

ANOVA: Analysis of Variance

16 Oct 2009
CPSY 501
Dr. Sean Ho
Trinity Western University

- **REB** forms are due!
- Please download:
Treatment4.sav

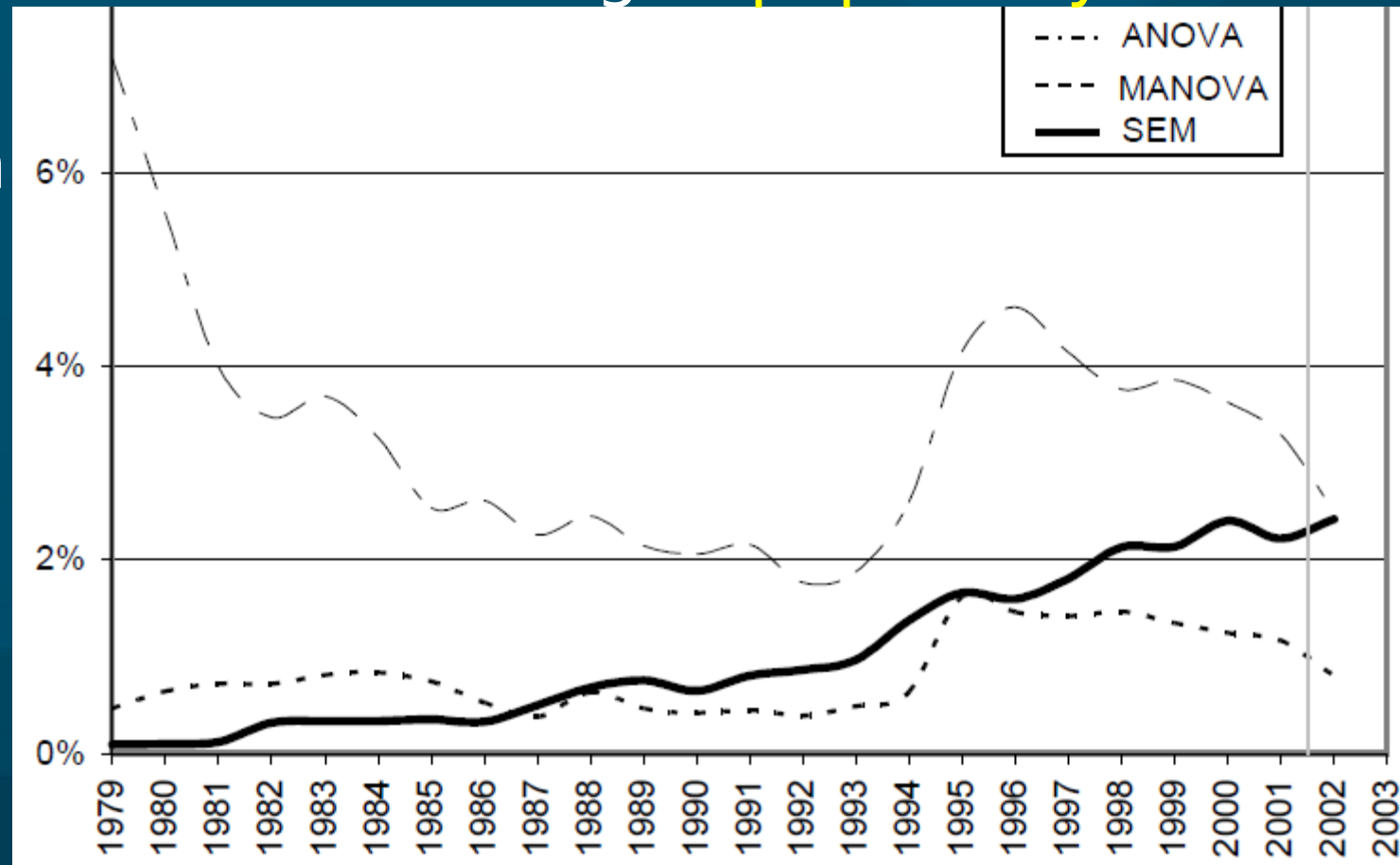
Outline for today

- Core concepts of ANOVA (with pictures!)
 - One-way, factorial, RM, MANOVA, ANCOVA
- Running ANOVA in SPSS
 - Interpreting output
 - Follow-up analysis: post-hoc comparisons
 - Follow-up analysis: planned comparisons
- Assumptions of ANOVA: Parametricity
- Introduction to ANCOVA
 - Use of ANCOVA in therapy research

Trends in Research: ANOVA

Nachtigall, C., Kroehne, U., Funke, F., & Steyer, R. (2003). (Why) Should we use SEM? Pros and cons of structural equation modeling. *Methods of Psychological Research - Online*, 8(2), 1-22.

- ANOVA is a **conceptual** framework / foundation
- **Other** methods like Structured Equation Modelling are also increasing in **popularity**
- **Citations** in PsychINFO, 1979-2003



ANOVA: Core Concepts

- Extension of *t*-test to multiple groups
- Why not just do a bunch of *t*-tests (all pairs)?
 - Multiple comparisons: to control Type-I error, we'd have to spread out our α across all *t*-tests, so each one is a stricter test
 - We will do this in post-hoc analysis!
- ANOVA can be thought of as a form of regression, with categorical predictors
- Global test for group differences (omnibus)
 - Post-hoc tests then locate the differences

Kinds of ANOVA

- All ANOVAs require **parametric outcome** var!
- **One-way** ANOVA: **One** categorical predictor
 - If IV has only two levels, this is a **t-test**!
- **Factorial** (“between subjects”) ANOVA:
Two or more predictors, plus **interactions**
- **Repeated Measures** (“within subjects”) ANOVA:
Multiple observations of IV on each participant
 - e.g., **time**: pre/post/follow-up

Kinds of ANOVA (continued)

- **Mixed** (between/within) ANOVA:
Some predictors are **between-groups** (factorial), some are **within-groups** (repeated measures)
- **MANOVA**: Two or more **outcome** variables, correlated, and in the same analysis
- **ANCOVA**: Any of the above designs, trying to control for an extraneous influence (**covariate**) on the DV

