Repeating the Repeated Measures ANOVA

Using the Mixed Procedure

MAASTRICHT UNIVERSITY

Welcome

This Session

- The Logic of Repeated Measures
- Sphericity
- Repeated Measures ANOVA output
- Repeated Measures ANOVA Limitations
- Repeated Measures with the Mixed Procedure

The Next Sessions

How to build Models and decide on covariance structures

(plus talk about what covariance structures are)

How to add covariates, random intercepts, and random slopes

(Making the Mixed Models truly Mixed)

Adding Planned and Polynomial Contrasts

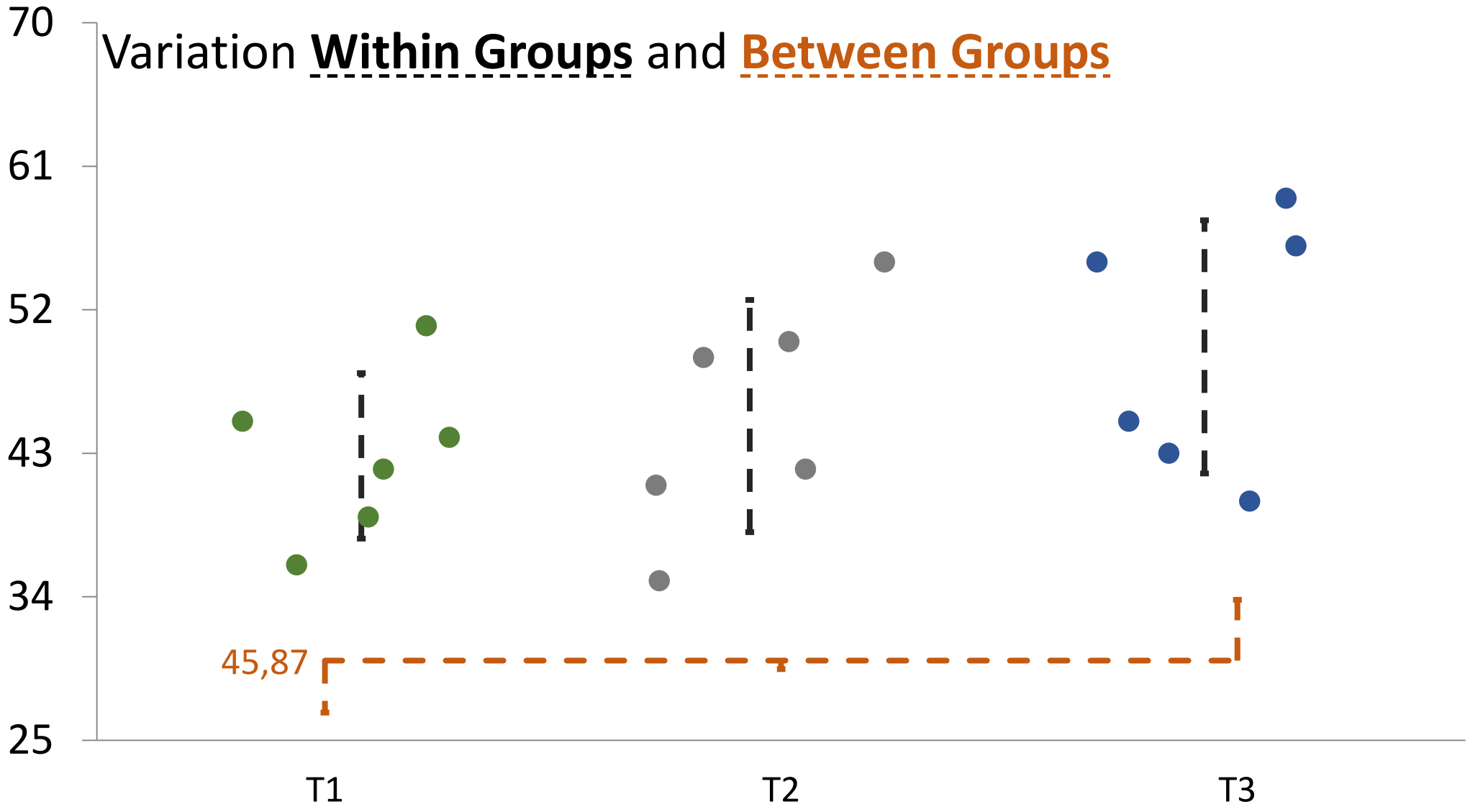
(Thinking ahead like a good data scientist)

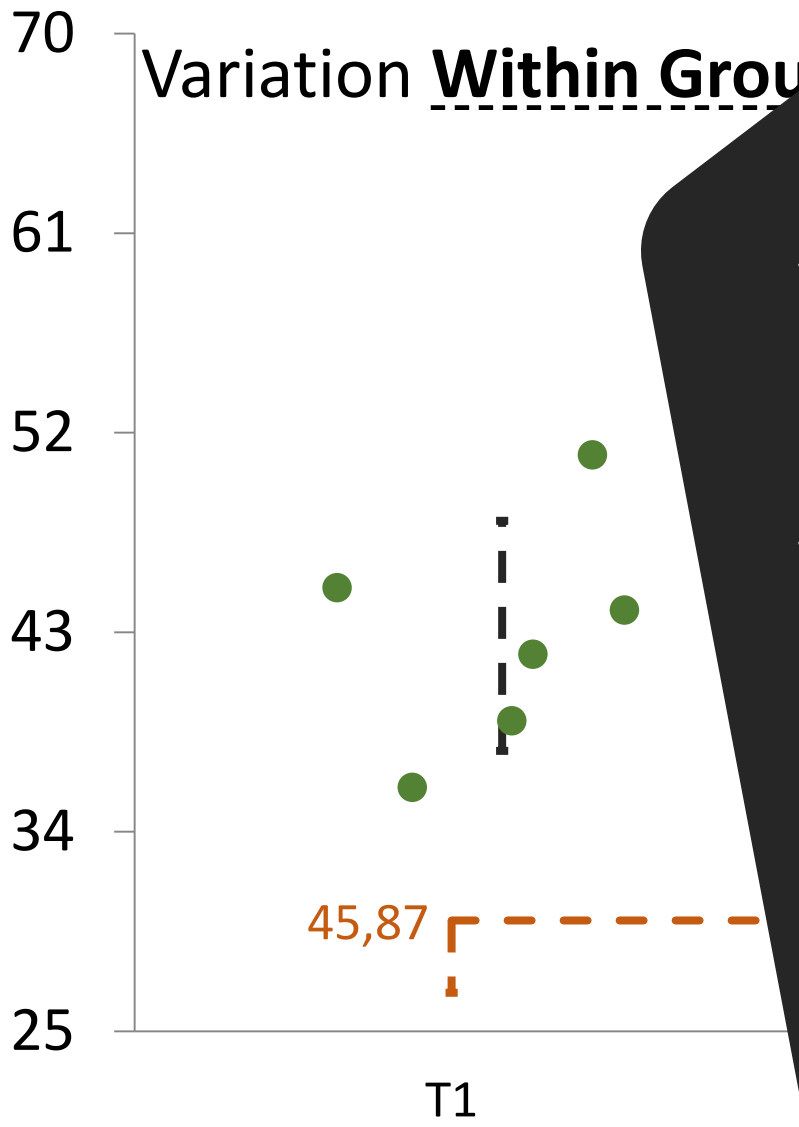
01

Repeated Measures

The Regular ANOVA

The ANOVA


Variation Within Groups and Between Groups
70
61
52
43
34
25
45,87
T1
T2
T3



Why is this group lower than the overall average?

1) People in T1 score lower than average (group effect)
2) The average is affected by error, pulling it up or down.

Why is this person higher or lower than the group average?

