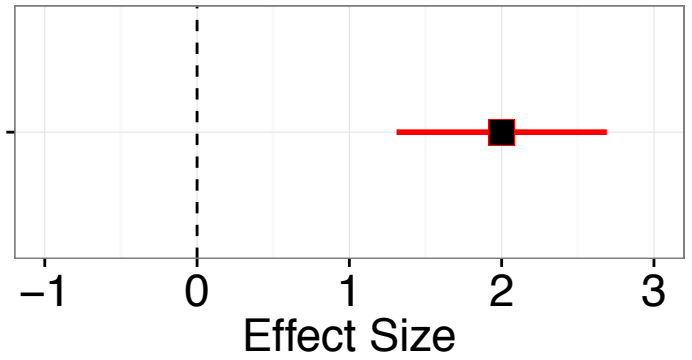
$$d = \frac{M_{group1} - M_{group2}}{SD_{pooled}}$$

$$SD_{pooled} = \sqrt{(SD^2_{group1} + SD^2_{group2})/2}$$

Cohen's *d* confidence interval

$n_1 = 24$
$n_2 = 24$





$$
\begin{aligned}
var_d &= \frac{n_1 + n_2}{n_1 * n_2} + \frac{d^2}{2(n_1 + n_2)} \\
&= \frac{24 + 24}{24 * 24} + \frac{2^2}{2(24 + 24)} \\
&= .125
\end{aligned}
$$

$$
\begin{aligned}
CI(d) &= Est(d) \pm z_{(\alpha/2)} * \sqrt{var(d)} \\
&= 2 \pm 1.96 * .35 \\
&= 2 \pm .69
\end{aligned}
$$

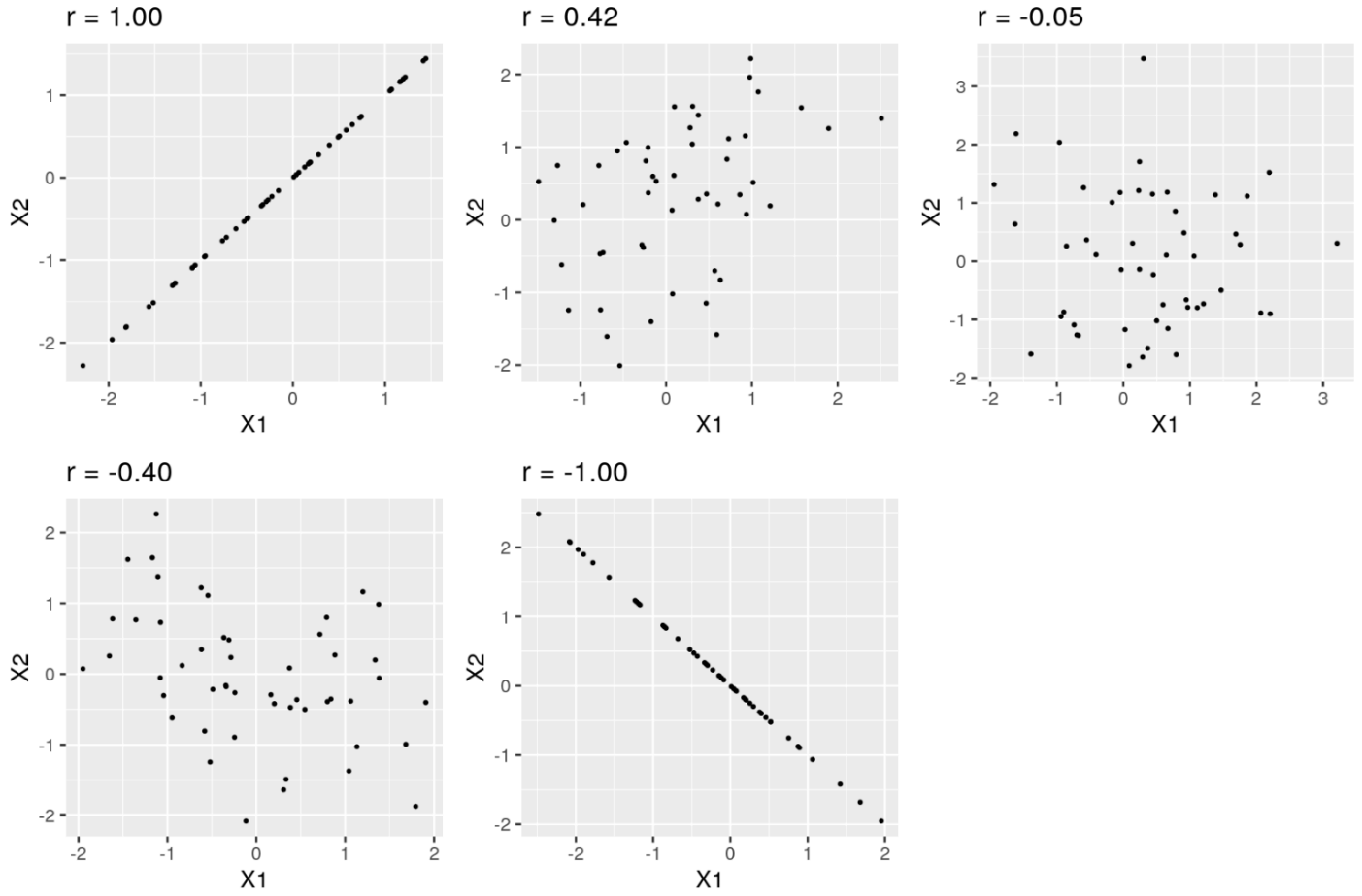
Pearson's *r*

Correlation coefficient

Standardized measure of the size of an effect when you have a continuous IV and a continuous DV.