ANOVA: The *F*-ratio

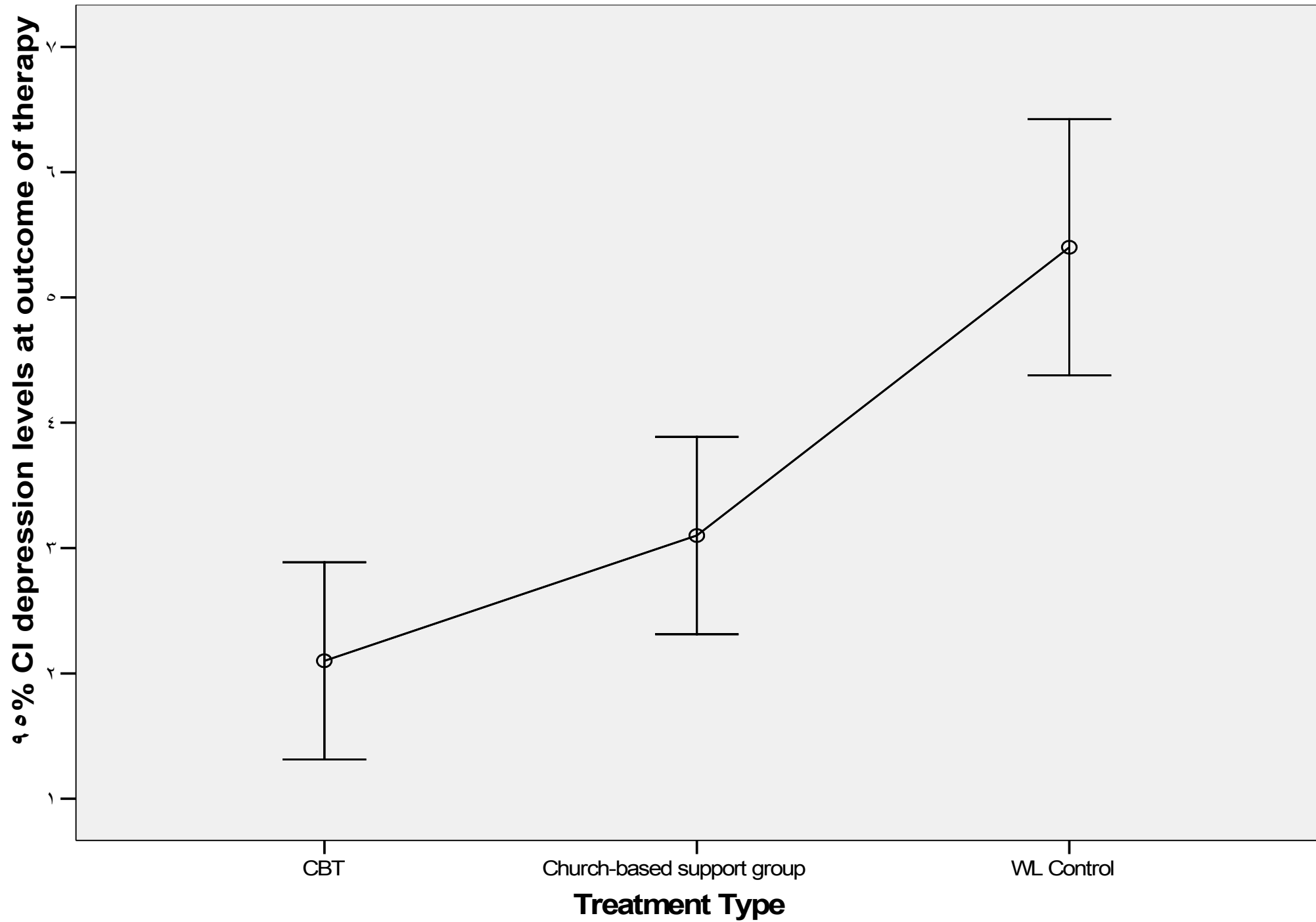
- ANOVA's test statistic is the *F*-ratio:

$$F = MS_{\text{model}} / MS_{\text{residual}}$$

- “**Model**”: variation/difference **among** cells
 - “**Residual**”: variation **within** each cell
 - Same as *F*-ratio in **regression**: **model** / **error**
- Categorical IVs divide the sample into **cells**:
all **combinations** of groupings from the IVs
 - e.g., Treatment4.sav:
 - 1 **IV** (**treatment type**) with 3 **levels**:
CBT (cog-beh), **CSG** (church), **WL** (control)

