Factorial ANOVA

Psychology 3256

Made up data..

- Say you have collected data on the effects of Retention Interval on memory
- So, you do the ANOVA and conclude that RI affects memory

	5 min	1 hr	24 hr
% corr	90	70	60

Made up data 2..

- What about Levels of Processing?
- So, you do the ANOVA and conclude that LOP affects memory

	Low	Med	High
% corr	70	80	90

Hmmm

- What level of RI should you have done your LOP experiment at?
- For that matter, what level of LOP should you have done the RI experiment at?

Combine ‘em

	5 min	1 hr	24 hr
Low	G1	G2	G3
Med	G4	G5	G6
High	G7	G8	G9

Here comes the model

- $x = \mu + \alpha + \beta + \alpha\beta + \varepsilon$

any score

grand mean

A effect

B effect

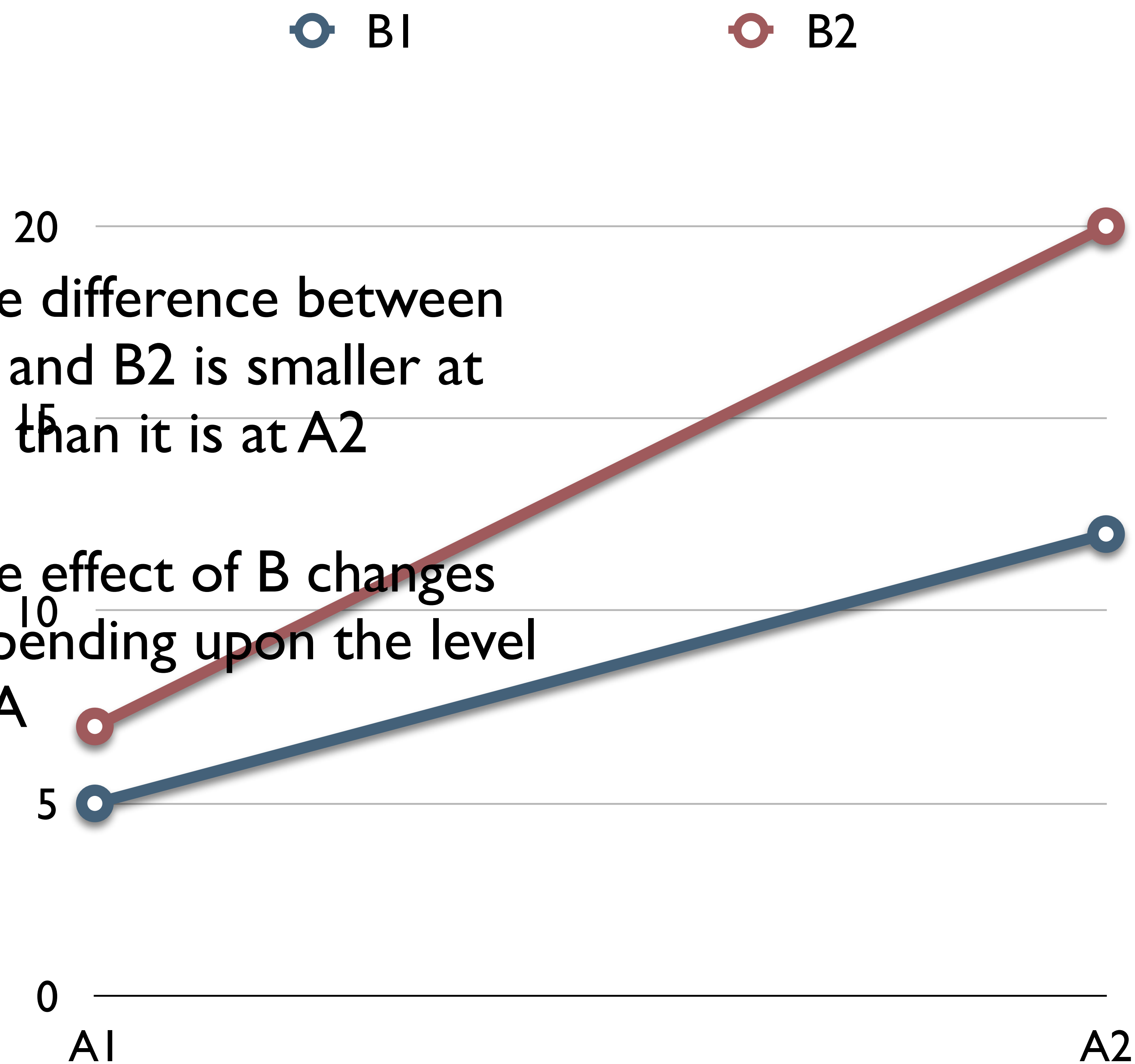
AB Interaction

error

A bit more explanation

- So, not only can we look at A and B, we also look at how A and B act together, how they interact
- Sort of a whole is more than the sum of its parts thing
- The effect of 1 variable changes depending upon the level of some second variable

picture = 1000(words)



- The difference between B1 and B2 is smaller at A2 than it is at A2
- The effect of B changes depending upon the level of A

Meanwhile, back at the structural model...

