Within-subjects designs

Quiz week 5 available now
Due Fri 11:59pm

Questions, concerns, curiosities?
Come to office hours! ☺
Thu 1:30-3:30

- Last week: *between-groups* designs
  - Compare group 1 to group 2 (to group 3…)

- This week: *within-subjects* designs
  - Compare each person to themselves

Within-subjects design: An elegant way to reduce nuisance variability

- Between-groups design
  - Each participant contributes data in **one** condition
  - Different participants perform differently: this adds **noise**
  - Which can make it hard to see your effects
  - To get 20 data points per condition, need 40 ppts (or 60, or 80…)
- Within-subjects design
  - Each participant contributes data in **all** conditions
  - Compare each participant TO THEMSELF in other conditions
  - Removes (at least some) participant-related noise
  - To get 20 data points per condition, need 20 ppts (still 20, still 20…)

Some specific types of within-subjects designs

- Pretest-posttest design*
  - Two conditions
  - Each participant gets both a pretest and a posttest
  - Usually some intervention occurs in between
- Repeated-measures design
  - Two or more conditions
  - Ex: Random dot motion coherence: 10%, 20%, 30%, 40%
- Longitudinal design**
  - Multiple measurements at different time points
  - Can be really informative!
  - Also can take a really really long time (maybe you want to get results sooner than that)