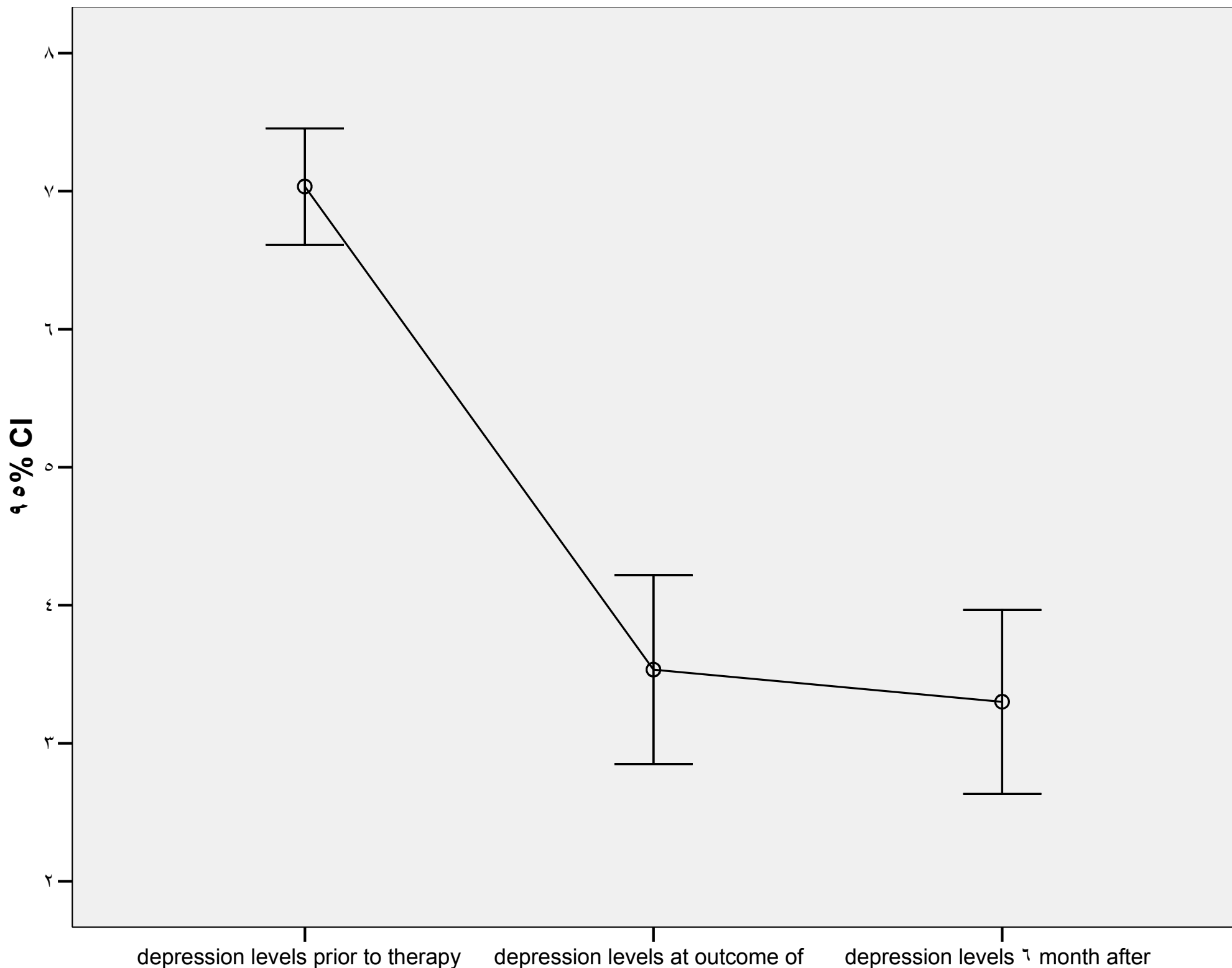
Picture: Within-Cell Variation

- Error bar charts can help visualize both the “between-cell” and “within-cell” variation
 - Confidence intervals around cell means describe within-cell variation (residual)
 - Cell mean differences describe between-cell variation (model effect of IV)
- SPSS: Graphs → Legacy Dialogs → Error bar → Simple and “groups of cases”:
 - Variable (DV): depression levels at outcome
 - Category Axis (IV): treatment type
 - (Also try: Line, with Options → “error bars”)