1) Because of factors other than being in the group

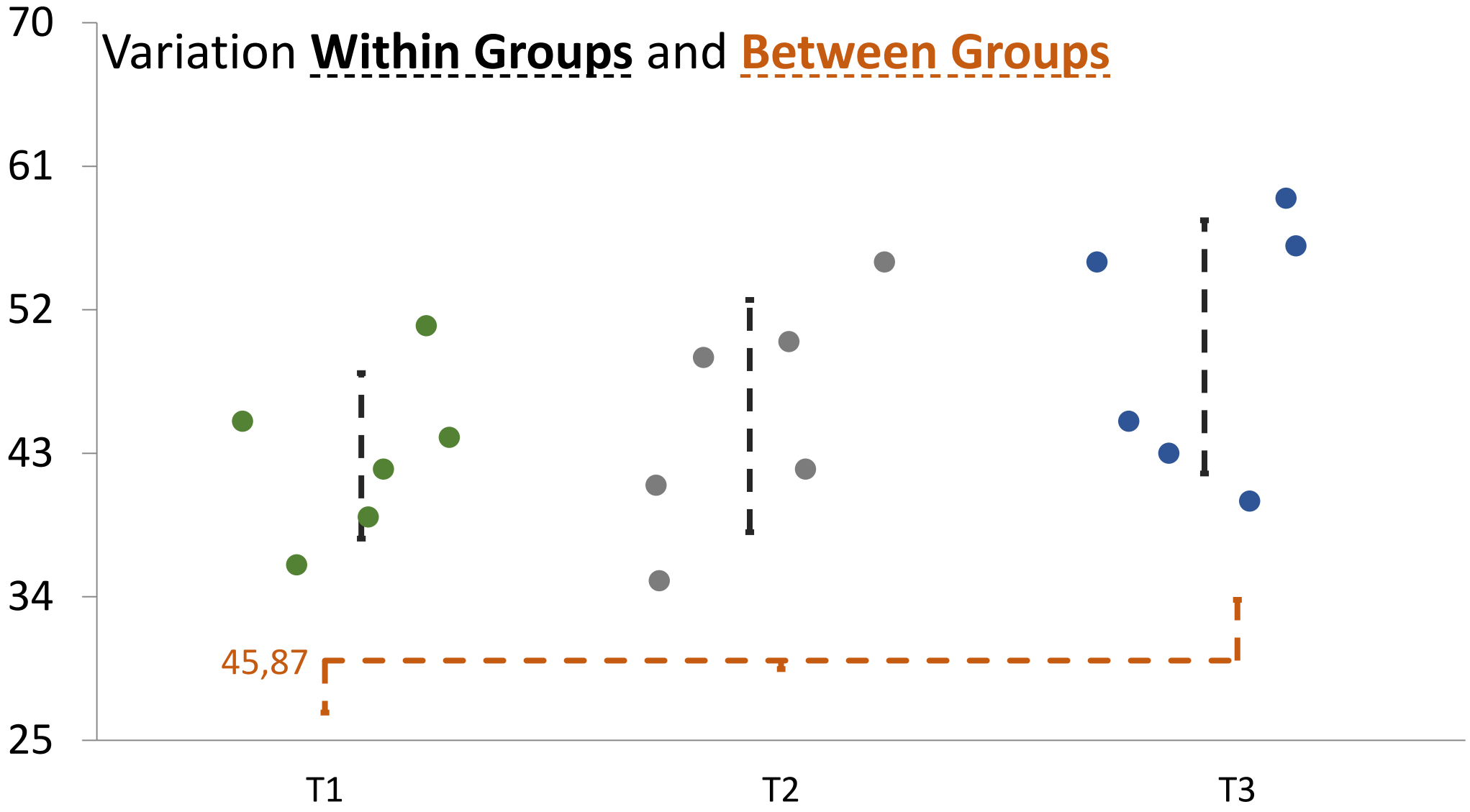
$$F = \frac{MS(Groups) + MS(Error)}{MS(Error)}$$

Variation caused by Making Groups (biased)

is larger than

Variation caused by Individuals (unbiased)

The ANOVA


Variation Within Groups and Between Groups
70
61
52
43
34
25
45,87
T1
T2
T3

Let’s run an ANOVA then…

PP	Value	Condition
1	46	1
1	55	2
1	68	3
2	44	1
2	44	2
2	61	3
3	72	1

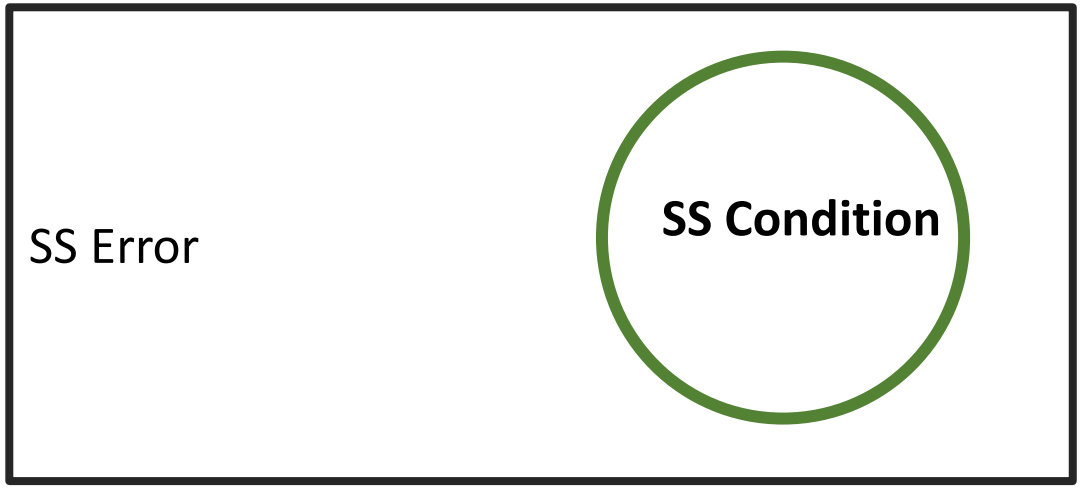
Tests of Between-Subjects Effects

Dependent Variable: Data

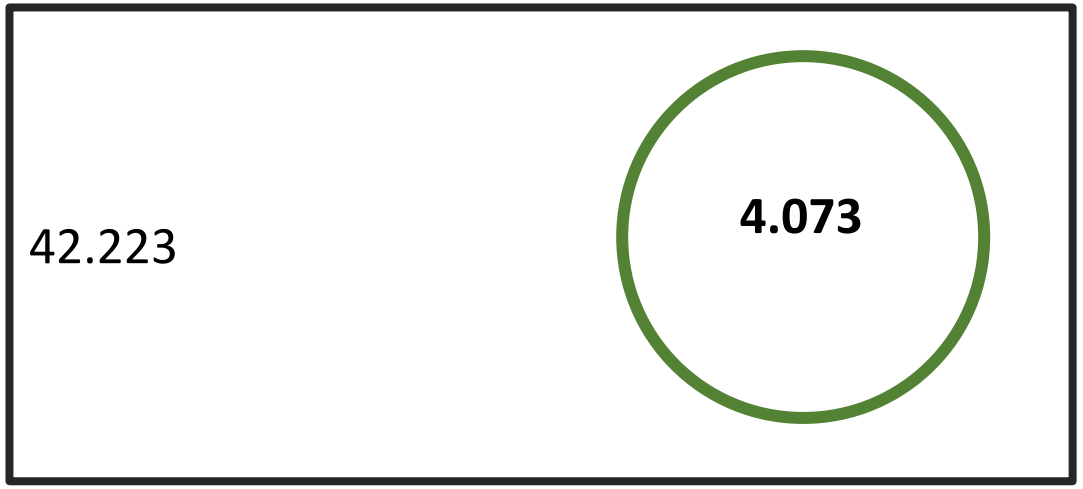
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4.073[a]	2	2.037	5.643	.005
Intercept	1741.289	1	1741.289	4825.065	.000
CONDITION	4.073	2	2.037	5.643	.005
Error	42.223	117	.361		
Total	1787.586	120			
Corrected Total	46.297	119			

a. R Squared = .088 (Adjusted R Squared = .072)

SS Total



46.297



The ANOVA

People can be very consistent

–

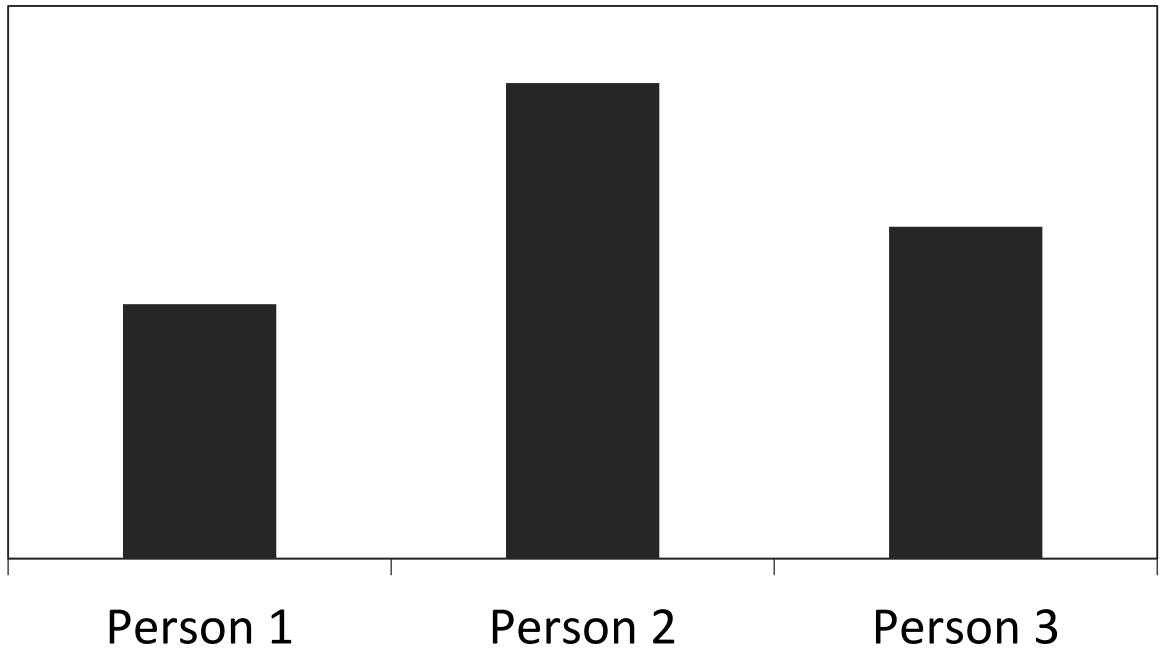
Part of the error is explained by having the same **Person** doing the same test

