

Factorial Analysis of Variance: Between-Between

Ovande Furtado

2023-04-23

Table of contents

1	Sample data	2
2	Intro to <i>f</i>ANOVA	3
3	Assumptions	4
4	Equations	5
5	F Distribution	7
6	Measure of effect size	9
7	Post-hoc analysis	10
8	Result interpretation	11
8.1	Important note!	12
9	Between-subjects <i>f</i>ANOVA example	13
9.1	Research question	13
9.2	Hypothesis Statements	13
9.3	Analyzing with jamovi	14
9.4	Analyzing with SPSS	15
9.5	Analyzing with R	17
9.6	Interpreting the results	18
9.7	APA Style	19
10	Nonparametric	19
10.1	R	19

10.2 jamovi and SPSS	20
10.2.1 jamovi:	20
10.2.2 SPSS:	20

11 Other resources **21**

Learning Objectives

1. Understand the concept of Factorial Analysis of Variance (ANOVA) in a between-between design and its application in research.
2. Learn the assumptions underlying Factorial ANOVA and how to verify them using various diagnostic techniques.
3. Familiarize with the key equations and calculations involved in conducting a Factorial ANOVA by hand.
4. Learn how to interpret the results of a Factorial ANOVA, including main effects and interactions.
5. Understand the appropriate measures of effect size for Factorial ANOVA and how to interpret them using Cohen’s criteria.
6. Explore post-hoc analysis options for Factorial ANOVA, including pairwise comparisons and multiple testing corrections.
7. Learn how to conduct a Factorial ANOVA using popular statistical software packages, including jamovi, SPSS, and R Studio.
8. Understand how to report the results of a Factorial ANOVA in APA style.
9. Learn about non-parametric alternatives to Factorial ANOVA for cases where assumptions are violated, such as the Kruskal-Wallis test and the Mann-Whitney U test.
10. Develop the ability to apply Factorial ANOVA to real-world datasets and draw meaningful conclusions based on the analysis.

1 Sample data

This dataset consists of 45 participants divided into three exercise programs (A, B, and C). Each participant’s flexibility was measured at three time points (Time1, Time2, and Time3). Additionally, the dataset includes information about each participant’s gender (Male or Female).

Download¹ the dataset: [flexibility.csv](#)

```
# Load required packages
if (!require("pacman")) install.packages("pacman")
```

Loading required package: pacman

¹Right-click on the link and save as...

```
pacman::p_load("tidyverse", "dplyr", "psych", "knitr", "kableExtra")

# Read the dataset
data <- read.csv("../data/flexibility.csv", header = TRUE)

# Convert Group and Gender to factors
data$Group <- as.factor(data$Group)
data$Gender <- as.factor(data$Gender)

# Compute the descriptives table
descriptives <- data %>%
  select(Group, Gender, Flex_Time1, Flex_Time2, Flex_Time3) %>%
  group_by(Group, Gender) %>%
  psych::describe() %>%
  select(-n, -se)

# Print the descriptives table in a nice format
kable(descriptives)
```

	vars	mean	sd	median	trimmed	mad	min	max	range	skew
Group*	1	2.000000	0.8257228	2	2.000000	1.4826	1	3	2	0.000000
Gender*	2	1.555556	0.5025189	2	1.567568	0.0000	1	2	1	-0.2161948
Flex_Time1	3	27.511111	5.1769692	27	27.216216	5.9304	20	39	19	0.4549922
Flex_Time2	4	30.022222	5.2418287	29	29.756757	5.9304	22	41	19	0.4228177
Flex_Time3	5	33.022222	5.2418287	32	32.756757	5.9304	25	44	19	0.4228177

2 Intro to *f*ANOVA

In this blog post, we will delve into Factorial Analysis of Variance (ANOVA), specifically focusing on Between-Between designs. As a kinesiology professor, I understand the importance of accurate and reliable research methods for advancing our understanding of human movement and physical performance. Therefore, I am eager to share my knowledge on this topic, which will help you comprehend the value and application of Factorial ANOVA in kinesiology studies.