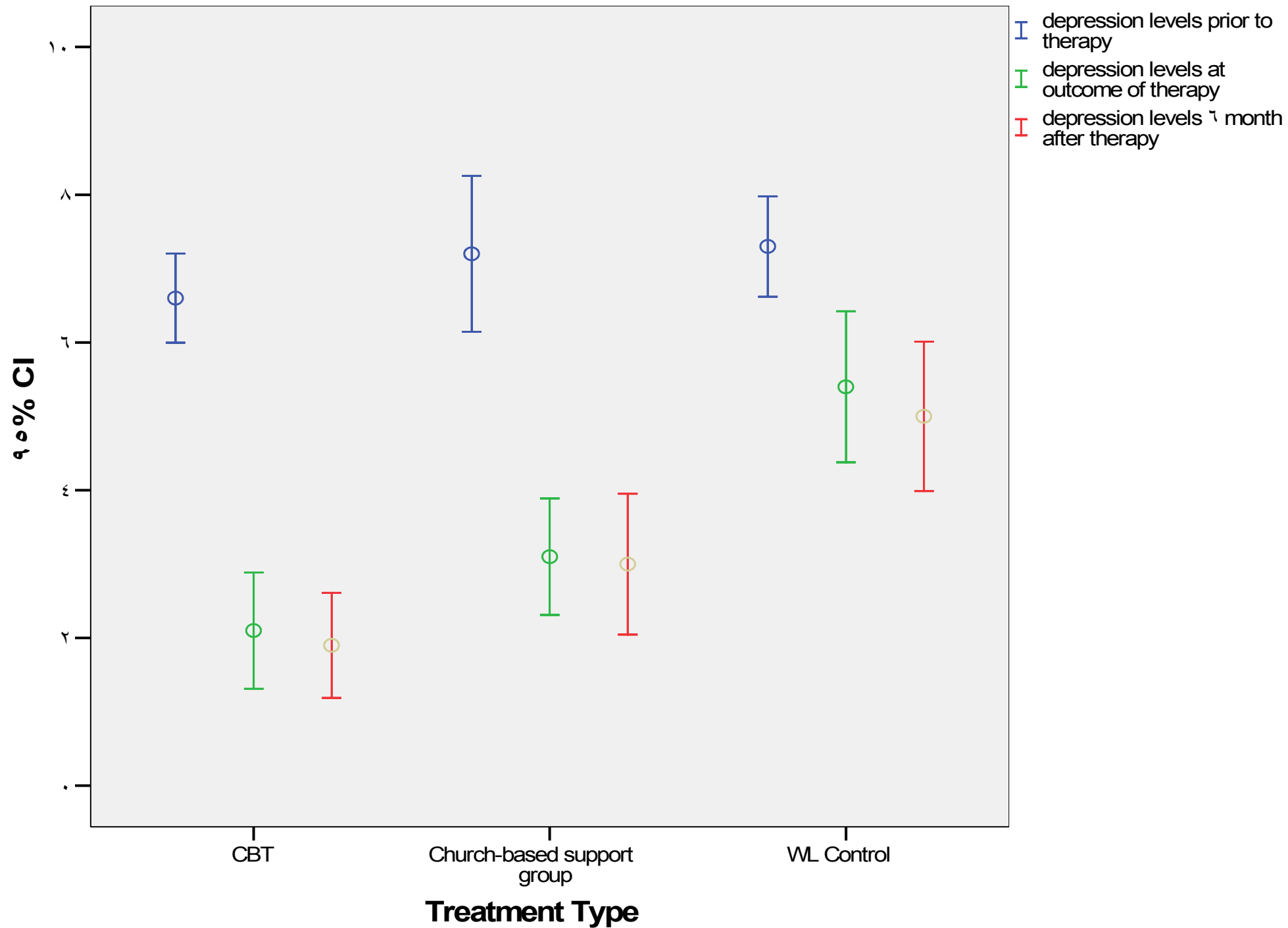
Picture: Repeated Measures

- The **same DV** (depression levels) is measured on the **same individual** at three **different times**:
 - Pre-treatment, at **outcome** of therapy, and at 6-month **follow-up**
- **SPSS**: Graphs → Legacy Dialogs → Error bar → Simple and “separate variables”:
 - **Error Bars** (DV): **depression** at 3 time points
- This graph ignores treatment group; i.e., **collapsed** across all **treatment** groups



Repeated Measures, by Group

- If we want to visualize the **treatment effect**, we can use **clustered** error bars to see the depression levels of each **treatment** group at each **time** point:
- **SPSS**: Graphs → Legacy Dialogs → Error bar → Clustered and “separate variables”:
 - **Variables** (DV): **depression** at 3 time points
 - **Category Axis** (IV): **treatment type**



Outline for today

- Core concepts of ANOVA (with pictures!)
 - One-way, factorial, RM, MANOVA, ANCOVA
- Running ANOVA in SPSS
 - Interpreting output
 - Follow-up analysis: post-hoc comparisons
 - Follow-up analysis: planned comparisons
- Assumptions of ANOVA: Parametricity
- Introduction to ANCOVA
 - Use of ANCOVA in therapy research

Running ANOVA in SPSS

- Analyze → General Linear Model → Univariate:
 - **Dependent Variable**: our outcome variable
 - ◆ e.g., depression levels at **outcome** of therapy
 - **Fixed Factors**: our (categorical) predictors
 - ◆ e.g., Treatment Type
 - Check: Options → “Estimates of effect size” and “Descriptive statistics”
- **Omnibus** test: H_0 is that **all** groups are same, H_A is that there is **some** group difference somewhere

Interpreting ANOVA Output

- There is a significant effect of **treatment type** on **depression**, $F(2,27) = 19.23, p < .001$.
- This is a strong / large effect, $\eta^2 = 59\%$

Tests of Between-Subjects Effects

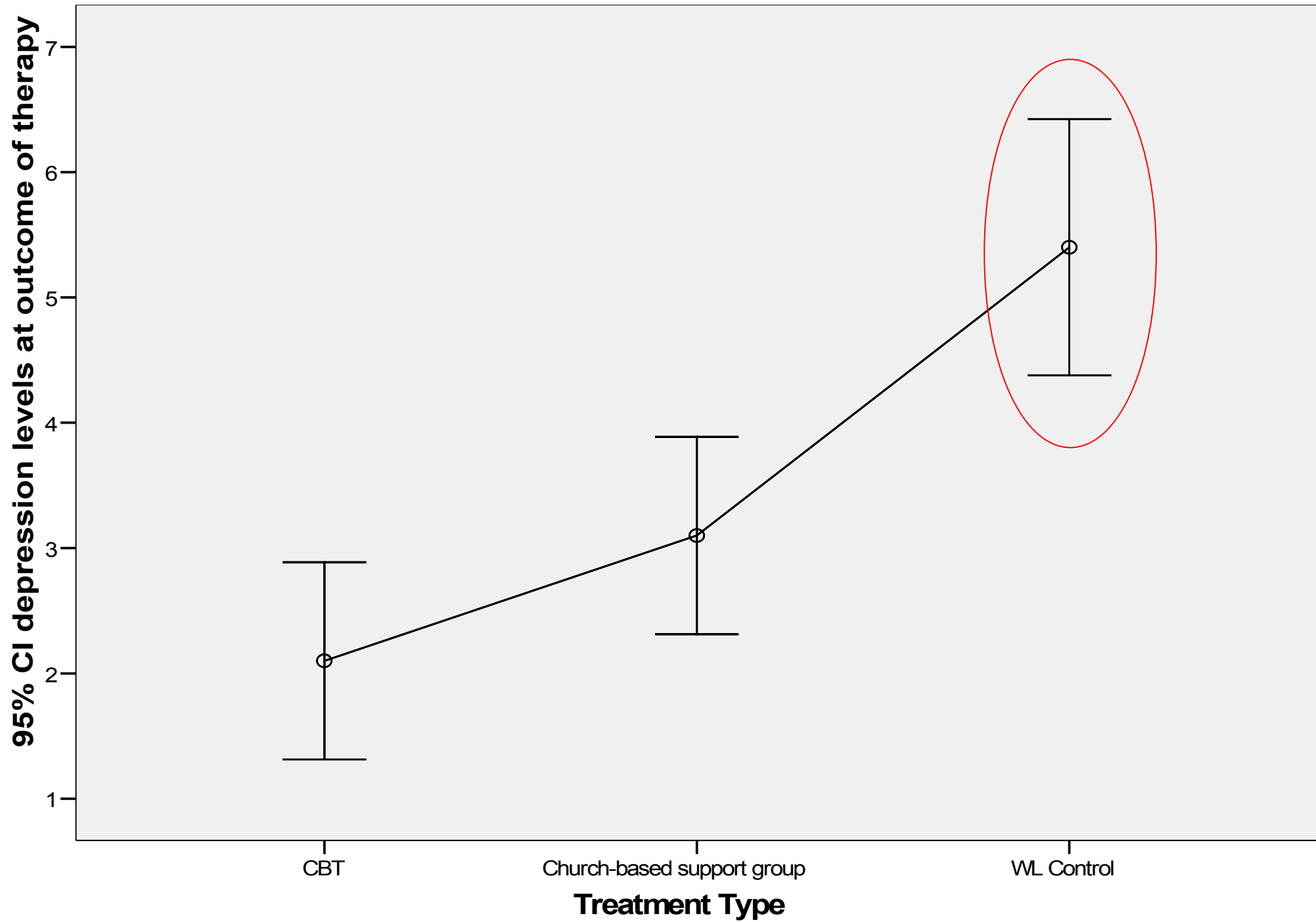
Dependent Variable: Level of trauma symptoms