The ANOVA

But we want to take the **general Person** effect into account

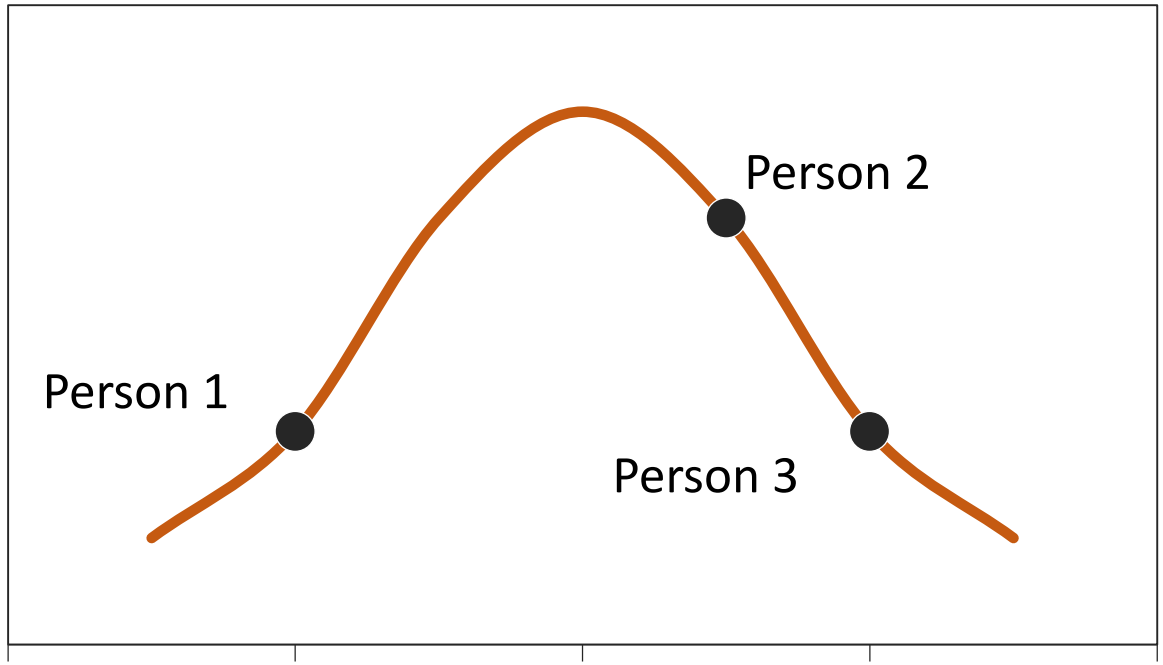
-

We add **Person** as a **Random Effect**



Person is just another variable, just another factor that can explain **variation**. Most of the time it explains a ton of it because people (or any kind of subject) are very consistent when you test them repeatedly.

But person is not a regular factor, it's a perfect example of a random factor.



A **fixed effect** looks at the difference between levels
A **random effect** assumes a population effect

Do we want to know the difference between people or do we want to know the general person effect?

Let’s add **Person** to the model…

PP	Value	Condition
1	46	1
1	55	2
1	68	3
2	44	1
2	44	2
2	61	3
3	72	1

Tests of Between-Subjects Effects

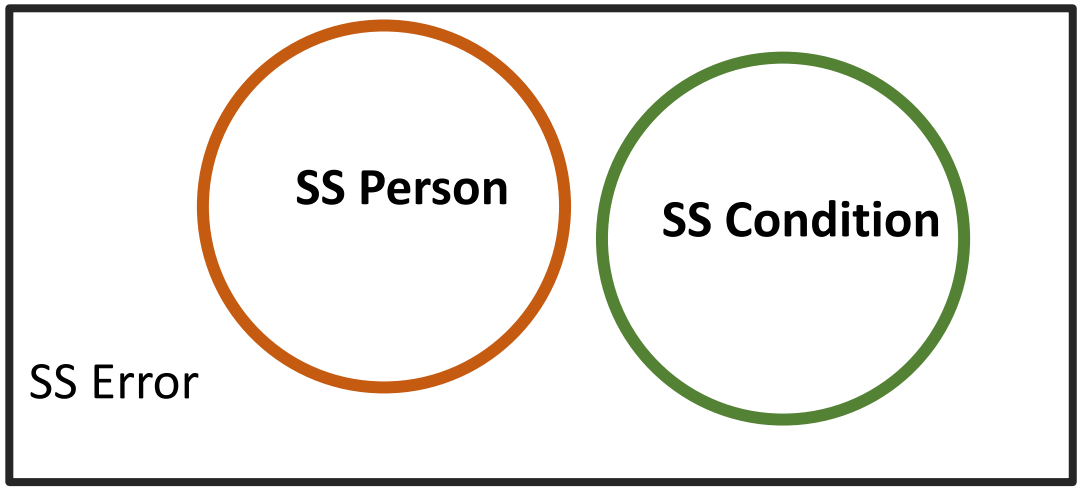
Dependent Variable: Data

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	1741.289	1	1741.289	1893.681	.000
	Error	35.862	39	.920[a]		
CONDITION	Hypothesis	4.073	2	2.037	24.969	.000
	Error	6.362	78	.082[b]		
Person	Hypothesis	35.862	39	.920	11.274	.000
	Error	6.362	78	.082[b]		

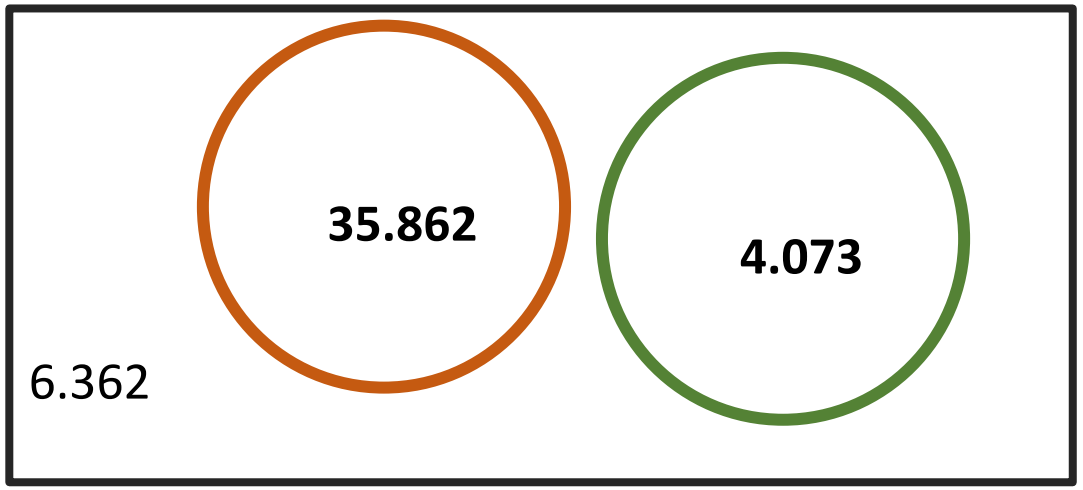
a. MS(PP)

b. MS(Error)