Factorial Analysis of Variance is a powerful and flexible statistical technique that allows researchers to examine the effects of multiple independent variables on a single dependent variable. Doing so provides a more comprehensive understanding of the relationships between variables and helps uncover potential interactions between them. The term “Between-Between”

refers to a specific type of factorial ANOVA design where all independent variables are manipulated between subjects. This means that different groups of participants are exposed to various combinations of the independent variables, and their performance or response is measured and compared.

In the field of kinesiology, this technique can be particularly valuable as it allows researchers to investigate the complex interactions between different factors that influence human movement and performance. For example, a kinesiology researcher may be interested in understanding the combined effect of training intensity and recovery time on muscle strength gains. A Between-Between factorial ANOVA design would allow them to systematically vary the training intensity and recovery time across different groups of participants and examine their combined effects on muscle strength.

As the complexity of the research question increases, so does the importance of employing appropriate experimental designs and statistical techniques. This is where the Between-Between Factorial ANOVA shines, enabling researchers to untangle the effects of multiple factors and investigate potential interactions that may have otherwise remained hidden in simpler research designs.

In this blog post, we will explore the fundamental concepts and steps involved in conducting a Between-Between Factorial ANOVA, such as hypothesis formulation, experimental design, data analysis, and interpretation of results. Furthermore, we will discuss this technique's advantages and potential limitations, along with practical tips for its application in kinesiology research.

So, without further ado, let's dive into the fascinating world of Factorial Analysis of Variance and unlock the potential of complex experimental designs for advancing our understanding of human movement and performance!

3 Assumptions

Certain assumptions must be met to conduct a Between-Between Factorial Analysis of Variance (ANOVA) accurately and obtain valid results. These assumptions are critical because they ensure that the statistical model underlying the ANOVA is appropriate for the given data. The key assumptions for a Between-Between Factorial ANOVA are as follows:

1. Independence of observations: This assumption requires that the data points collected from different participants must be independent of one another. In other words, one participant's performance or response should not influence another participant's performance or response. This assumption is often met by randomly assigning participants to different experimental groups.

2. Normality: The normality assumption states that the dependent variable's distribution should be approximately normally distributed within each combination of the independent variables' levels (i.e., within each cell of the factorial design). While the ANOVA is generally robust against violations of normality, extreme departures from this assumption can affect the validity of the results. Researchers may consider data transformations or non-parametric alternatives to the ANOVA in such cases.
3. Homogeneity of variances: This assumption requires that the variances of the dependent variable within each cell of the factorial design are roughly equal. The homogeneity of variances assumption is important because the ANOVA involves comparing the variances within and between groups. If this assumption is not met, the results may be biased. Researchers can use tests such as Levene's Test to check for the homogeneity of variances and, if necessary, consider using alternative techniques like the Welch ANOVA or data transformations to address this issue.
4. Additivity and linearity: The factorial ANOVA assumes that the relationship between the independent and dependent variables is additive and linear. This means that the effect of one independent variable on the dependent variable is constant across all levels of the other independent variable(s). In other words, the only interaction between the independent variables should exist if explicitly tested for in the ANOVA model.
5. Random sampling: Ideally, the participants in the study should be randomly sampled from the population of interest. This helps to ensure the generalizability of the study's findings to the broader population. However, true random sampling is often difficult to achieve in practice, and researchers should aim for as representative a sample as possible.

It is essential to check these assumptions before conducting a Between-Between Factorial ANOVA. Violations of these assumptions can lead to biased or unreliable results. In cases where assumptions are not met, researchers should consider alternative statistical methods, data transformations, or adjustments to the experimental design to address these issues.

4 Equations

Performing a Between-Between Factorial ANOVA by hand involves several calculations. For example, suppose you have a 2x2 factorial design with two independent variables A and B, each with two levels. The following equations are needed for hand calculations:

1. **Grand mean (GM)**: Calculate the overall mean of the dependent variable across all groups.

$$GM = \frac{\sum \sum Y}{N_{total}}$$

where $\sum \sum Y$ is the sum of all dependent variable scores, and N_{total} is the total number of observations.

