KIN 610: Quantitative Methods in Kinesiology

Chapter 16: Factorial Analysis of Variance

Ovande Furtado Jr., PhD.

2026-04-15

Table of contents

1 Interpreting a Factorial ANOVA

2 FYI

This presentation is based on the following books. The references are coming from these books unless otherwise specified.

Main sources:

- Weir, J. P., & Vincent, W. J. (2021). *Statistics in kinesiology* (5th ed.). Human Kinetics.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- Furtado, O., Jr. (2026). *Statistics for movement science: A hands-on guide with SPSS* (1st ed.). https://drfurtado.github.io/sms

ClassShare App

You may be asked in class to go to the ClassShare App to answer questions.

- https://classshare-b44e7.web.app/

SPSS Tutorial

- SPSS Tutorial: Factorial ANOVA

3 Learning Objectives

By the end of this chapter, you should be able to:

- Define a factorial design and explain how it differs from running separate one-way ANOVAs
- Distinguish between main effects and interaction effects
- Describe how variance is partitioned in a two-way between-subjects ANOVA and a mixed factorial ANOVA
- Identify the correct error term for each F-ratio in a mixed factorial ANOVA
- Interpret an interaction plot and determine whether the interaction is qualitative or quantitative
- Conduct simple effects analysis to decompose a significant interaction
- Compute and interpret partial eta-squared (²_p) and partial omega-squared (²_p) for each effect
- Report factorial ANOVA results in APA format

4 Symbols

Symbol	Name	Pronunciation	Definition
SS_A	Factor A SS	"SS A"	Variance due to the main effect of Factor A
SS_B	Factor B SS	"SS B"	Variance due to the main effect of Factor B
$SS_{A\times B}$	Interaction SS	"SS A by B"	Variance due to the combined A × B effect
SS_{error}	Error SS	"SS error"	Within-cell residual variance
SS_{BS}	Between-subjects SS	"SS between subjects"	Stable individual differences (mixed ANOVA)
SS_{WS}	Within-subjects error SS	"SS within-subjects error"	Individual inconsistency across time (mixed ANOVA)
η^2_p	Partial eta-squared	"partial eta squared"	Effect relative to its own error term

Symbol	Name	Pronunciation	Definition
ω_p^2	Partial omega-squared	"partial omega squared"	Less-biased population effect size estimate
f	Cohen's f	"f"	Standardized effect size for power calculations

5 Why Factorial ANOVA?

The limitation of one factor at a time:

A one-way ANOVA answers: "Does the training program improve strength?"

But modern movement science asks richer questions:

- Does training improve strength? (**main effect of Group**)
- Does improvement differ for males vs. females? (**interaction**)
- How does strength change over time? (**main effect of Time**)
- Do groups follow different trajectories? (**Group × Time interaction**)

All of these questions require at least two factors operating simultaneously.

Why not run separate one-way ANOVAs?

Three problems with separate ANOVAs:

1. **Inflated Type I error** — each extra test raises the family-wise false-positive rate
2. **Cannot detect interactions** — the most scientifically important finding in factorial research simply does not exist in a series of one-way tests
3. **Larger error term** — separate ANOVAs cannot remove the variance attributable to the other factor(s) before estimating error

The factorial ANOVA solves all three problems simultaneously.

- Emphasize that interactions are the primary reason to use factorial ANOVA — they reveal *who benefits*, *when*, and *under what conditions*.

6 Factorial Design Vocabulary

Factor — each independent variable in the design (e.g., Sex, Group, Time)

Level — a discrete category of a factor (e.g., Female/Male; Control/Training; Pre/Mid/Post)

Cell — a unique combination of factor levels; each cell has its own mean

Design notation — a 2×3 design has 2 levels of one factor and 3 levels of another → 6 cells

Table 1: A 2(Sex) × 2(Group) design with 4 cells.

	Control	Training
Female	Female-Control	Female-Training
Male	Male-Control	Male-Training

Balanced design — equal cell sizes → factors are orthogonal (independent), clean SS partitioning

Unbalanced design — unequal cell sizes → SPSS uses Type III SS by default (handles this correctly)

- Draw the cell table on the board and fill in hypothetical means to preview the interaction concept.

7 Main Effects and Interactions

Main effect = the *overall* effect of one factor, averaging (collapsing) across all levels of the other factor(s)

- Main effect of **Group**: Training (85 kg) vs. Control (77 kg), averaged across Sex
- Main effect of **Sex**: Male vs. Female, averaged across Group

Main effects tell the simple, unconditional story.