SS Total



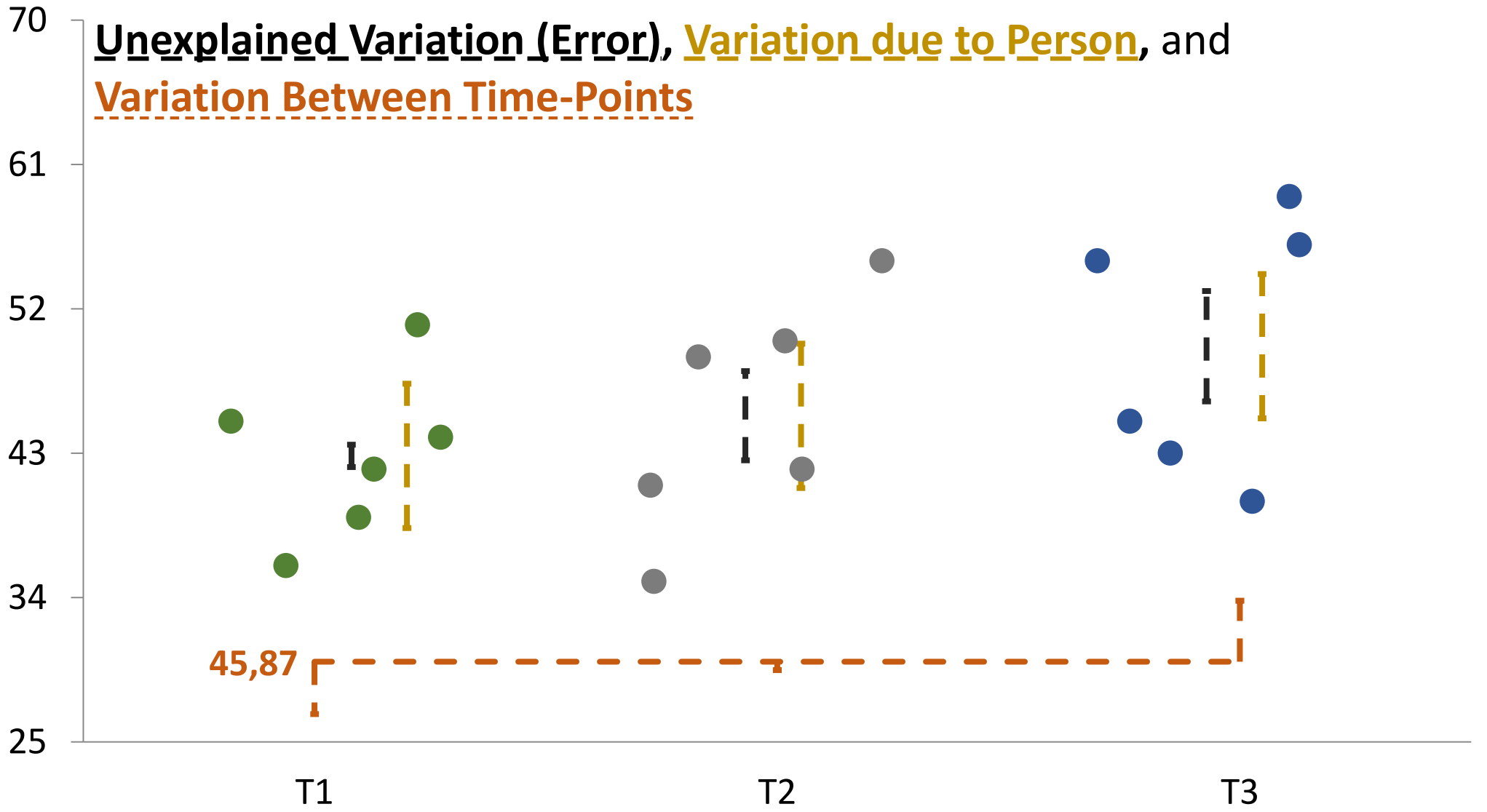
46.297

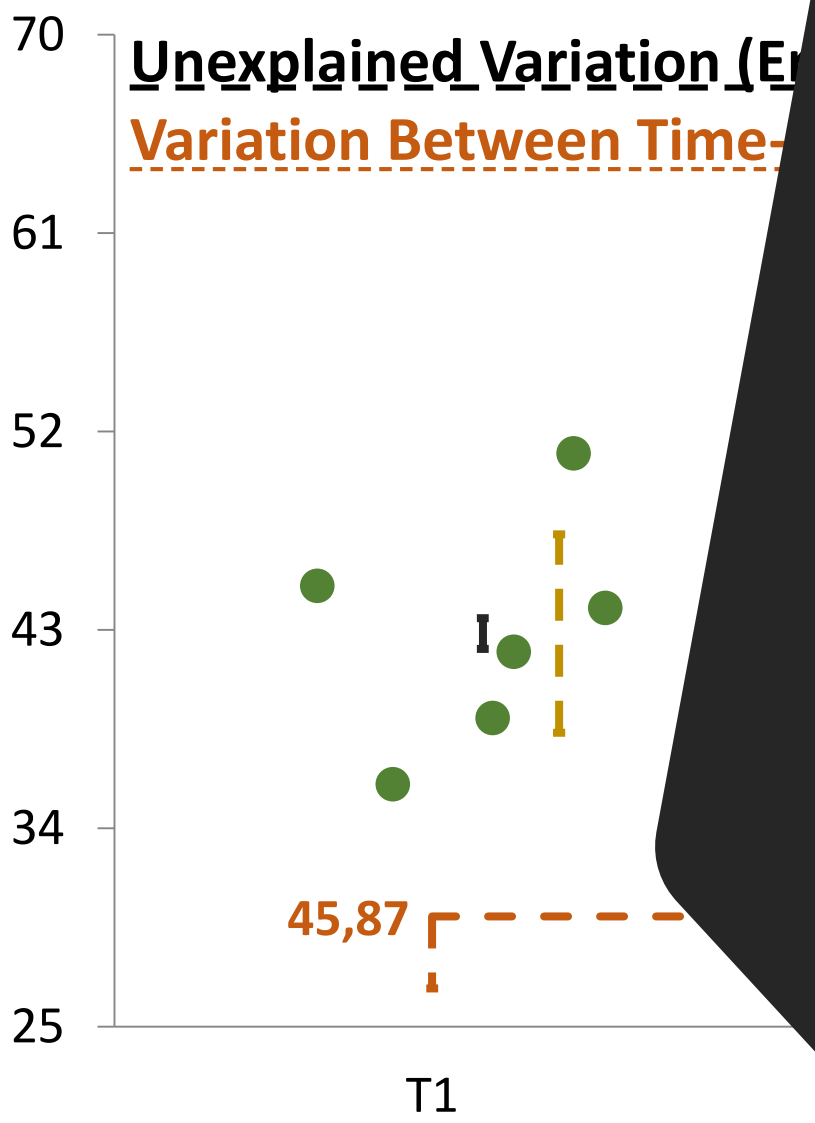


Person explains a lot of error, this improves the model (there's less unaccounted for)

It also turns 120 observations into 40 observations measured three times

The RM ANOVA


Unexplained Variation (Error), Variation due to Person, and
Variation Between Time-Points
70
61
52
43
34
25
45,87
T1
T2
T3



Unexplained Variation (Er

Variation Between Time-

Why is T1 lower than the overall average?

1) People in at Time-Point 1 scored lower than average
2) Error due to factors not controlled by the experiment.
3) These specific people are in T1 (some people are simply better/worse on this task).

Why does this specific person have a lower/higher score?

1) Error due to factors out of our control.
2) The person effect.

$$F = \frac{MS(Time) + MS(Time * PP) + MS(Error)}{MS(Time * PP) + MS(Error)}$$

Person improves the model by making it **Repeated Measures**, it explains more **variance** and lowers the **standard error**

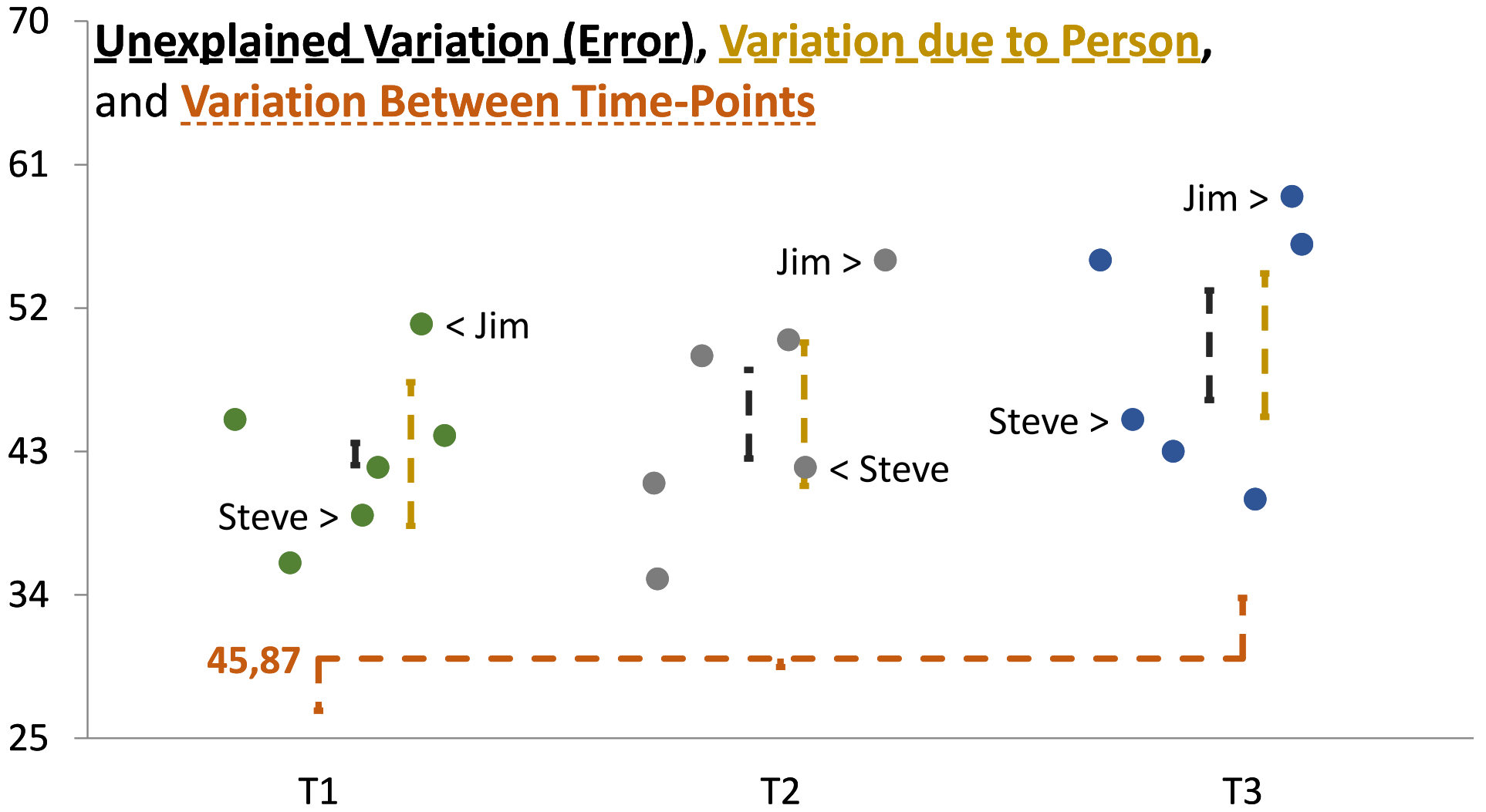
(getting us closer to the population mean)

02

Sphericity

In a normal ANOVA the variances must be equal in each group (Homogeneity of variances). This means that **MSE is an Unbiased Estimator of the error variance** (we expect it to only express error).

When we add a **random person** effect we assume that the effect of person is the same at each time-point. The variation caused by person doesn't change and is still unbiased.