2. **Sum of Squares Total (SST)**: Calculate the overall variability in the dependent variable.

$$SST = \sum \sum (Y - GM)^2$$

where Y represents individual scores, and GM is the grand mean.

3. Calculate the **main effect means**:

$$A_{mean} = \frac{(\sum Y_{A1}/N_{A1} + \sum Y_{A2}/N_{A2})}{2}$$

$$B_{mean} = \frac{(\sum Y_{B1}/N_{B1} + \sum Y_{B2}/N_{B2})}{2}$$

where $\sum Y_{A1}$, $\sum Y_{A2}$, $\sum Y_{B1}$, and $\sum Y_{B2}$ represent the sum of scores for each level of the independent variables A and B, and N_{A1} , N_{A2} , N_{B1} , and N_{B2} represent the number of observations in each level.

4. **Sum of Squares for the main effects**:

$$SS_A = \sum [(A_{mean} - GM)^2 * N_{Ai}]$$

$$SS_B = \sum [(B_{mean} - GM)^2 * N_{Bi}]$$

where N_{Ai} and N_{Bi} represent the number of observations in each level of the independent variables A and B.

5. Calculate the **interaction means**:

$$AB_{mean} = \frac{\sum Y_{AB}}{N_{AB}}$$

where $\sum Y_{AB}$ is the sum of scores for each combination of levels of A and B, and N_{AB} is the number of observations in each combination.

6. **Sum of Squares for the interaction**:

$$SS_{AB} = \sum [(AB_{mean} - A_{mean} - B_{mean} + GM)^2 * N_{AB}]$$

7. **Sum of Squares Within (SSW)**: Calculate the variability within each group.

$$SSW = SST - SS_A - SS_B - SS_{AB}$$

8. Calculate the **degrees of freedom**:

$$df_A = (\text{levels of A} - 1)$$

$$df_B = (\text{levels of B} - 1)$$

$$df_{AB} = df_A * df_B$$

$$df_W = N_{total} - (\text{levels of A} * \text{levels of B})$$

9. Calculate the **Mean Squares**:

$$MS_A = \frac{SS_A}{df_A}$$

$$MS_B = \frac{SS_B}{df_B}$$

$$MS_{AB} = \frac{SS_{AB}}{df_{AB}}$$

$$MS_W = \frac{SSW}{df_W}$$

10. Calculate the **F-ratios**:

$$F_A = \frac{MS_A}{MS_W}$$

$$F_B = \frac{MS_B}{MS_W}$$

$$F_{AB} = \frac{MS_{AB}}{MS_W}$$

Once you have calculated the F-ratios, you can compare them to the critical F-values from the F-distribution table, considering the degrees of freedom and chosen significance level (typically $= 0.05$). Suppose the calculated F-ratio is greater than the critical F-value. In that case, you can reject the null hypothesis and conclude that there is a significant effect for that independent variable or interaction.

5 F Distribution

The F distribution Furtado (2023), also known as the Fisher-Snedecor distribution, is a continuous probability distribution that is widely used in statistical hypothesis testing, particularly in the analysis of variance (ANOVA). It is named after Ronald A. Fisher and George W. Snedecor, two prominent statisticians who contributed significantly to its development.

The F-distribution used in the Between-Between Factorial ANOVA is the same as that used in One-Way ANOVA. The F-distribution is a continuous probability distribution that arises frequently as the null distribution of the test statistic in ANOVA, regardless of whether it is a One-Way or Factorial ANOVA.

However, the degrees of freedom for the F-distribution will differ between One-Way ANOVA and Factorial ANOVA. In One-Way ANOVA, the degrees of freedom are associated with the number of levels of a single independent variable.

In Factorial ANOVA, the degrees of freedom are associated with the number of levels of multiple independent variables and their interactions. When comparing F-ratios to critical F-values, you need to consider the appropriate degrees of freedom for your specific test. In both One-Way and Factorial ANOVA, you look up the critical F-value in an F-distribution table based on the numerator and denominator degrees of freedom and the chosen significance level (usually $= 0.05$). If the calculated F-ratio is greater than the critical F-value, you can reject the null hypothesis and conclude that there is a significant effect.