Interaction = the effect of one factor *depends on* the level of the other

- If training adds 12 kg for males but only 4 kg for females → Group × Sex **interaction**
- The Group effect is no longer constant across levels of Sex

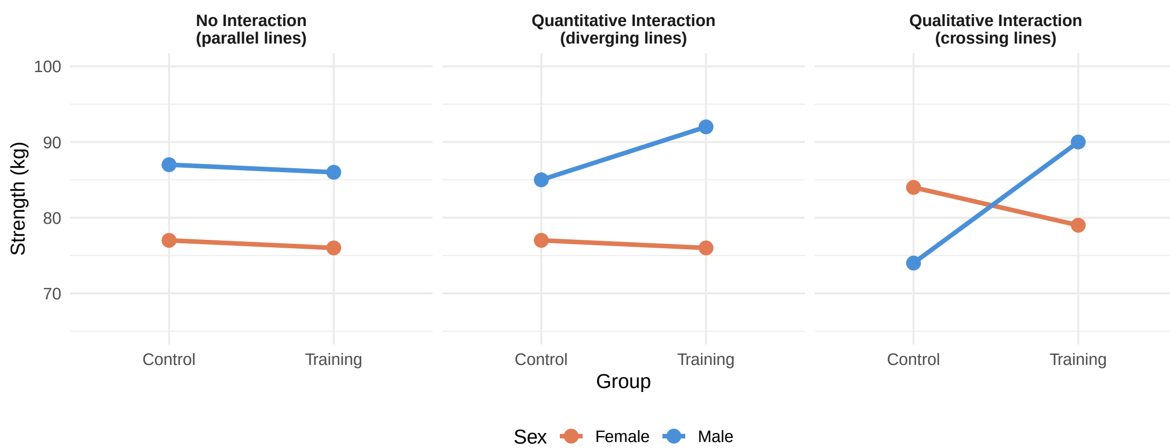
Two types of interactions:

Type	Pattern	Lines in plot
Quantitative (ordinal)	Same direction, different magnitude	Diverge — do not cross
Qualitative (disordinal)	Direction reverses	Cross

! Cardinal rule

Always check the interaction before interpreting main effects.
A significant interaction means the main effects are incomplete summaries — and can be misleading.

8 Reading an Interaction Plot



Key visual rule: Are the lines **parallel**?

- Yes → No interaction (main effects can be interpreted directly)
- No (diverging) → Quantitative interaction (same direction, different magnitude)
- No (crossing) → Qualitative interaction (direction reverses — the most dramatic type)

- Spend time on this slide — reading interaction plots is a core skill that students use constantly.

9 Between-Subjects Factorial ANOVA: Variance Partitioning

In a **two-way between-subjects ANOVA**, each participant appears in exactly one cell. Total variance splits four ways:

$$SS_{\text{total}} = SS_A + SS_B + SS_{A\times B} + SS_{\text{error}}$$

All three effects share the same within-cell error term:

$$F_A = \frac{MS_A}{MS_{\text{error}}}, \quad F_B = \frac{MS_B}{MS_{\text{error}}}, \quad F_{A\times B} = \frac{MS_{A\times B}}{MS_{\text{error}}}$$

Source	df	Uses MS_error?
Factor A	$a-1$	Yes
Factor B	$b-1$	Yes
A × B Interaction	$(a-1)(b-1)$	Yes
Error	$N-ab$	—

where a = levels of A, b = levels of B, N = total sample size, ab = number of cells.

Note

SPSS handles all calculations automatically. Focus on *what* each source tests and *which* error term it uses.

10 Worked Example: 2(Sex) × 2(Group) Between-Subjects ANOVA

Research question: Did training improve post-test strength, and did this effect differ by sex?

Design: 2(Sex: Female, Male) × 2(Group: Control, Training), $N = 60$

Sex	Group	n	M (kg)	SD
Female	Control	21	77.54	14.70
Female	Training	12	87.18	8.21
Male	Control	9	76.22	12.92
Male	Training	18	83.64	14.72
—	**Control total**	30	77.14	13.98

Sex	Group	n	M (kg)	SD
—	**Training total**	30	85.06	12.48

Source	SS	df	MS	F	p	^{2}_p
Sex	0.22	1	0.22	0.00	.974	.000
Group	939.31	1	939.31	5.22	.026	.085
Sex × Group	101.14	1	101.14	0.56	.457	.010
Error	10,084.70	56	180.08			

Interpretation: The interaction was not significant → main effects can be interpreted directly. Training improved post-test strength regardless of sex ($p = .026$, ^{2}_p = .085 — medium effect). Sex had no significant effect.