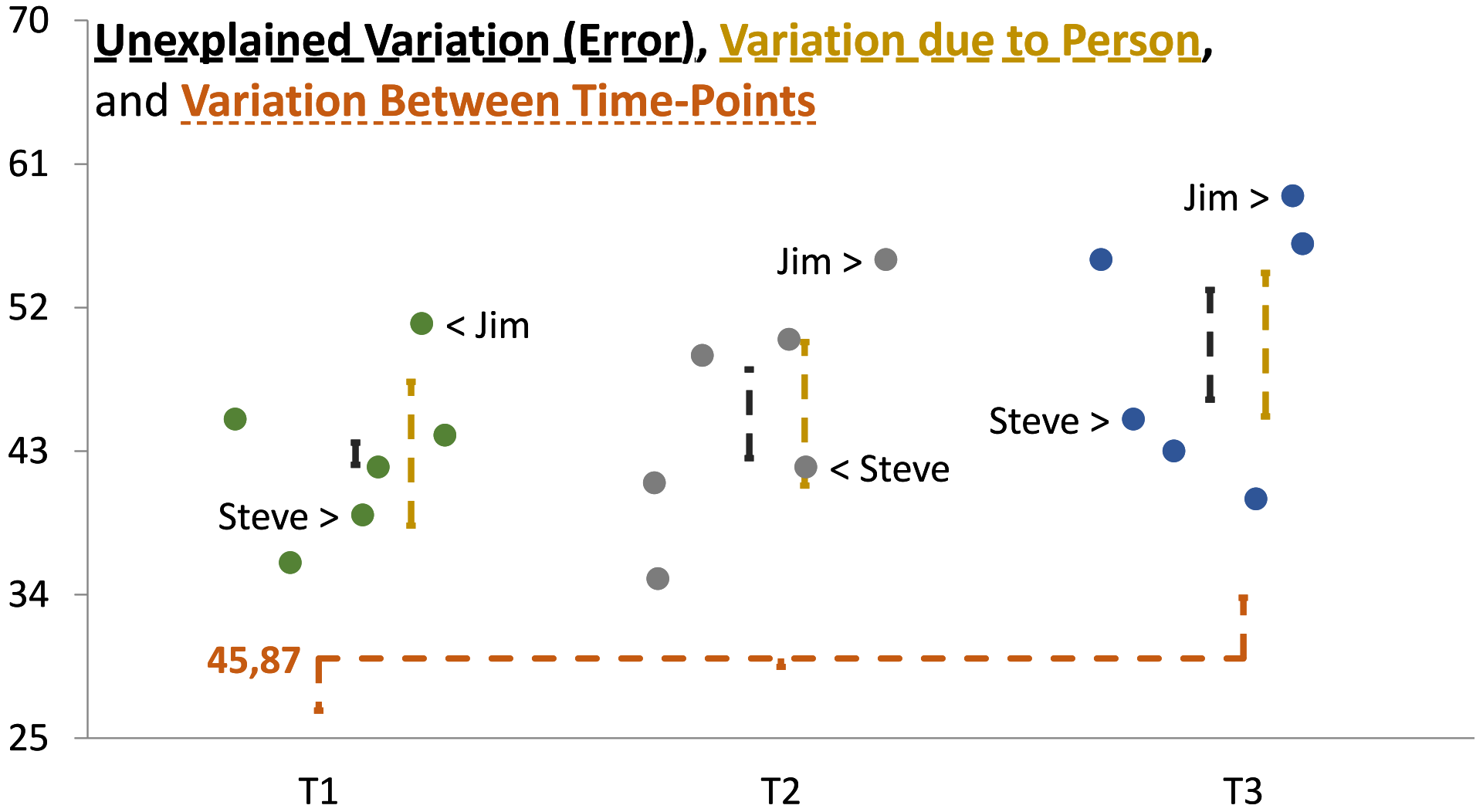


Sphericity

The effect of person is always the same

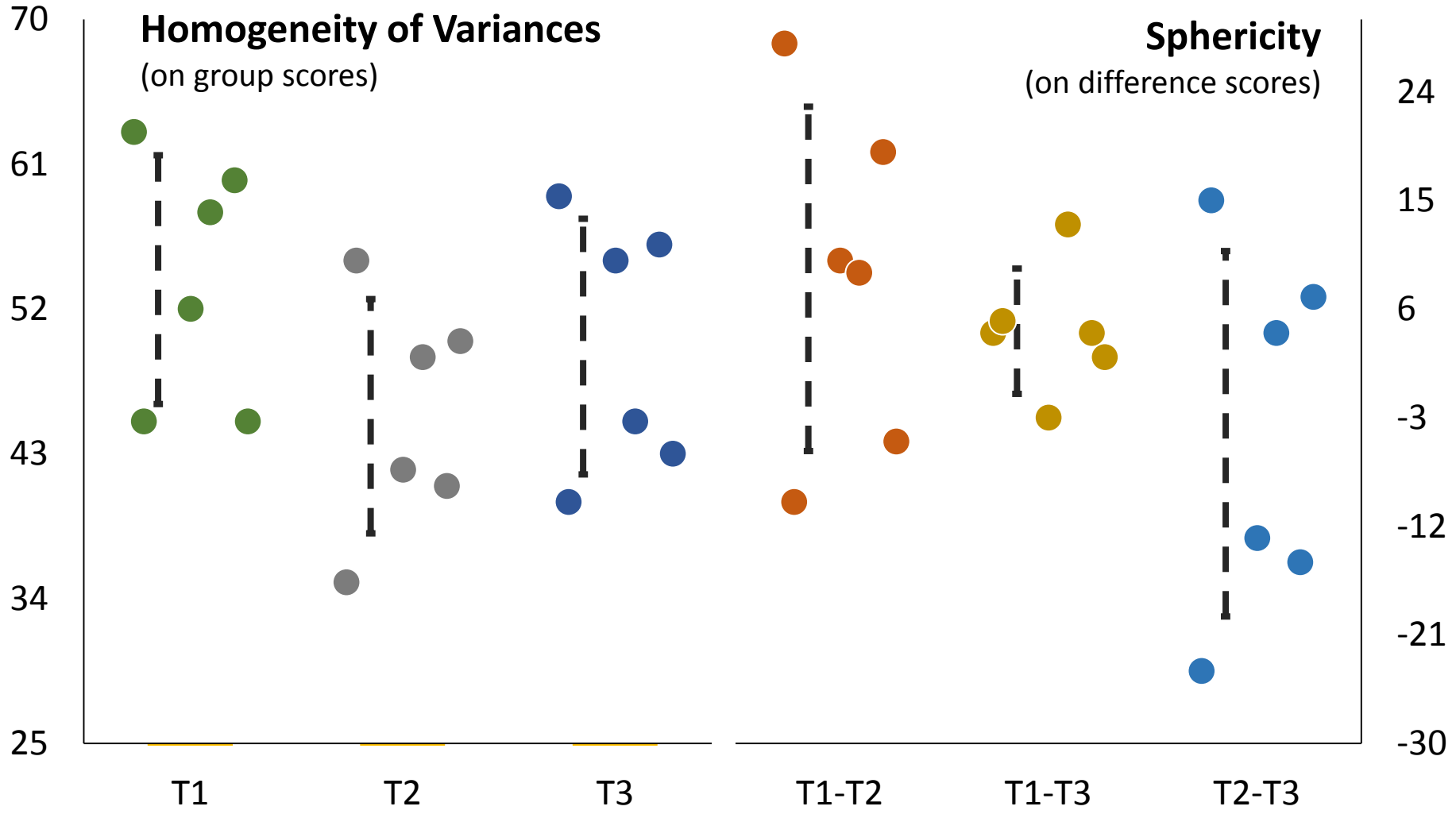
It’s not about variation between individuals but about the correlation between replications

Jim has a high score, it's a repeated measures so **Jim** has a high score each time. The **Person** effect is the same everywhere, so **Jim** is always 10 points higher than **Steve**.
This is the assumption we make, does this make sense?



What if Jim and Steve switch in the second time-point? Instead of Jim being 10 points better than Steve, it's Steve who is 10-points better than Jim. What if everyone switches? The lowest becomes the highest and vice versa.

The variances are still the same, the differences between participants in all groups are more or less the same, even if we reverse who has the highest score in T2. But because the scores reverse in T2, the time-points no longer share the same correlation.