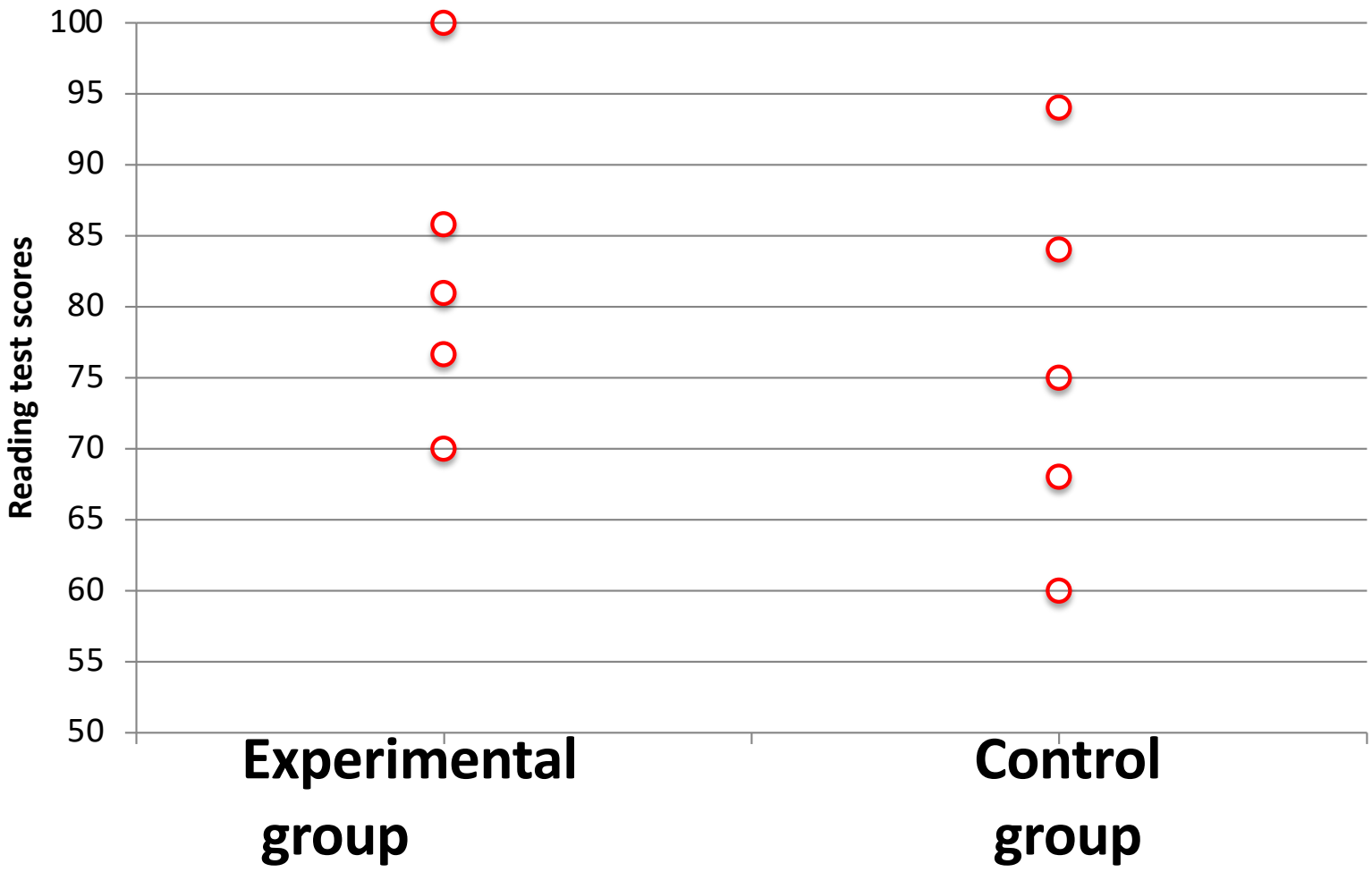
Benefits of within-subjects design

- Need fewer participants!
  - Half as many; 1/3 as many; 1/4 as many…
  - Especially good if each condition is short
- Lower your noise (nuisance variability)!