- $x = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
- Assumptions (model)
- $\Sigma\alpha_i = 0$
- $\Sigma\beta_j = 0$
- $\Sigma\alpha_i\beta = 0$
- ε NID$(0,\sigma^2)$

Assumptions for F

- Homogeneity of variance
- random samples
- normal populations

Numerical example

- $x = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
- take out the grand mean
- $(9+7+3+1)/4=20/4=5$

	A1	A2
B1	9	7
B2	3	1

Subtract 5 out of each cell

- So with the grand mean removed we can go on to the effects of A and B
- B1 6/2 = 3
- B2 -6/2 = -3

	A1	A2	sum
B1	4	2	6
B2	-2	-4	-6

Subtract 3 from the B1s and -3 from the B2s

- A1 1
- A2 -1
- Now do the same as before but for the As

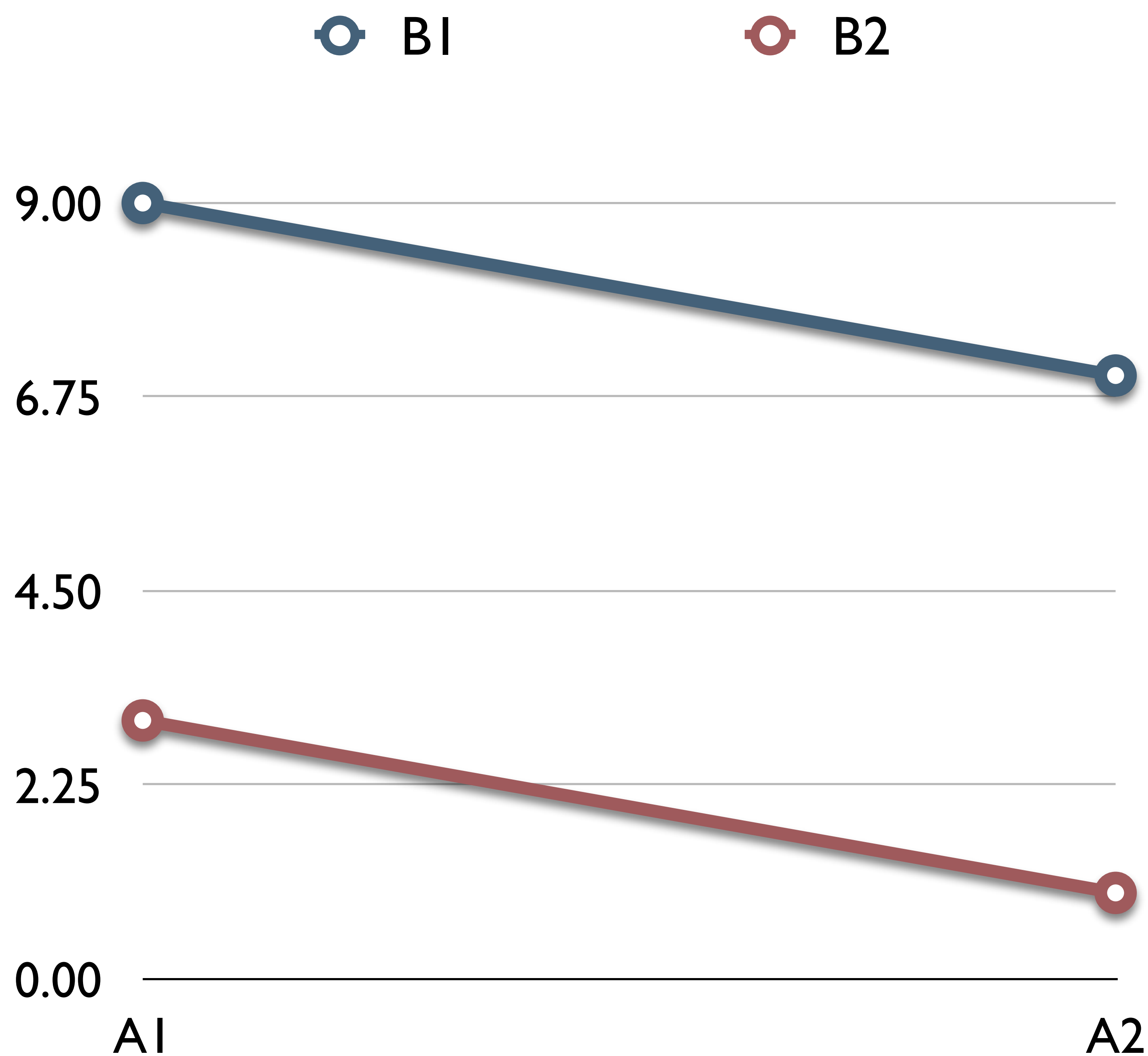
	A1	A2
B1	1	-1
B2	1	-1
sum	2	-2

And we are left with

- Well with nothing, so there is no interaction, just a mean and effects of A and B

	A1	A2
B1	0	0
B2	0	0

Graph it



Another example

- OK, first get the grand mean (20+0-10+2)/4=3
- Now, we remove the grand mean

	A1	A2
B1	20	0
B2	-10	2

Out comes the grand mean

- The A effect is 2 for A1 and -2 for A2
- Note how they always sum to 0

	A1	A2
B1	17	-3
B2	-13	-1
sum	4	-4

Take out 2 from A1 and -2 from A2

- OK, notice how the cells sum to 0 (the grand mean is gone) AND so do the columns, as we have taken out A
- So for B we have B1 7 B2 -7
- Well take it out