11 Mixed Factorial ANOVA: The Workhorse of Training Research

What is a mixed design?

Combines:

- 1 **between-subjects factor** (e.g., Group: Control vs. Training)
- 1 **within-subjects factor** (e.g., Time: Pre, Mid, Post)

The key question it answers:

Do the groups follow different trajectories of change over time?

This is the scientific heart of most intervention studies in movement science — and it requires a mixed ANOVA to detect.

Why two error terms?

Effect	Error term	Why?
Group (between)	$MS_{\text{Subjects/Group}}$	Stable individual differences
Time (within)	$MS_{\text{Time × Subjects/Group}}$	Within-person inconsistency only
Group × Time	$MS_{\text{Time × Subjects/Group}}$	Same within-subjects error

The within-subjects error is **much smaller** than the between-subjects error → Time and interaction F-ratios are typically much larger (and better powered) than the Group F.

Warning

Never compare F values across between-subjects and within-subjects effects to judge practical importance. Use ²_p.

12 Mixed ANOVA: Variance Partitioning

For a 2(Group) × 3(Time) mixed ANOVA, n participants per group:

Between-subjects partition:

$$SS_{\text{between subjects}} = SS_{\text{Group}} + SS_{\text{Subjects/Group}}$$

$SS_{\text{Subjects/Group}}$ = between-subjects error → denominator for F_{Group}

Within-subjects partition:

$$SS_{\text{within subjects}} = SS_{\text{Time}} + SS_{\text{Group} \times \text{Time}} + SS_{\text{Time} \times \text{Subjects/Group}}$$

$SS_{\text{Time} \times \text{Subjects/Group}}$ = within-subjects error → denominator for F_{Time} and $F_{\text{Group} \times \text{Time}}$

Three F-ratios:

$$F_{\text{Group}} = \frac{MS_{\text{Group}}}{MS_{\text{Subjects/Group}}} \qquad F_{\text{Time}} = \frac{MS_{\text{Time}}}{MS_{\text{Time} \times \text{Subj/Group}}} \qquad F_{\text{Group} \times \text{Time}} = \frac{MS_{\text{Group} \times \text{Time}}}{MS_{\text{Time} \times \text{Subj/Group}}}$$

- The key insight: the between-subjects error is large (all individual differences in absolute strength); the within-subjects error is tiny (only how inconsistently people respond across time).

13 Worked Example: 2(Group) × 3(Time) Mixed ANOVA

Research question: Did strength change across 12 weeks, and did trajectories differ by group?

Design: 2(Group) × 3(Time) mixed ANOVA; $n = 30$ per group

Group	Pre	Mid	Post
Control	76.34 (13.70)	76.85 (13.87)	77.14 (13.98)
Training	79.67 (12.26)	81.69 (12.26)	85.06 (12.48)