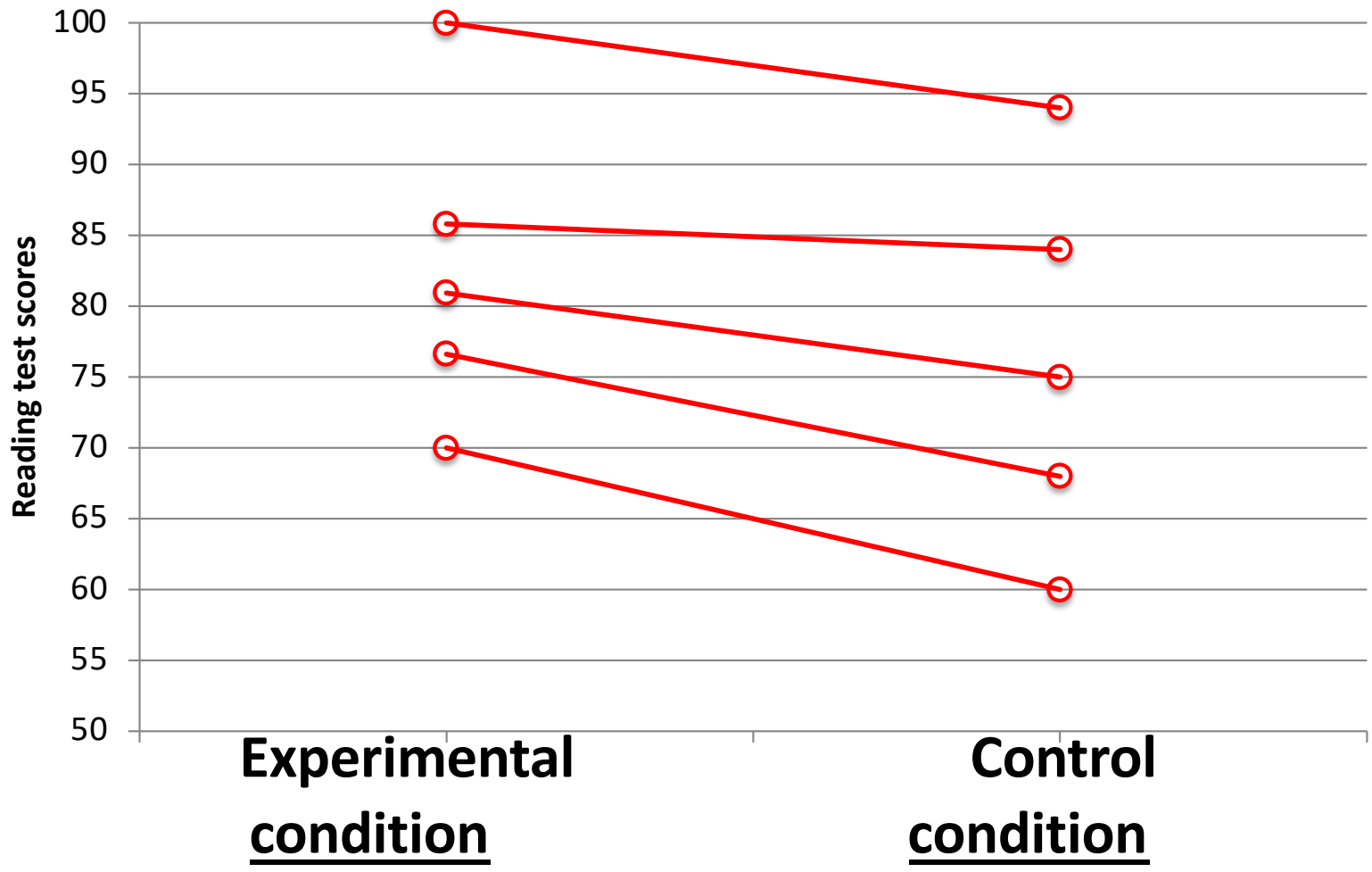
  = a more powerful test

$$\frac{\textbf{Variability between conditions}}{\textbf{Variability within conditions}}$$

Remember these data?



Remember these data?



This sounds great. Is there a catch?

- Not necessarily a catch. Sometimes it's just good experimental design.
- **Demand characteristics**
  - If you [participant] get all of the conditions, you're more likely to figure out what the manipulation is ("unblind" yourself)
- **Carryover effects**
  - If you do Condition A first, it might affect how you do on Condition B
  - **Fatigue effect**: worse on B because tired, bored
  - **Practice effect**: *better* on B because "warmed up," improved at task, learned to do task
- History effect
- Maturation effect
- Testing effect

Other concerns

- **History effect**: Something occurs during experiment (for example, between pretest and posttest; or partway thru longitudinal)
  - Midterms (+ anxiety, + sleep loss)
  - Major world event (+ anxiety, + distraction)
  - University wins/loses major game
  - Sun God Festival
  - History is *confounded* with the manipulation

Other concerns

- **Maturation effect**: Maybe improvement (or worsening) is due to passage of time, biological maturation—not your intervention
- Children's vocabularies improve with age
  - Maybe it wasn't your reading skills intervention
- Hearing gets worse due to biological aging, general noise exposure
  - Maybe it wasn't the particular blood thinner you gave to people
- Might want to do pretest-posttest on control group—measure rate of change *without* intervention

Other concerns

- **Testing effect**: if someone gets the same test twice (over a long time span), performance the second time is affected by first time
  - Know what to expect
  - Know particular test items
  - Know how to pace yourself
  - Might try to get particular result (e.g. might try to show less bias on Implicit Association Test)

Order effects

What you don’t want:

- Condition A on Day 1.
- Condition B on Day 2.
- Condition B shows better performance!

- Condition A run first; Condition B 30 min later.
- Condition A shows better performance!

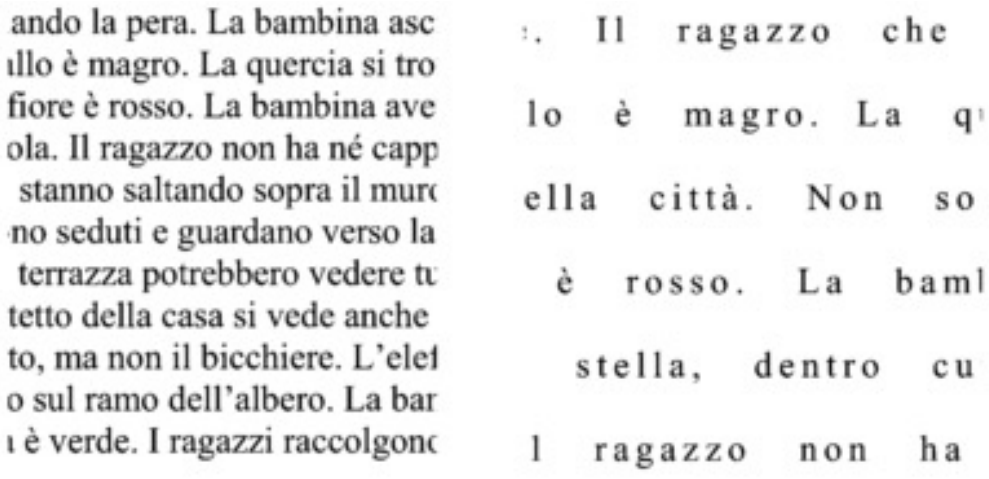
- What’s wrong with this picture?