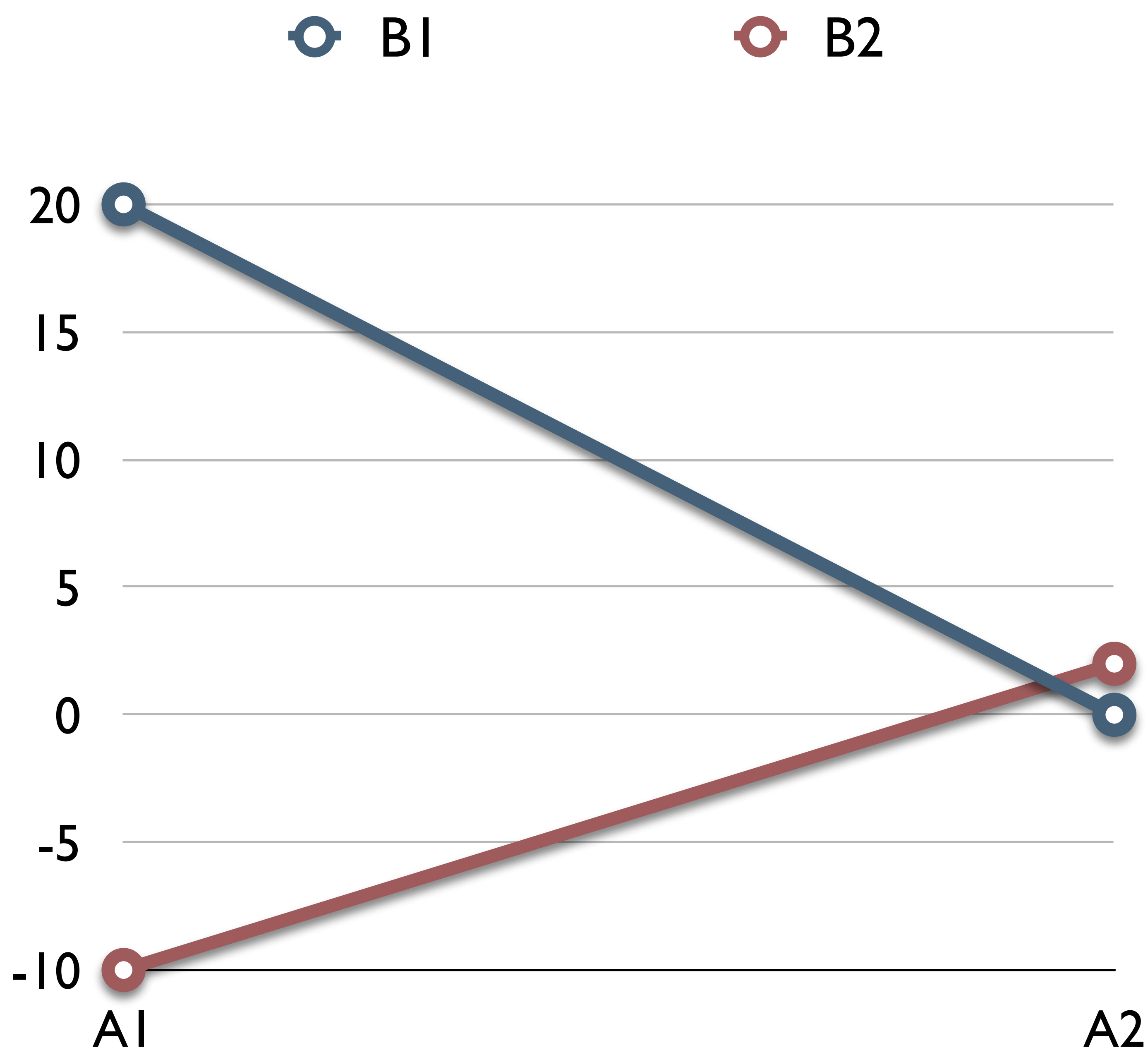
	A1	A2	sum
B1	15	-1	14
B2	-15	1	-14

What is left is the interaction

- So we have an interaction
- Note that the effects sum to 0 in every possible way
- if it was just 0s we would have no interaction

	A1	A2
B1	8	-8
B2	-8	8

Graph it



Interpreting interactions

- Be careful when you are interpreting main effects in the presence of interactions
- Not so bad with an ordinal interaction
- harder with a disordinal interaction, probably impossible really

Partitioning the df and SS

- remember the model
- $x = \mu + \alpha + \beta + \alpha\beta + \varepsilon$
- SSTO = SSA+SSB+SSAB+SSE

Think of it this way

- SSA
  - Squared deviations of column means from grand mean
- SSB
  - Squared deviations of row means from grand mean

Keep thinking….

- SSAB
  - Squared deviations of cell means from what we would expect given row and column means
- SSE
  - Squared deviations of individual scores from cell means

More Precisely...

$$\frac{\sum\left(x-\bar{x}_g\right)^2}{N-1}=nq\frac{\sum\left(\bar{x}_{j.}-\bar{x}_g\right)^2}{a-1}+np\frac{\sum\left(\bar{x}_{.i}-\bar{x}_g\right)^2}{b-1}+n\frac{\sum\sum\left(x-\bar{x}_{j.}-\bar{x}_{.i}+\bar{x}_g\right)^2}{(a-1)(b-1)}+\frac{\sum\sum\sum\left(\bar{x}_{ji}-\bar{x}_g\right)^2}{ab(n-1)}$$

Expected values

- Remember for the simple ANOVA the E(MST) = $\varepsilon + \tau$ and the E(MSE) = ε so we would divide MST by MSE to find out if we had an effect