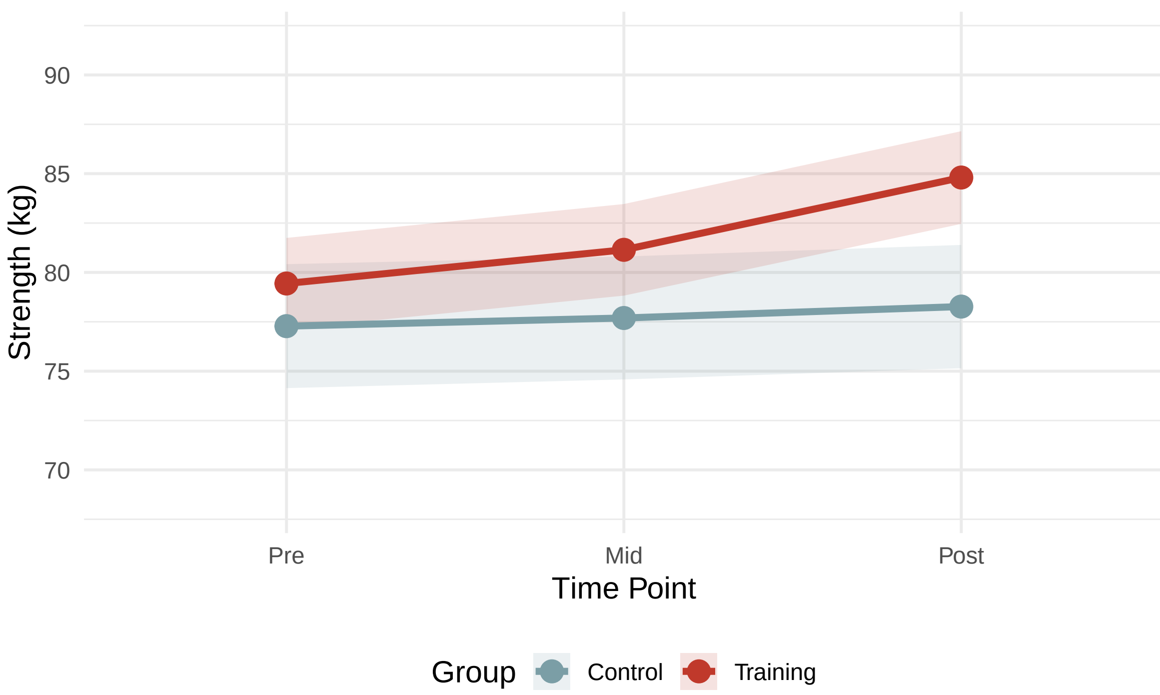
Group	Pre	Mid	Post

Mauchly's $W(2) = .93$, $p = .059$ → sphericity not violated → use Sphericity Assumed rows.

Source	*SS*	*df*	*MS*	*F*	*p*	²_p	²_p
Group	1,294.44	1	1,294.44	2.52	.117	.042	.008
Subjects/Group (error B)	29,772.65	58	513.32				
Time	290.49	2	145.24	108.55	$< .001$	.652	.544
Group × Time	163.26	2	81.63	61.00	$< .001$	.513	.400
Time × Subjects/Group (error W)	155.22	116	1.34				

Step 1 (always): The interaction is significant → decompose before interpreting main effects.

14 Visualizing the Group × Time Interaction



Error bands = ±1 SE. Diverging trajectories = significant Group × Time interaction.

Key observations:

- Control group: essentially flat across all three time points (Δ 0.8 kg total)
- Training group: progressive gains — +2.02 kg by mid, +5.39 kg by post
- Lines **diverge** (quantitative interaction) — training benefits are larger, but control does not worsen

15 Simple Effects Analysis

When to run simple effects:

Only after a **statistically significant interaction**. Running them after a non-significant interaction inflates Type I error.

Two perspectives (choose based on research question):

1. **Simple effect of Time within each Group:**
   - Does Time matter within Control? → No (F 0, p > .05)
   - Does Time matter within Training? → Yes, large increase
2. **Simple effect of Group at each Time point:**
   - Groups differ at Pre? → No (start comparable)
   - Groups differ at Mid? → Emerging difference
   - Groups differ at Post? → Significant divergence

Both perspectives are valid; the first is more common in training studies.

In SPSS:

General Linear Model → Repeated Measures → Options → Compare main effects → Bonferroni

Or: split file by Group, run separate one-way repeated measures ANOVAs.

> Examine the plot first
>
> Before running tests, look at the interaction plot. It tells you *where* the interaction is — which combinations are theoretically meaningful to test.

16 Sphericity in the Mixed ANOVA

The sphericity assumption applies to **within-subjects** effects in a mixed ANOVA:

- **Time** main effect → within-subjects error

- **Group × Time** interaction → within-subjects error

Both are affected because both use the same within-subjects error term.

Decision rule (same as Chapter 15):

Mauchly's result	Action
p > .05	Use "Sphericity Assumed" rows for Time and Group × Time
p < .05, _GG < .75	Apply **Greenhouse-Geisser** correction to Time and Group × Time
p < .05, _GG .75	Apply **Huynh-Feldt** correction to Time and Group × Time

The Group (between-subjects) effect is NOT affected by sphericity — it uses the between-subjects error, which has no sphericity assumption.

Always report which correction was applied and the epsilon value in your APA write-up.

17 Effect Sizes for Factorial ANOVA

Partial eta-squared (^{2}_p) is reported by SPSS for every effect:

$$\eta_p^2 = \frac{SS_{\text{effect}}}{SS_{\text{effect}} + SS_{\text{error}}}$$

The "error" changes by effect in a mixed ANOVA:

Effect	Error in denominator
Group	$SS_{\text{Subjects/Group}}$ (between-subjects error)
Time	$SS_{\text{Time} \times \text{Subjects/Group}}$ (within-subjects error)
Group × Time	$SS_{\text{Time} \times \text{Subjects/Group}}$ (within-subjects error)

Partial omega-squared (^{2}_p) — less biased, recommended for small-to-moderate samples:

$$\omega_p^2 = \frac{SS_{\text{effect}} - df_{\text{effect}} \cdot MS_{\text{error}}}{SS_{\text{effect}} + (N \cdot p - df_{\text{effect}}) \cdot MS_{\text{error}}}$$

Cohen's benchmarks (small .01, medium .06, large .14) apply cautiously — well-controlled lab training studies routinely produce ^{2}_p > .50 for within-subjects effects.