Do children understand speech better when it's in their own accent?
(These children speak London English)

The story spoken in the London accent was presented first (with the accompanying pictures and text), followed by the 20 single words in the London accent (alternately List A or List B). The experimental condition, consisting of the second short story and the twenty words in the Glaswegian accent (alternately, List B or List A) then followed.

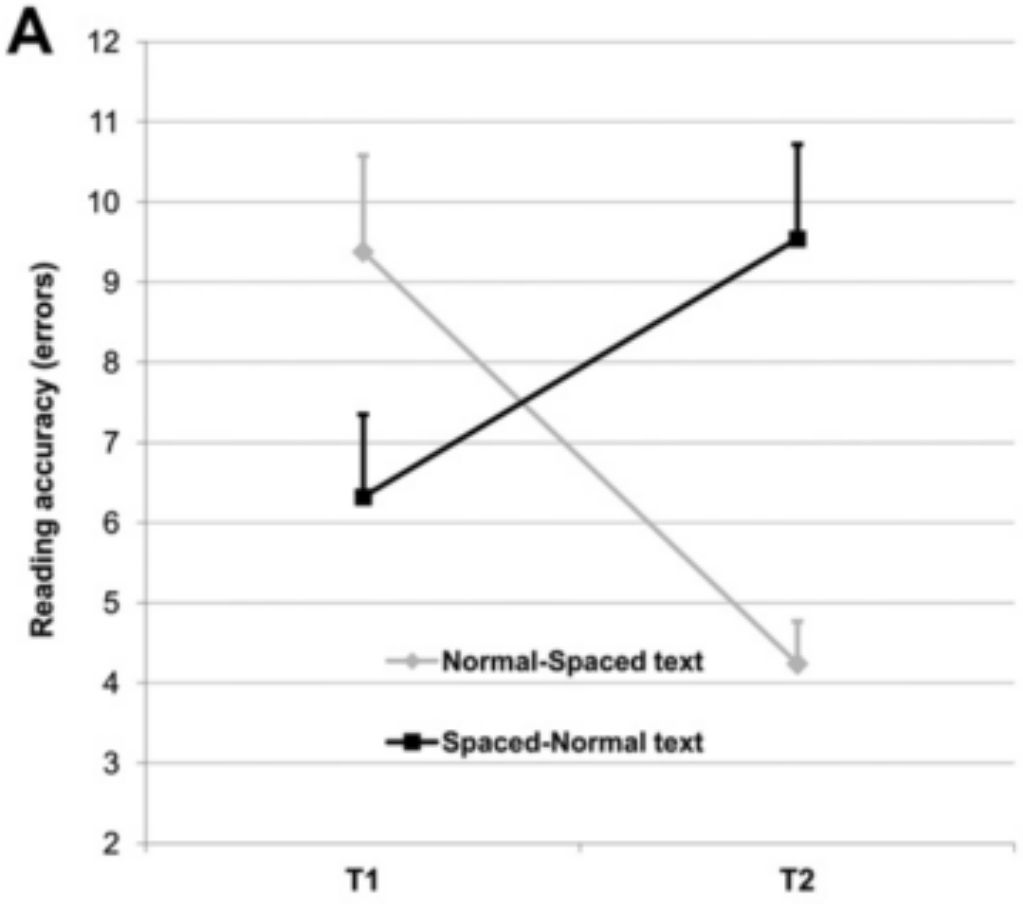
TABLE 2. *Means and standard deviations of factors age and accent*

	London accent		Glaswegian accent	
	Mean	S.D.	Mean	S.D.
4-year-olds	16·38	2·22	8·67	2·46
7-year-olds	18·83	0·82	14·21	1·91



regular spaced

Hypothesis: children with dyslexia will read better with widely-spaced text than regular text. (not just anybody)
Kids in each country were assigned to normal-spaced or spaced-normal.
Within or between?
Complete [everybody gets all orders] or incomplete [everybody gets all conditions but NOT all orders]?
Does spaced text actually help?



Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., … Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences, 109*(28), 11455–9. doi:10.1073/pnas.1205566109

This week

- Within-subjects designs

- Finish Week 5 quiz by 11:59pm Friday

Order effects and counterbalancing

- Problem: order effects happen in within-subjects designs.
- One solution: run a between-subjects design
- Another solution: **counterbalance** the order of the conditions

Counterbalancing

- There are a number of ways to do this.
- There are a lot of things you can counterbalance.
- As with between-subjects designs, counterbalancing can get pretty complex pretty quickly—but it’s worth it!

Counterbalancing

Complete within-subjects designs

Every subject gets all possible orders of conditions

- **ABBA design** (only two conditions, A and B)
  - As many times in a row as needed (ABBAABBAABBA)
  - Controls for practice effects—*if* effects linear/slow
  - Doesn't control for rapid effect of practice (~1 trial)
- **Block randomization** (3+ conditions)
  - *Block*: a single presentation of all *n* conditions
  - For 3 conditions, 6 (3!) possible: ABC ACB BAC BCA CAB CBA
  - Order these 6 blocks randomly (possibly multiple times)
  - Great, but what if 4 conditions? 5? 6? (OK if REALLY short)

Counterbalancing

Complete within-subjects designs

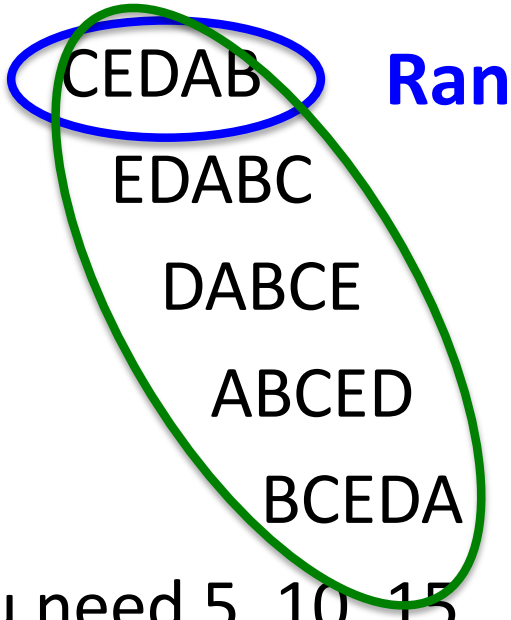
Every subject gets all possible orders of conditions

***Incomplete* within-subjects designs**

Every subject gets every condition, but *not* all possible orders

Incomplete within-subjects designs

- Every subject gets every condition, but *not* all possible orders
- **Random order** with **rotation**
  - Ppt 1: CEDAB
  - Ppt 2: EDABC
  - Ppt 3: DABCE
  - Ppt 4: ABCED
  - Ppt 5: BCEDA



Random order

Rotation

  - So you need 5, 10, 15, … participants for this design
  - (N of participants has to be multiple of # of conditions)

Incomplete within-subjects designs

- Every subject gets every condition, but *not* all possible orders
- **Latin-square design**
  - Each condition occurs in each ordinal position equally often
  - Each condition occurs after each other one equally often

	1st	2nd	3rd	4th
Version 1	A	B	C	D
Version 2	B	D	A	C
Version 3	C	A	D	B
Version 4	D	C	B	A

  - Even #'s: N conditions
  - Odd #'s: 2*N conditions
  - Participants: multiple of # of rows

Cog Sci counterbalancing: additional considerations

- Often, very 'noisy' dependent variables so you have many, many trials (=test items)
- Do you *counterbalance* or *randomly intermix* conditions?
- Visual search task:
  - Easy Easy Easy Easy Easy Easy Easy Easy Hard Hard Hard Hard Hard Hard Hard Hard
  - OR..
  - Easy Hard Easy Easy Hard Hard Easy Hard Hard Easy Hard Easy Easy Hard Hard Easy
- Intermixing basically cancels out order effects
- Counterbalance if:
  - It will be confusing to participants to intermix
  - Few items, or each item is really long

Limits of counterbalancing

- You can’t un-take a drug or un-have a surgery
- You can’t unlearn a memory technique
- You can’t have people un-see a website, or re-donate to public TV/NPR*