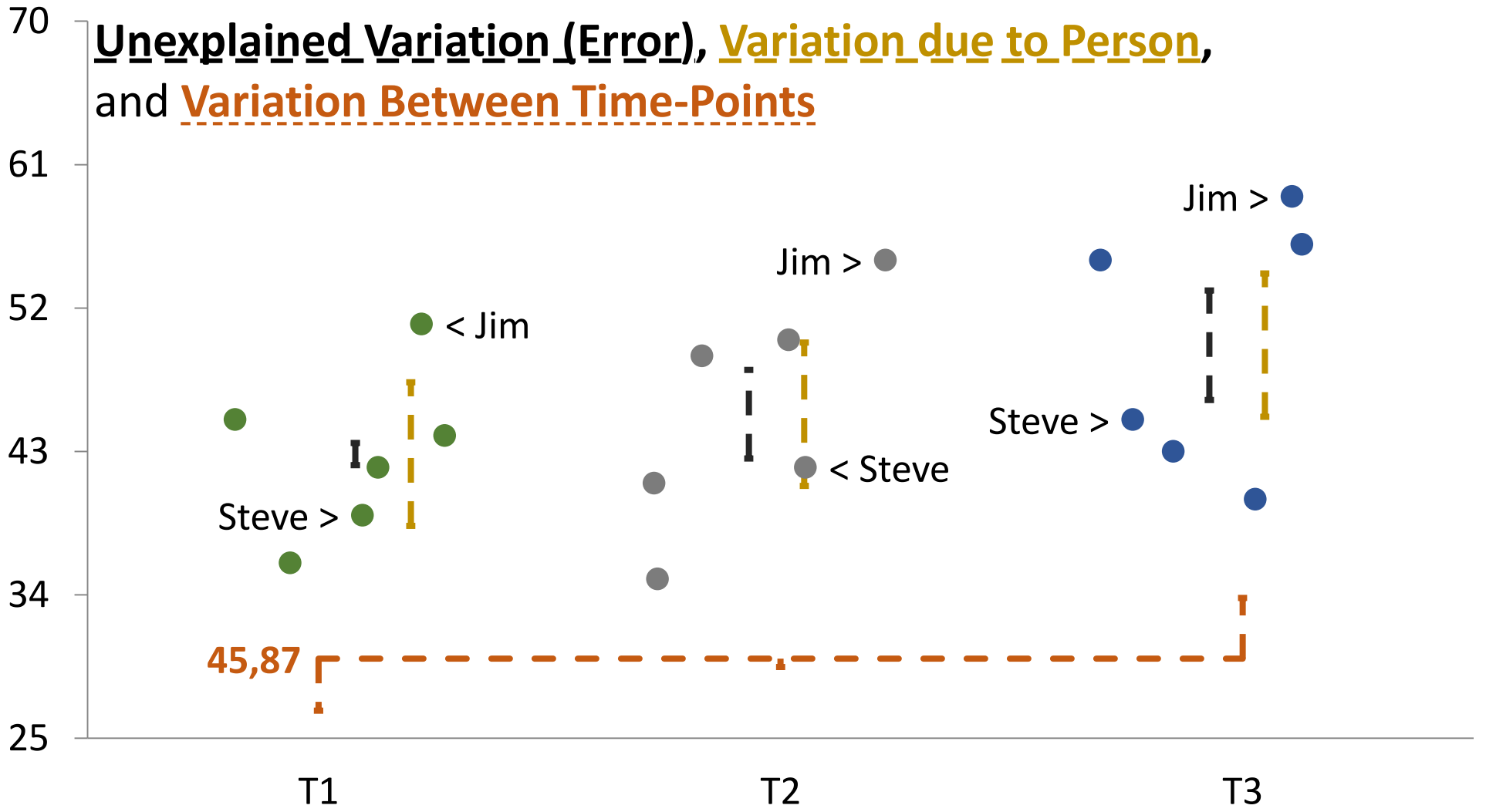
Sphericity

Sphericity is a concern when you have more than two replications (more than one correlation)

-or-

When you **assume** that correlations are equal

The RM ANOVA


Unexplained Variation (Error), Variation due to Person,
and Variation Between Time-Points
70
61
52
43
34
25
< Jim
Steve >
Jim >
< Steve
Jim >
Steve >
45,87
T1
T2
T3

The RM ANOVA

Variation caused by Replications (Time or Groups)

is larger than

Variation caused by Individuals (unbiased)

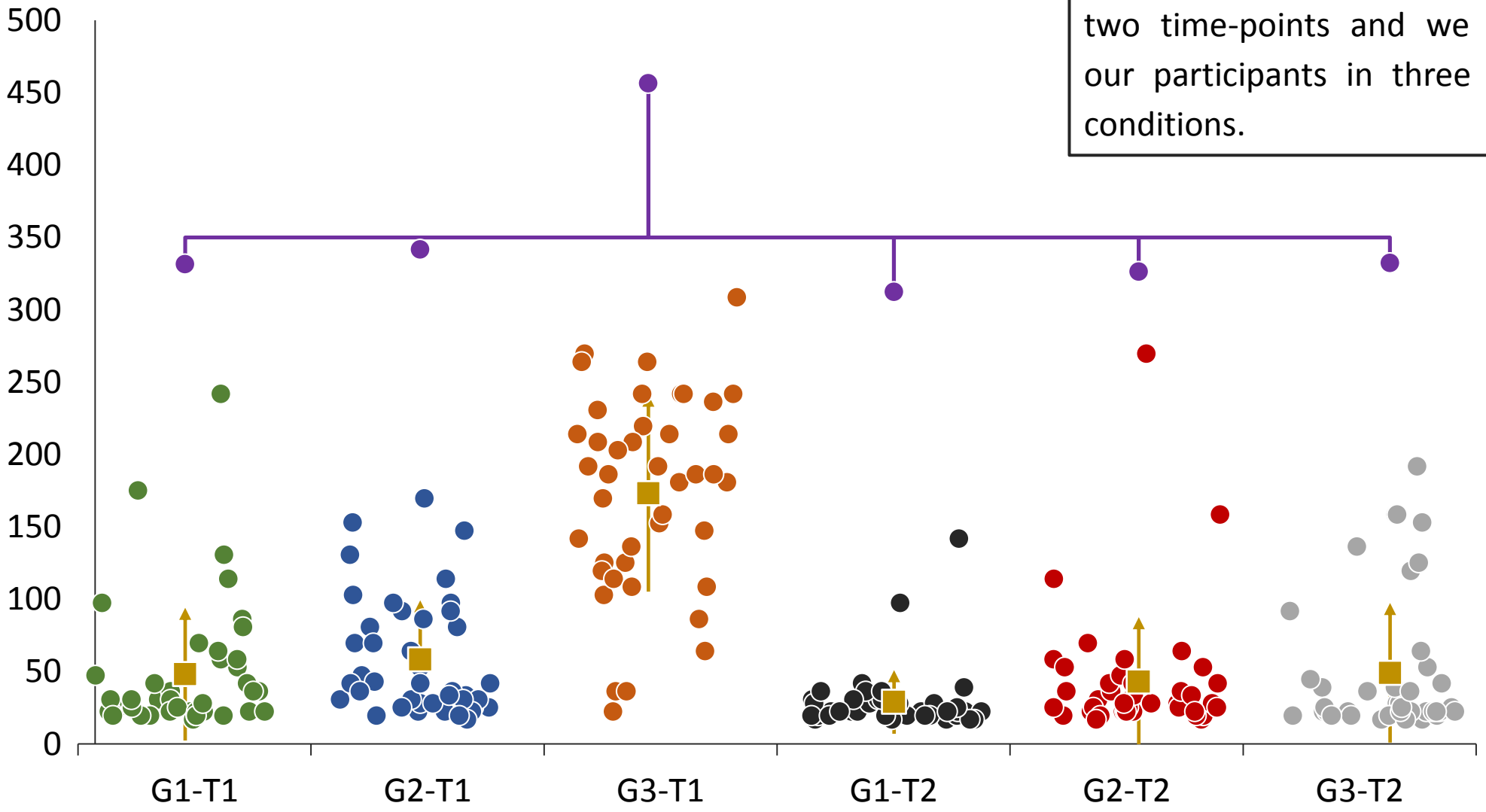
Short Break

03

Example

The Example

For the example I chose a slightly more complicated design, a crossed 2x3 within subject design. There are two time-points and we measure our participants in three different conditions.



The Repeated Measures ANOVA

The Multivariate Test

PP	T1			T2		
	G1	G2	G3	G1	G2	G3
1	-	-	-	-	-	-
2	-	-	-	-	-	-

Multivariate Tests

Effect	Value	F	Hypothesis df	Error df	Sig.
Time	.787	151.475	1	41	.000
Group	.738	56.379	2	40	.000
Time * Group	.654	37.742	2	40	.000

The **Multivariate Tests** ignores the correlation between replications, only taking into account the fact that replications come from the same subjects. It is literally a MANOVA and has no assumption of **sphericity**.

The Multivariate Test

The Multivariate Test is a MANOVA on Repeated Measures where Sphericity plays no role

-

It assumes no dependency between replications except for the dependency caused by a common source (person)

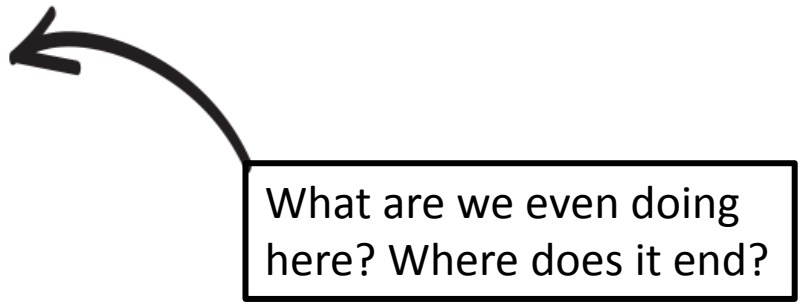
PP	T1			T2		
	G1	G2	G3	G1	G2	G3
1	-	-	-	-	-	-
2	-	-	-	-	-	-

Mauchly's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
Time	1.000	.000	0	.
Group	.691	14.766	2	.001
Time * Group	.662	16.477	2	.000

Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	176352.868	1.000	176352.868	151.475	.000
Error(Time)	47733.841	41.000	1164.240		
Group	254081.038	1.528	166254.728	82.023	.000
Error(Group)	127005.312	62.659	2026.935		
Time * Group	159956.084	1.495	106981.366	56.584	.000
Error(Time*Group)	115902.626	61.302	1890.675		



Looking at the Test of Within-Subject Effects we can see main effects of Time ($F_{(1,41)}$=151.47, p<.001), Group ($F_{(1.53,62.66)}$=82.02, p<.001), and their interaction ($F_{(1.50,61.30)}$=56.58, p<.001).