- Well we have to do the same thing for MSA MSB and MSAB (and of course MSE)

Here you go, as you would expect

- $E(MSA) = \alpha + \varepsilon$
- $E(MSB) = \beta + \varepsilon$
- $E(MSAB) = \alpha\beta + \varepsilon$
- $E(MSE) = \varepsilon$
- so divide them all by MSE to sort of isolate the effect

However….

- Those expected values are only for the case where you are only interested in the particular values of A and B that you have in your experiment, no others!
- This is called a Fixed effect model
- What if we randomly chose the levels?

Random effects model

- E(MSA) = $\alpha + \alpha\beta + \varepsilon$
- E(MSB) = $\beta + \alpha\beta + \varepsilon$
- E(MSAB) = $\alpha\beta + \varepsilon$
- E(MSE) = ε
- So divide MSA and MSB by MSAB and MSAB by MSE

Mixed model, A fixed, B random

- $E(MSA) = \alpha + \alpha\beta + \varepsilon$
- $E(MSB) = \beta + \varepsilon$
- $E(MSAB) = \alpha\beta + \varepsilon$
- $E(MSE) = \varepsilon$
- No, that is not a typo.. and yes it is counterintuitive

So....

- We are assuming with a random effects model that the levels of the random factor are randomly selected and independent of each other
- Really, we are usually doing a random effects or mixed model, sort of...
- Did you really randomly select the levels?

ANOVA summary table

Source of Variation	df	MS	F
A	a-1	SSA/(a-1)	MSA/MSE
B	b-1	SSB(b-1)	MSB/MSE
AB	(a-1)(b-1)	SSAB/ (a-1)(b-1)	MSAB/MSE
Error	ab(n-1)	SSE/ ab(n-1)	
TOTAL	N-1		

FIXED EFFECTS ONLY!

You can make these designs bigger!