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	57.267 ^a	2	28.633	19.231	.000	.588
Intercept	374.533	1	374.533	251.552	.000	.903
TREATMNT	57.267	2	28.633	19.231	.000	.588
Error	40.200	27	1.489			
Total	472.000	30				
Corrected Total	97.467	29				

a. R Squared = .588 (Adjusted R Squared = .557)

Finding Group Differences

- Yay, so the global *F*-test tells us there *is* a group difference! Now what?
 - Effect size: $\eta^2 = 59\%$: analogous to R^2
- *Where* do the differences lie? Highest group? Lowest group? *Clusters* of similar groups?
- Two strategies:
 - *Post-hoc*: *t*-tests between *all pairs* of cells
 - ◆ Watch out for α of multiple comparisons!
 - *Planned comparisons*: we tell SPSS to *focus* on a few pairings that may show differences



Outline for today

- Core concepts of ANOVA (with pictures!)
 - One-way, factorial, RM, MANOVA, ANCOVA
- Running ANOVA in SPSS
 - Interpreting output
 - Follow-up analysis: post-hoc comparisons
 - Follow-up analysis: planned comparisons
- Assumptions of ANOVA: Parametricity
- Introduction to ANCOVA
 - Use of ANCOVA in therapy research

Post-Hoc Comparisons

- In **post-hoc** analysis, we ask SPSS to run ***t*-tests** between **all** possible pairings of cells
 - Potentially lots of *t*-tests!
 - Analyze → GLM → Univariate → Post-Hoc
- Useful in **exploratory** analysis – no prior **hypotheses** about which groups might differ
- Watch out for **Type-I error**! Must **distribute** our **$\alpha=0.05$** amongst all comparisons
 - So each *t*-test gets a tiny **slice** of **α**
 - Or equivalently, adjust ***p*-values** up

Types of Post-Hoc Tests

- **Tukey** or **REGW-Q** (Ryan, Einot, Gabriel, Welch):
 - Best choice if all groups are of **equal size** and **equal variance**
- **Gabriel**: when **sizes** are roughly **similar** (~10%) and **variances** are **equal**
- **Hochberg's GT2**:
 - For **different** group **sizes** but **equal variances**
- **Games-Howell**: if **unequal variances**
 - You can select this one anyway and compare it with other methods

Post-Hoc Analysis: Notes

- SPSS's **menu** system for ANOVA has limited options for post-hoc and planned comparisons.
 - For more complex options, try **multiple regression** or SPSS **syntax**
- **Pairwise comparisons** tables help to show where specific differences lie, but the
- **Confidence intervals** must be adjusted for multiple comparisons: try **Bonferroni** or **Sidak**
 - Otherwise the **p-values** will be smaller than they ought to be

Output: Pairwise Comparisons

Pairwise Comparisons

Dependent Variable: depression levels at outcome of therapy

(I) Treatment Type	(J) Treatment Type	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
CBT	Church-based support group	-1.000	.546	.234	-2.393	.393
	WL Control	-3.300*	.546	.000	-4.693	-1.907
Church-based support group	CBT	1.000	.546	.234	-.393	2.393
	WL Control	-2.300*	.546	.001	-3.693	-.907
WL Control	CBT	3.300*	.546	.000	1.907	4.693
	Church-based support group	2.300*	.546	.001	.907	3.693

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Post-Hoc: Output

Multiple Comparisons

Dependent Variable: depression levels at outcome of therapy

			Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
(I) Treatment Type	(J) Treatment Type	Lower Bound				Upper Bound	
Tukey HSD	CBT	Church-based support group	-1.00	.546	.178	-2.35	.35
		WL Control	-3.30*	.546	.000	-4.65	-1.95
	Church-based support group	CBT	1.00	.546	.178	-.35	2.35
		WL Control	-2.30*	.546	.001	-3.65	-.95
	WL Control	CBT	3.30*	.546	.000	1.95	4.65
		Church-based support group	2.30*	.546	.001	.95	3.65
Bonferroni	CBT	Church-based support group	-1.00	.546	.234	-2.39	.39
		WL Control	-3.30*	.546	.000	-4.69	-1.91
	Church-based support group	CBT	1.00	.546	.234	-.39	2.39
		WL Control	-2.30*	.546	.001	-3.69	-.91
	WL Control	CBT	3.30*	.546	.000	1.91	4.69
		Church-based support group	2.30*	.546	.001	.91	3.69
Games- Howell	CBT	Church-based support group	-1.00	.492	.133	-2.26	.26
		WL Control	-3.30*	.571	.000	-4.76	-1.84
	Church-based support group	CBT	1.00	.492	.133	-.26	2.26
		WL Control	-2.30*	.571	.002	-3.76	-.84
	WL Control	CBT	3.30*	.571	.000	1.84	4.76
		Church-based support group	2.30*	.571	.002	.84	3.76

Based on observed means.

*. The mean difference is significant at the .05 level.