Ranges from -1 to 1

Don't have to calculate it (typically reported in paper)



Effect size measures

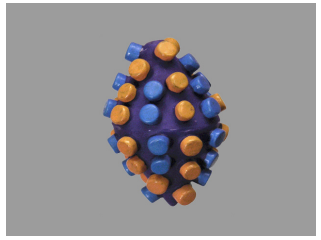
- For any statistical test you conduct, can compute effect size (in principle)
- ES depends on design
- Can convert between ES metrics

Let’s calculate an effect size for performance in an experiment testing mutual exclusivity.

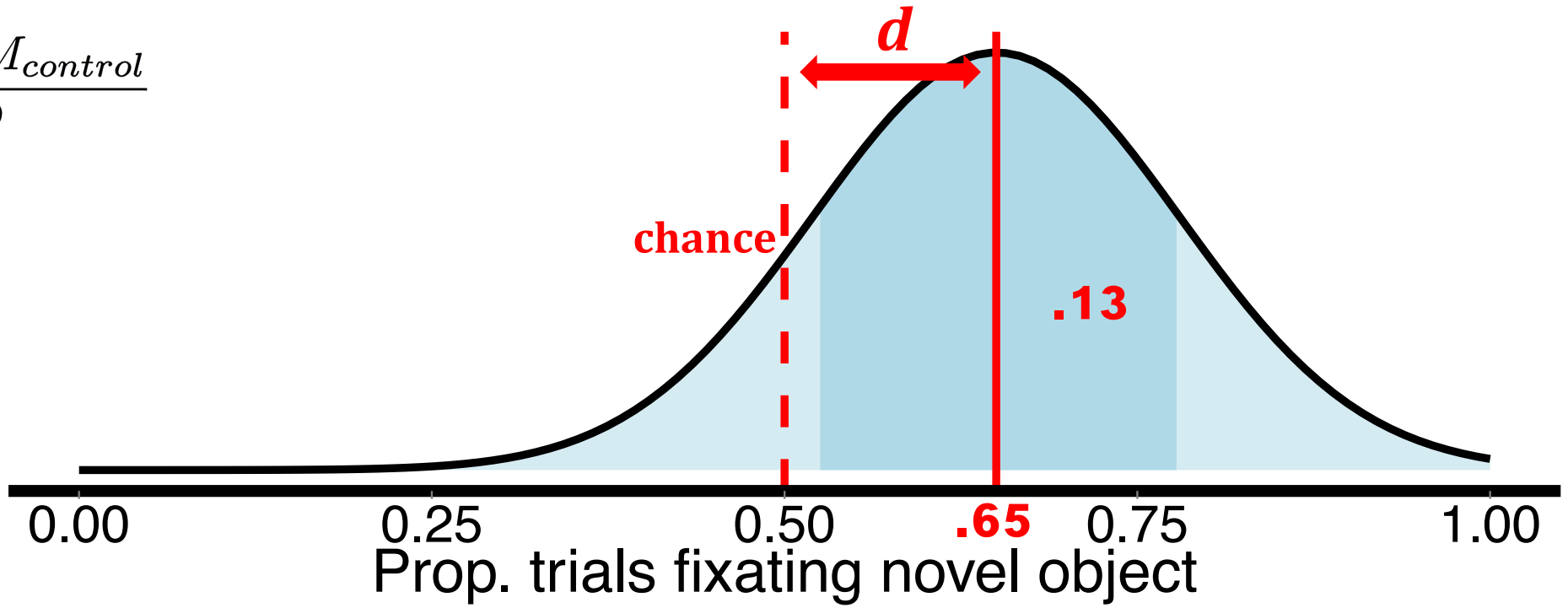
- What is mutual exclusivity?
- What kind of design is it?
- What is the appropriate effect size measure.

Calculating an effect size from a paper

Where's the dofa?



$$d = \frac{M_{exp} - M_{control}}{SD}$$



Coding an effect size from a paper

- Requires you to understand what part of the paper has data relevant to the experiment you care about
- Here, we're interested in coding effect sizes for experiments that test a version of "Mutual Exclusivity"
- May not report the data in the way you need (e.g. number of correct choices rather than proportion)
- Sometimes they may not even report the means at all! (just the statistical test)

Cognition 126 (2013) 39–53

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months

Ricardo A.H. Bion [a,*], Arielle Borovsky [a,b], Anne Fernald [a]

[a] *Stanford University – Department of Psychology, 450 Serra Mall, Stanford, CA 94305, United States*
[b] *University of California, San Diego – Center for Research in Language, 9500 Gilman Drive MC 0526, La Jolla, CA 92093-0526, United States*

Your job #1

- In your breakout groups, calculate an effect size for each of the three age groups (18, 24, 32 months) in Experiment 2 of Bion et al. (2013)
- Note that another name for "mutual exclusivity" is "disambiguation"
- You'll have to dig into the paper a little bit to find the relevant numbers.
- If you have time, you can also calculate the confidence intervals on the effect sizes.

Paper: https://bit.ly/2QvwHq1 (also linked on website)

J. Child Lang. **28** (2001), 787–804. © 2001 Cambridge University Press
DOI: 10.1017/S0305000901004858 Printed in the United Kingdom

NOTE

By any other name: when will preschoolers produce several labels for a referent?*

GEDEON O. DEÁK AND LOULEE YEN
Vanderbilt University

JEREMY PETTIT
David Lipscomb University

(*Received* 23 *February* 2000. *Revised* 8 *December* 2000)

Your job #2

- In your breakout groups, calculate an effect size for the two age groups in Deak, et al. (2001) in Experiment 1.
- If you have time, you can also calculate the confidence intervals on the effect sizes.

Paper: https://bit.ly/3a8GmuH (also linked on website)