	C1	C1	C2	C2
	A1	A2	A1	A2
B1	G1	G2	G5	G6
B2	G3	G4	G7	G8

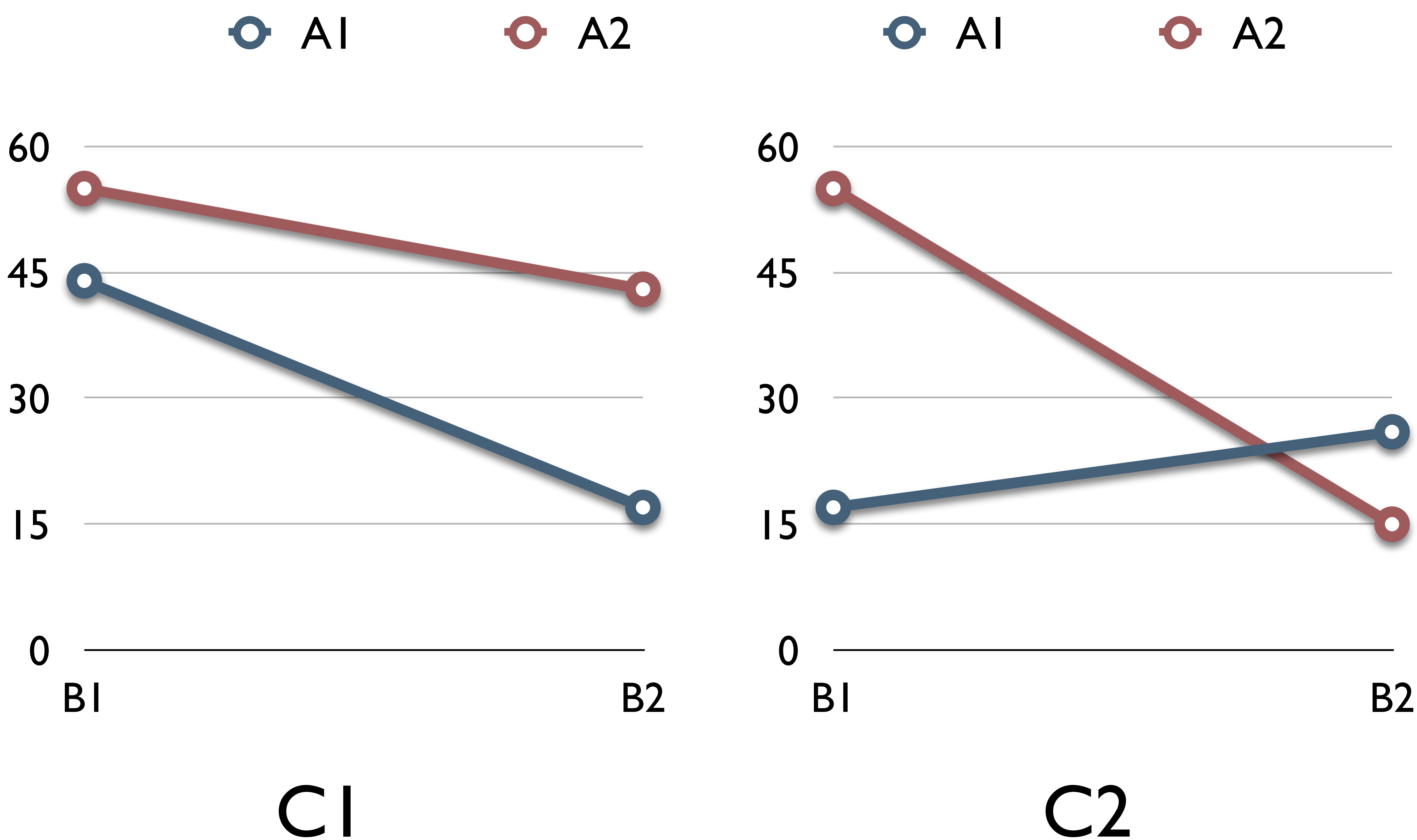
Now….

- Now you have 3 main effects (A B and C)
- 3 two way interactions (AB AC and BC)
- and a 3 way interaction (ABC)
  - when a 2 way interaction changes depending on the level of some third variable

The model now is..

- $x = \mu + \alpha + \beta + \gamma + \alpha\beta + \alpha\gamma + \beta\gamma + \alpha\beta\gamma + \varepsilon$

Looks like this



Advantages of these designs

- We can study interactions
- indeed many of our theories have interactions in them
- relatively simple to interpret once you have done it a few times

The down side...

- Fixed, random or mixed?
- They can get HUGE fast