Equality of variances not assumed



Post-Hoc: Interpretation

- The various options for testing all say:
 - **Control** group (WL) is significantly different from **treatment** groups (CSG & CBT), but
 - The **treatment** groups (CSG & CBT) are **not different** from one another
- Some choices of post-hoc test are more “**conservative**” – with lower significance levels reported (e.g., Games-Howell)

Outline for today

- Core concepts of ANOVA (with pictures!)
 - One-way, factorial, RM, MANOVA, ANCOVA
- Running ANOVA in SPSS
 - Interpreting output
 - Follow-up analysis: post-hoc comparisons
 - Follow-up analysis: planned comparisons
- Assumptions of ANOVA: Parametricity
- Introduction to ANCOVA
 - Use of ANCOVA in therapy research

Planned Comparisons

- “A Priori” (“before the fact”) / planned comparisons / contrasts
- Instead of trying all pairs, use **theory** and our **hypotheses** to focus on comparisons of interest
 - “sniper” instead of “shotgun”
 - Only test **conceptually relevant** contrasts
- Planned comparisons help to overcome the problem of inflated **Type-I error** due to conducting **multiple** significance tests
- Allows for **sets** of groups to be compared (not just 1:1 **pairwise** comparisons)

Specifying Contrasts: Weights

- To specify a desired contrast, set **weights** on each **group** (each level of the categorical IV)
- All the **positive**-weight groups will be compared against all the **negative**-weight groups
- Total **sum** of weights must balance to **0**
- **Zero**-weighted groups will be omitted
- A group on **one** side cannot subsequently be **combined** with groups from the **other** side
 - So that the contrasts are **orthogonal** (test non-overlapping portions of variance)

Planned Comparisons: Example

- Treatment4.sav: 3 groups: CBT, CSG, WL
 - Contrast 1: control (WL) vs. treatment (both CBT and CSG together)
 - Contrast 2: compare two treatment methods (CBT vs. CSG)
- SPSS: (Contrast1): 1, 1, -2; (Contrast2): 1, -1, 0
- Orthogonal: 2 degrees of freedom, 2 contrasts
- What if we also had anti-anxiety drug treatment (DT) and relaxation-class control (RC)?
 - Possible orthogonal sets of contrasts?

Planned Comparisons: SPSS

- Try this in **One-Way** ANOVA (also in Univariate):
Analyze → Compare Means → One-Way ANOVA
 - Set Dependent List (**DV**) and Factor (**IV**)
 - **Contrasts**: enter weightings, in order
 - ◆ Contrast 1: (1, 1, -2). Contrast 2: (1, -1, 0)
 - **Options**: “Descriptive” (group **means**) and “Homogeneity of variance” (**Levene's** test)
- **Output**: “Contrast Tests” gives results for both “**Assume** equal variances” and “Does **not** assume equal variances”:
Use the appropriate one (from Levene's test)

Planned Comparisons: Output

Test of Homogeneity of Variances

depression levels at outcome of therapy

Levene Statistic	df1	df2	Sig.
.795	2	27	.462

Tests

Contrast 1: (1, 1, -2)

Contrast 2: (1, -1, 0)

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
depression levels at outcome of therapy	1	-5.60	.945	-5.925	27	.000
	2	-1.00	.546	-1.833	27	.078
Does not assume equal variances	1	-5.60	1.030	-5.439	14.480	.000
	2	-1.00	.492	-2.032	18.000	.057

Assume equal variances

Planned Comp.: Interpretation

- From the **results** of our planned contrasts, we conclude:
- **Contrast 1**: The control group is significantly different from the two treatment groups ($p < .001$)
- **Contrast 2**: The difference between the two treatment groups is not significant ($p = .078$)

Planned Comparisons: Notes

- Plan them out when **designing** your study, not after you have already run your ANOVA
 - Tied **conceptually** to your variables
- May need **multiple runs** to get all your desired comparisons
- SPSS provides tools in **One-Way** and GLM **Univariate** ANOVA (less convenient in Factorial)
 - Use the “**Contrasts**” button
- More **complex** designs can also be addressed using **Multiple Regression** methods
 - Use dummy-coding, include only desired IVs

Outline for today

- Core concepts of ANOVA (with pictures!)
 - One-way, factorial, RM, MANOVA, ANCOVA
- Running ANOVA in SPSS
 - Interpreting output
 - Follow-up analysis: post-hoc comparisons
 - Follow-up analysis: planned comparisons
- Assumptions of ANOVA: Parametricity
- Introduction to ANCOVA
 - Use of ANCOVA in therapy research

Assumption of Parametricity

- Interval level DV: from research design
- Independence of scores: from sampling process
- Normally distributed DV: Check for outliers, run Kolmogorov-Smirnov & Shapiro-Wilks tests:
 - Analyze → Descriptive Statistics → Explore → Plots: “Normality plots with tests”
 - Technically, need DV normal within each cell
- Equality of variances: Levene’s test
 - Analyze → General Linear Model → Univariate → Options → “Homogeneity tests”

Robustness of ANOVA

- But! ANOVA is pretty **robust** to some violations
- Show-stoppers:
 - **Non-interval** level DV
(try **non-parametric** methods instead)
 - **Dependent** scores
(try **Repeated-Measures** ANOVA or MLM)
- ANOVA becomes **more robust** when:
 - **sample sizes** are larger
 - the groups are closer to being **equal in size**
 - violations are **minor** rather than extreme

Handling Non-Normality

- If DV is **not normal** (after dealing with **outliers**):
- Check **histogram**: if **close** to normal, proceed
- Otherwise, check histograms for **each group**:
if they are all skewed in a **similar** way, proceed
 - Graphs → Legacy Dialogs → Histogram →
Rows: put IV here
- If the groups are skewed in **different** ways,
use a **non-parametric** comparison test

Handling Unequal Variances

- If **variances** are **not equal** (Levene's test):
- If **sample sizes** for each group are close to **equal**, you may proceed with ANOVA
- Otherwise, use **Welch's F** instead of regular F
 - Adjusts within-group df
 - Only available in **One-Way** ANOVA in SPSS
 - Analyze → Compare Means → One-Way ANOVA → Options: **Welch**
 - e.g., “Welch's $F(2, 17.78) = 16.25, p < .001$ ”
- Remember appropriate **post-hoc** tests (**Games-Howell**)

Assumptions Testing: Practise

- Dataset: Treatment4.sav
- DV: “depression at follow-up”
- IV: “age” (treat as categorical rather than scale)
- Check assumptions of parametricity:
 - What assumptions are violated?
 - For each violation, what should we do?