The Univariate Test

The Univariate Test assumes Sphericity

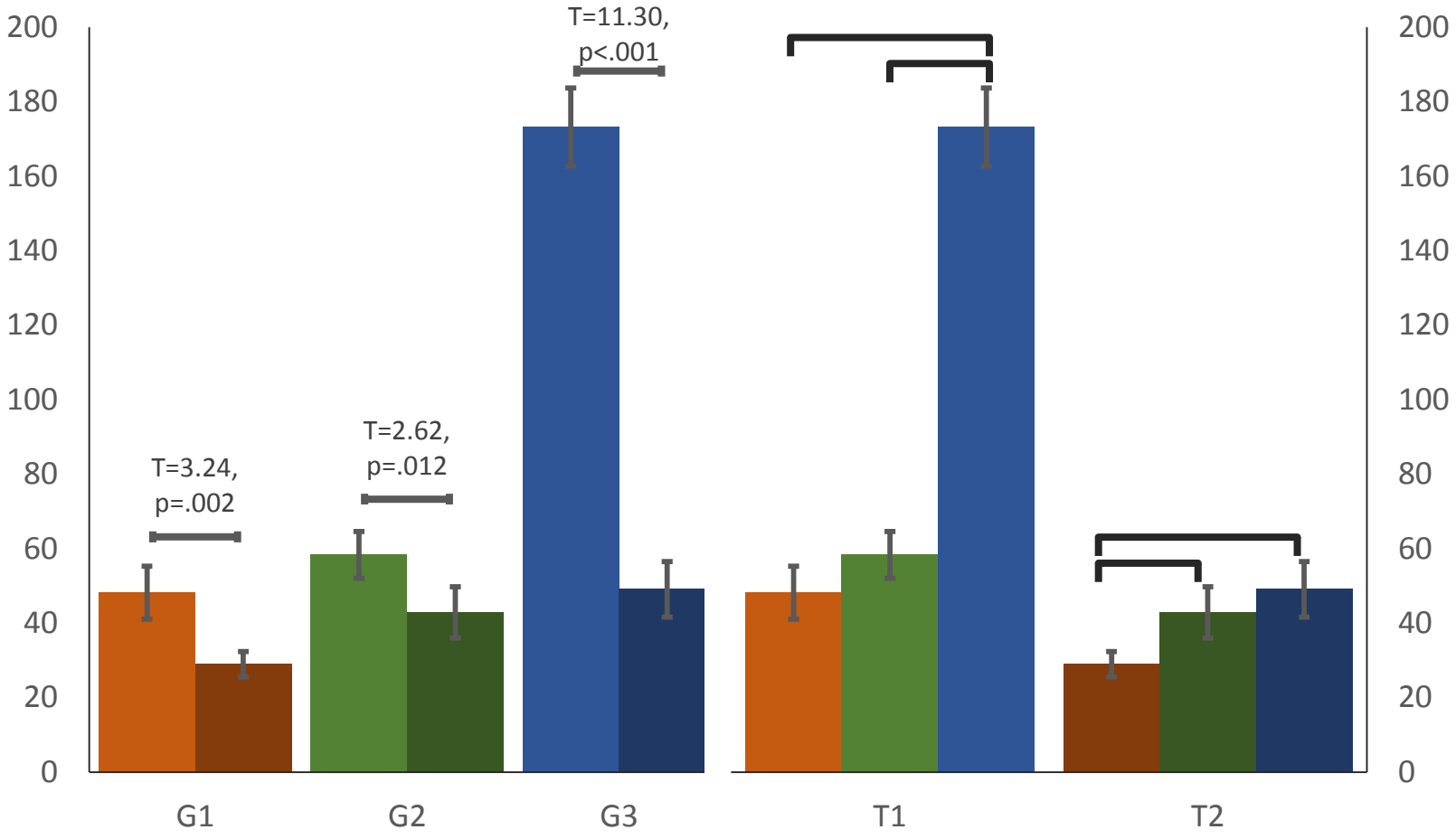
-

It assumes dependency between replications and dependency caused by a common source

-

This is a ridiculous assumption and needs to be corrected for

Pairwise Comparisons



we have Time effects in each group, which is nice. The other way around we see a group effect in both time-points. In T1 it's group 3 that is significantly higher than the other two, while in T2 it's group 1 that is significantly lower than the other two.

Pairwise Comparisons

Group	Mean Difference (T1-T2)	Std. Error	Sig. Sidak
T1	19,218*	5,924	,002
T2	15,373*	5,860	,012
T3	124,133*	10,983	,000

Pairwise Comparisons

Time	Group	Group	Mean Difference	SE	Sig.
T1	G1	G2	-10.110	6.915	.389
		G3	-125.021*	12.483	.000
	G2	G3	-114.912*	12.004	.000
T2	G1	G2	-13.955*	3.852	.002
		G3	-20.106*	5.556	.002
	G2	G3	-6.151	5.454	.605

The Repeated Measures ANOVA

EVERYTHING WRONG WITH

The Repeated Measures ANOVA

Mauchly's Test of Sphericity

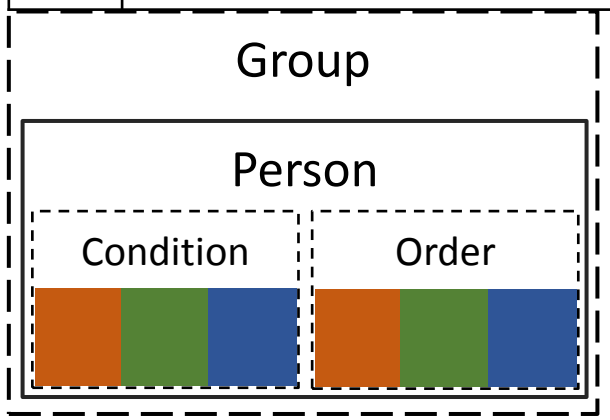
Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon Correction		
					GG	HF	LB
Time	1.000	.000	0	.	1.000	1.000	1.000
Group	.691	14.766	2	.001	.764	.788	.500
Time * Group	.662	16.477	2	.000	.748	.769	.500

What are we even doing here? Where does it end?

Unrealistic expectations about the data.
Needing a test to test the test we are using to test our hypothesis.

Group	PP	Condition	Order
1	1	1	1
1	1	2	2
1	1	3	3
1	2	1	2
1	2	2	3
1	2	3	1
2	3	1	3
2	3	2	1
2	3	3	2
2	4	1	1
2	4	2	2
2	4	3	3
3	5	1	2

PP	Cond1	Cond2	Cond3
PP	Cond1-Ordr1	Cond1-Ordr2	Cond1-Ordr3
1	45	--	--
2	--	32	--
3	--	--	34
4	54	--	--

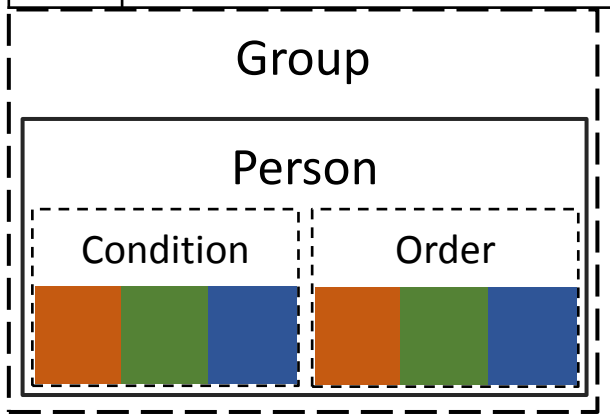


That's never going to run

No support for multiple levels or variables on the same level

Group	PP	Condition	Order
1	1	1	1
1	1	2	2
1	1	3	3
1	2	1	2
1	2	2	3
1	2	3	1
2	3	1	3
2	3	2	1
2	3	3	2
2	4	1	1
2	4	2	2
2	4	3	3
3	5	1	2

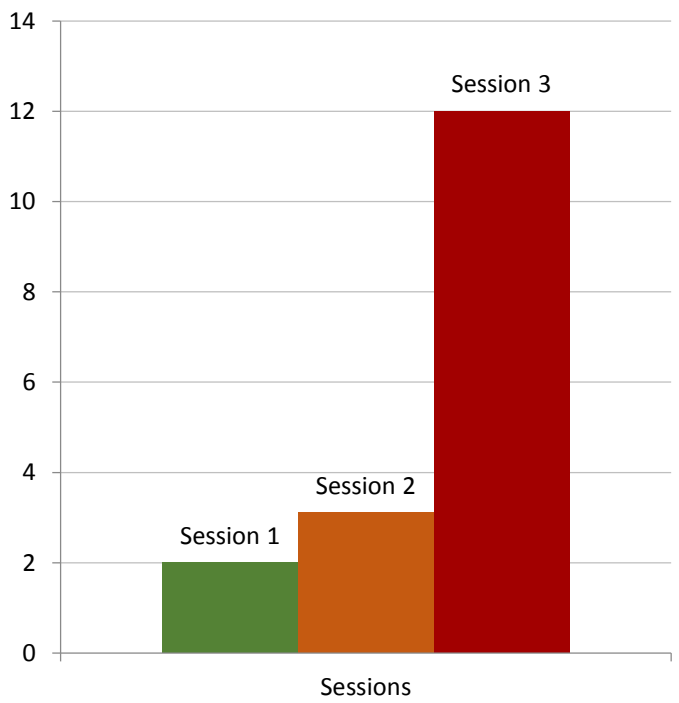
PP	Cond1	Cond2	Cond3
PP	Cond1-Ordr1	Cond1-Ordr2	Cond1-Ordr3
1	45	--	--
2	--	32	--
3	--	--	34
4	54	--	--