Warning

Report ^{2}_p AND ^{2}_p for **every** source, not only the "significant" ones. Readers need all values to evaluate power and practical significance.

18 APA Reporting Template

For a mixed ANOVA with a significant interaction:

"A 2 ([Factor A levels]) × [k] ([Factor B levels]) mixed ANOVA was conducted with [DV] as the dependent variable. Mauchly's test indicated that the sphericity assumption [was/was not] violated, W([df]) = [W], p = [p]. [Correction applied (= value) if violated.] The [Factor A × Factor B] interaction was [significant/not significant], F([df1], [df2]) = [F], p = [p], ^{2}_p = [value], ^{2}_p = [value]. [Simple effects follow-up.] The main effect of [Factor B] was [significant], F([df1], [df2]) = [F], p = [p], ^{2}_p = [value], ^{2}_p = [value]. The main effect of [Factor A] was [not significant], F([df1], [df2]) = [F], p = [p], ^{2}_p = [value]."

Full example (Group × Time):

"A 2 (Group: control, training) × 3 (Time: pre, mid, post) mixed ANOVA was conducted with strength (kg) as the dependent variable. Mauchly's test indicated that the sphericity assumption was not violated, W(2) = .93, p = .059. The Group × Time interaction was statistically significant, F(2, 116) = 61.00, p < .001, ^{2}_p = .51, ^{2}_p = .40, indicating that the two groups followed different strength trajectories over 12 weeks. Simple effects analysis revealed a significant effect of Time within the training group, with progressive gains from pre (M = 79.67 kg) to post (M = 85.06 kg), while Time was not significant within the control group. The main effect of Time was significant, F(2, 116) = 108.55, p < .001, ^{2}_p = .65, ^{2}_p = .54. The main effect of Group was not significant, F(1, 58) = 2.52, p = .117, ^{2}_p = .04."

19 Common Pitfalls

1. **Interpreting main effects when the interaction is significant** — The main effect of Group averaged over two very different sex-specific patterns describes neither pattern accurately. Always test the interaction first. If it is significant, decompose it before making any statements about individual factors.

2. **Concluding "no interaction" from a non-significant p-value** — In an underpowered study, even a moderate interaction may be undetected. Always report ^{2}_p for the interaction regardless of significance, so readers can evaluate power.

3. **Using the wrong error term in a mixed ANOVA** — Group uses the between-subjects error; Time and the interaction use the within-subjects error. Never compare F values across these effects to judge practical importance — use ^{2}_p or ^{2}_p.
4. **Failing to report effect sizes for all effects** — Report ^{2}_p and ^{2}_p for every source — Group, Time, and the interaction. Journals in kinesiology increasingly require this for all effects, not only the statistically significant ones.
5. **Running simple effects without a significant interaction** — Simple effects tests are only justified after a significant omnibus interaction. Running them after a non-significant interaction capitalizes on chance and inflates Type I error.

20 Chapter Summary

Key takeaways:

- Factorial ANOVA tests main effects and the interaction simultaneously — the interaction is the exclusive contribution of the factorial approach
- **Between-subjects factorial ANOVA**: all factors are between-subjects; all three effects share one within-cell error term
- **Mixed factorial ANOVA**: combines between-subjects (Group) and within-subjects (Time) factors; uses **two different error terms** — between-subjects error for Group, within-subjects error for Time and interaction
- The interaction must be **interpreted first** — when significant, main effects cannot be interpreted in isolation
- **Simple effects analysis** decomposes a significant interaction: test one factor separately at each level of the other
- **Sphericity** (Mauchly's test, GG/HF corrections) applies to within-subjects effects in mixed ANOVA — not to the between-subjects Group effect
- Report ^{2}_p (SPSS default) and ^{2}_p (less biased) for **every** source of variance
- A complete APA report includes: Mauchly's test, the interaction first, simple effects follow-up, both main effects, and descriptive statistics for all cells

References

1. Furtado, O., Jr. (2026). *Statistics for movement science: A hands-on guide with SPSS* (1st ed.). https://drfurtado.github.io/sms/