Outline for today

- Core concepts of ANOVA (with pictures!)
 - One-way, factorial, RM, MANOVA, ANCOVA
- Running ANOVA in SPSS
 - Interpreting output
 - Follow-up analysis: post-hoc comparisons
 - Follow-up analysis: planned comparisons
- Assumptions of ANOVA: Parametricity
- Introduction to ANCOVA
 - Use of ANCOVA in therapy research

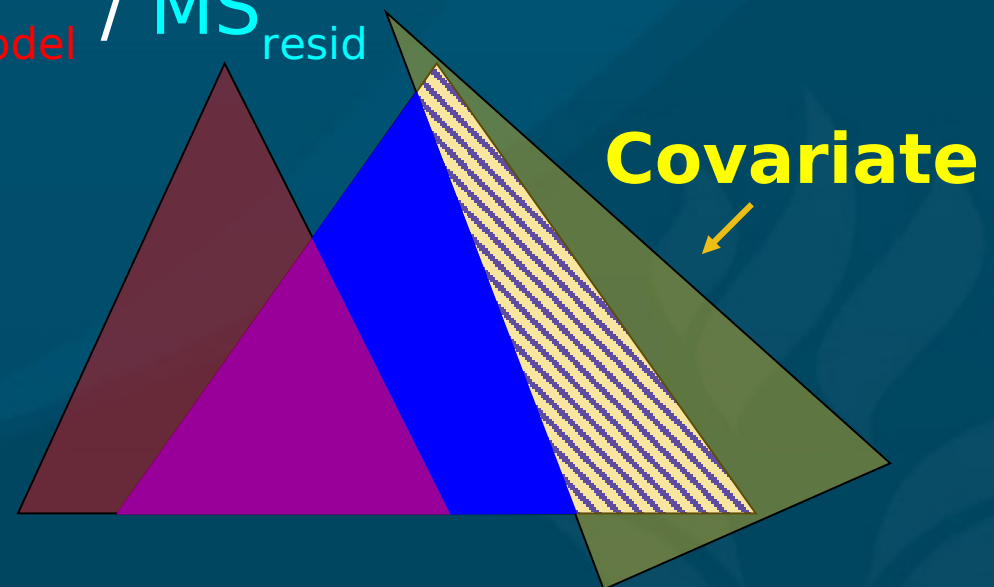
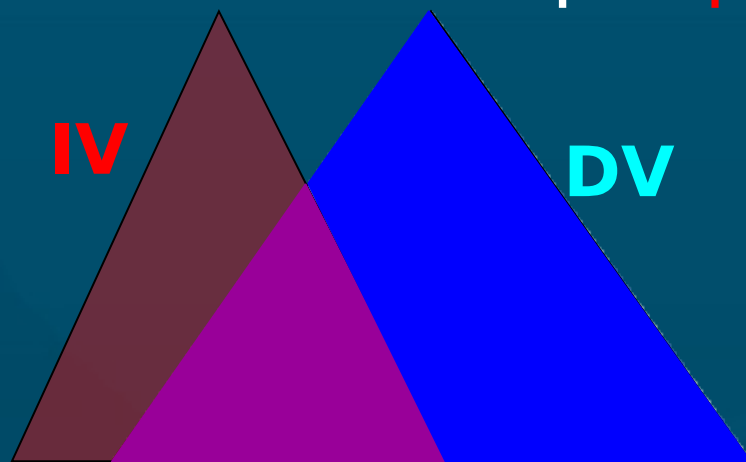
Introduction to ANCOVA

- **Covariates** are continuous “predictor” variables used as “**control**” factors to help power
- Covariates may be **promoted** to IVs if **conceptually linked** to other **IVs** or to the **DV**
- ANCOVA **factors out** the portion of **variance** in the DV that is accounted for by the covariates
 - Affects both **MS_{model}** and **MS_{residual}**
- Caution is required when covariates are **correlated** with IVs – creating **conceptual links**

Using ANCOVA in Research

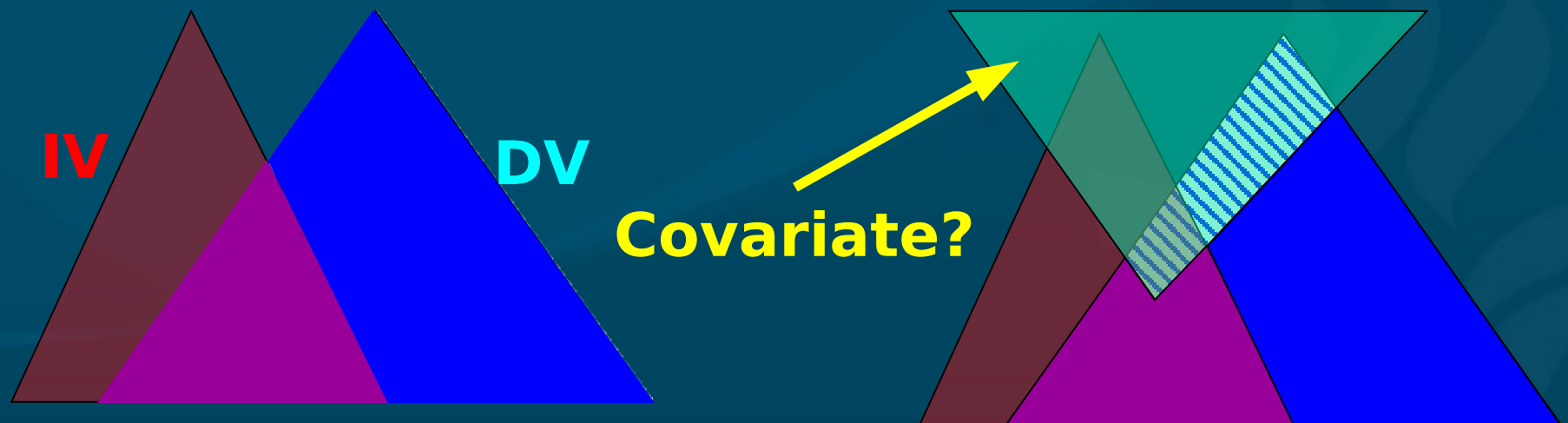
- Reduction of **error variance**: Including covariate(s) related to the DV in the model accounts for some **within-group error variance**, thus **reducing** MS_{resid} and increasing the F -ratio.