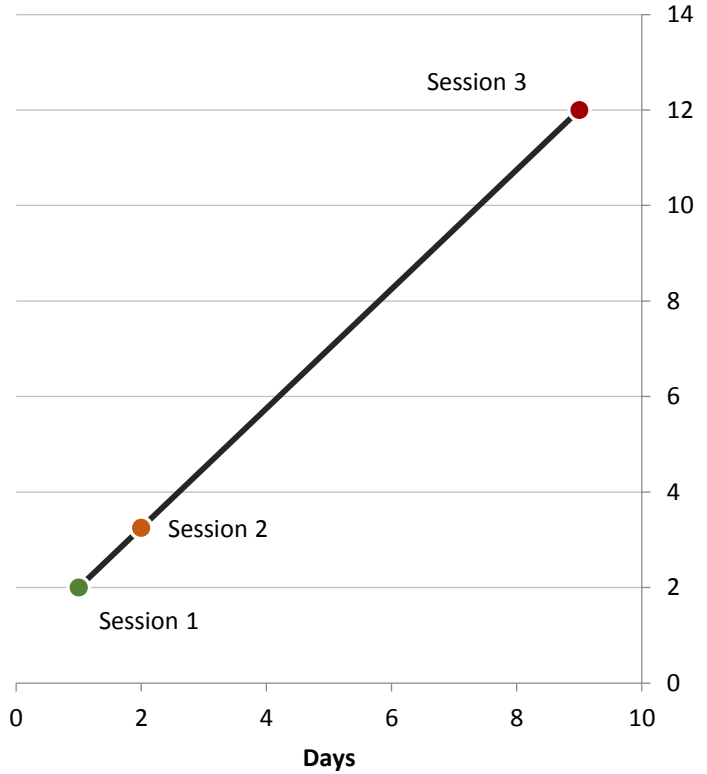


That's never going to run

Too zealous with missing data, removing entire participants when they miss a single observation

Session 1 | Session 2 | Session 3





Time is always a factor when we know time isn't that categorical

The Linear Mixed Procedure

The Model Summary

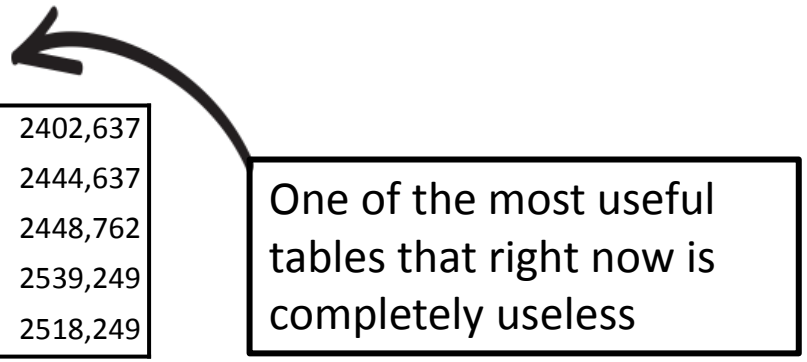
PP	Time	Group
1	1	1
1	1	2
1	1	3
1	2	1
1	2	2
1	2	3
2	1	1
2	1	2
2	1	3
2	2	1
2	2	2
2	2	3

Model Dimension

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables	Number of Subjects
Fixed Effects	Intercept	1		1		
	Time	2		1		
	Group	3		2		
	Time * Group	6		2		
Repeated Effects	Time * Group	6	Unstructured	21	PP	42
Total		18		27		

Information Criteria

-2 Restricted Log Likelihood	2402,637
Akaike's Information Criterion (AIC)	2444,637
Hurvich and Tsai's Criterion (AICC)	2448,762
Bozdogan's Criterion (CAIC)	2539,249
Schwarz's Bayesian Criterion (BIC)	2518,249



The Fixed Effects Model

Type III Tests of Fixed Effects[a]

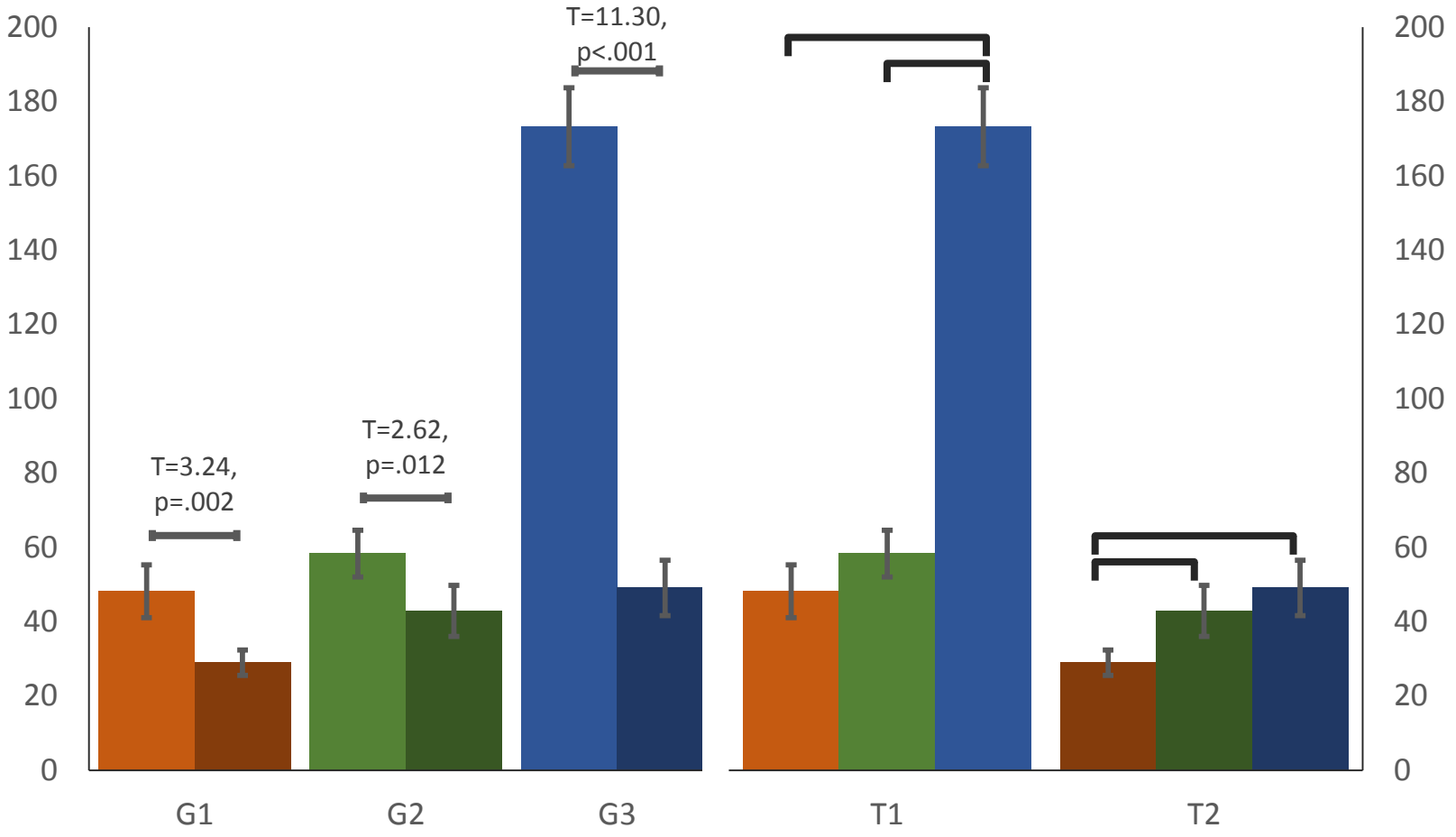
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	41.000	183.959	.000
Time	1	41.000	151.474	.000
Group	2	41.000	57.788	.000
Time * Group	2	41.000	38.684	.000

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	49.074048	7.437278	41	6.598	.000
[Time=1]	124.133571	10.983209	41	11.302	.000
[Group=1]	-20.106190	5.555638	41	-3.619	.001
[Group=2]	-6.150714	5.454389	41	-1.128	.266
[Time=1] * [Group=1]	-104.915714	13.819458	41	-7.592	.000
[Time=1] * [Group=2]	-108.760000	12.385732	41	-8.781	.000

$$\hat{Y} = \beta_o + \beta_1[Time = 1] + \beta_2[Group = 1] + \beta_3[Group = 2] + \beta_4[Time = 1; Group = 1] + \beta_5[Time = 1; Group = 2]$$

Pairwise Comparisons



Pairwise Comparisons

Group	Mean Difference (T1-T2)	Std. Error	Sig. Sidak
T1	19,218*	5,924	,002
T2	15,373*	5,860	,012
T3	124,133*	10,983	,000

Pairwise Comparisons

Time	Group	Group	Mean Difference	SE	Sig.
T1	G1	G2	-10.110	6.915	.389
		G3	-125.021*	12.483	.000
	G2	G3	-114.912*	12.004	.000
T2	G1	G2	-13.955*	3.852	.002
		G3	-20.106*	5.556	.002
	G2	G3	-6.151	5.454	.605

In its simplest form

Mixed Models can produce the same results and conclusions as **RM ANOVA**

-

But it’s capable of so much more, less bound by arbitrary limitations or statistical assumptions

This Session

- The Logic of Repeated Measures
- Sphericity
- Repeated Measures ANOVA output
- Repeated Measures ANOVA Limitations
- Repeated Measures with the Mixed Procedure

The Next Sessions

How to build Models and decide on covariance structures

(plus talk about what covariance structures are)

How to add covariates, random intercepts, and random slopes

(Making the Mixed Models truly Mixed)

Adding Planned and Polynomial Contrasts

(Thinking ahead like a good data scientist)

Session Evaluation