$$F = MS_{\text{model}} / MS_{\text{resid}}$$



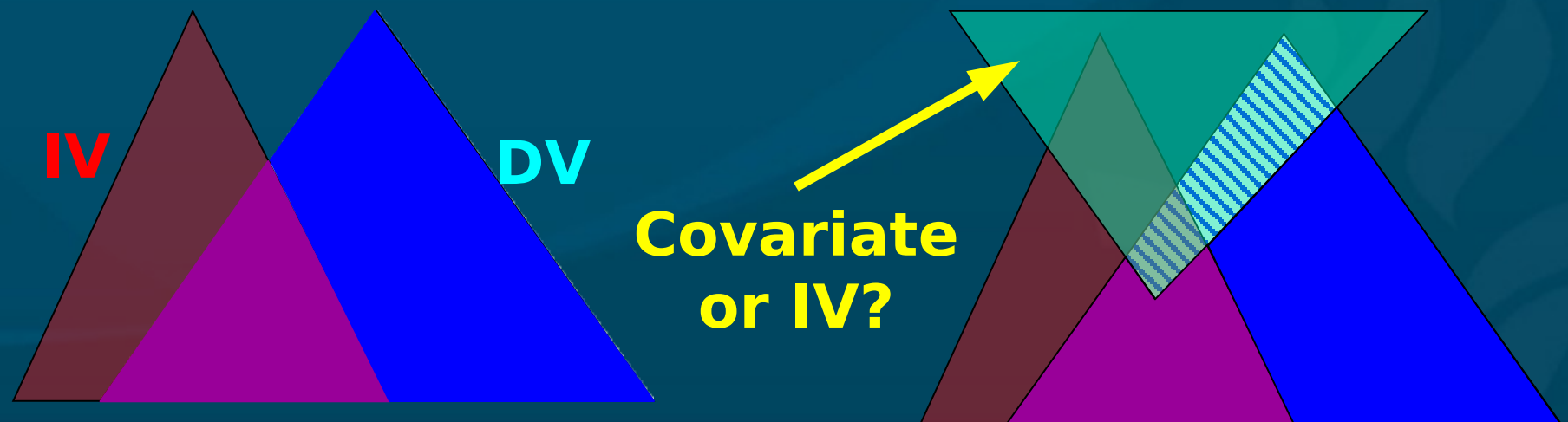
ANCOVA: Confounding Vars

- “**Confounding**” variables: External variables that may systematically **influence** an experimental manipulation.
- They can be **identified** through **theory**
- **Control** for them by entering them as **covariates** (though this may or may not **improve** *F*-ratio)



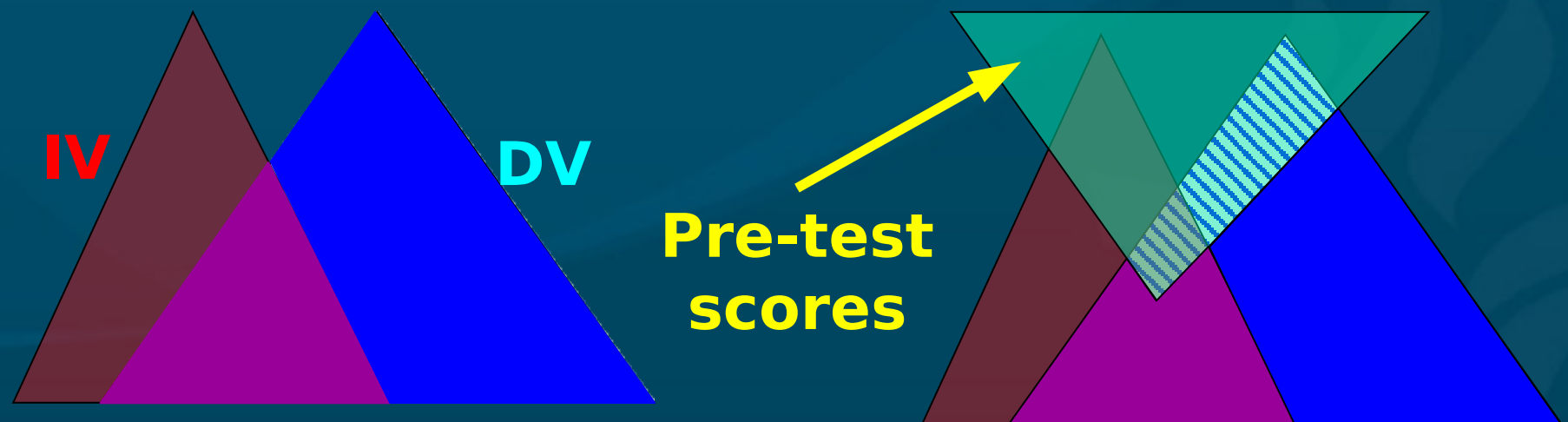
Covariates vs. Predictors

- If a covariate is 'linked' conceptually / **theoretically** with another IV or with the DV, then treat the covariate as an **IV**.
- Any **interactions** or interpretable **IV-Cov** correlations then become part of the analysis.



ANCOVA in Therapy Research

- In **therapy** studies, different treatment **groups** often have different **pre-treatment** scores
 - How to **compare** treatments when the **starting points** are different?
- **Solution**: When in doubt, treat **pre-treatment** scores as another **IV**, not as a covariate.



Assumptions of ANCOVA

- Parametricity of DV (as with regular ANOVA)
- Homogeneity of regression slopes:
 - Regression of the DV on the Cov is the same for all groups
 - Test as interactions between IVs & Cov
- Conceptual independence of Cov & IVs
 - So that the shared variance is “external” to our RQ

ANCOVA: SPSS

- Analyze → General Linear Model → Univariate
 - Add variables to “Covariates” box
- **Reporting:** “Controlling/accounting for the influence of <Cov>, the effect of <IV> on <DV> is/is not significant,
 $F(df_{IV}, df_{error}) = _, p = _.$ ”