Some key characteristics of the F distribution are:

1. It is always non-negative, as it represents the ratio of two chi-square distributions.
2. It is asymmetric and positively skewed, with a longer tail on the right side.
3. The peak of the distribution shifts to the right as the degrees of freedom increase.
4. As both degrees of freedom approach infinity, the F distribution converges to a normal distribution.


```

# Load required packages quietly
if (!require("pacman")) install.packages("pacman", quiet = TRUE)
suppressMessages(pacman::p_load("ggplot2", "ggthemes"))

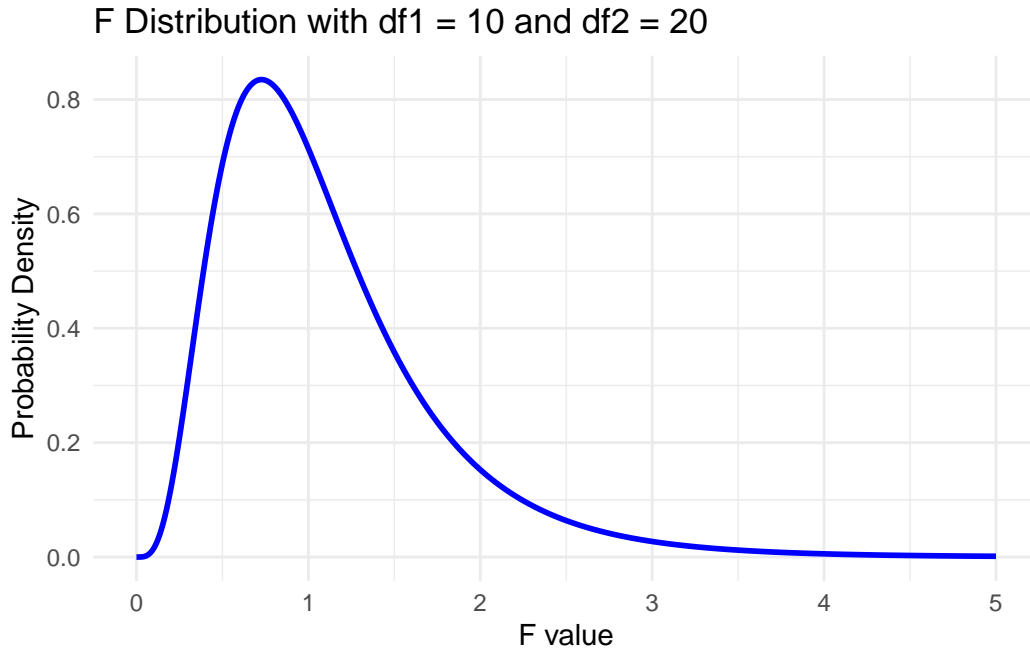
# Set the parameters for the F distribution
df1 <- 10 # degrees of freedom for the numerator
df2 <- 20 # degrees of freedom for the denominator

# Create a function to calculate the probability density function (pdf) of the F distribution
f_pdf <- function(x) {
  df(x, df1, df2)
}

# Define the range of x values to plot
x_range <- seq(0, 5, length.out = 1000)

# Plot the F distribution using ggplot2
suppressWarnings(
  ggplot(data.frame(x = x_range, y = f_pdf(x_range)), aes(x = x, y = y)) +
    geom_line(color = "blue", size = 1) +
    ggtitle(paste("F Distribution with df1 =", df1, "and df2 =", df2)) +
    xlab("F value") +
    ylab("Probability Density") +
    theme_minimal()
)

```



6 Measure of effect size

The most appropriate measure of effect size for a Between-Between Factorial ANOVA is partial eta squared (η_p^2). Partial eta squared describes the proportion of variance in the dependent variable that can be explained by an independent variable or interaction while controlling for other factors in the model. Here is the information in a Markdown format with LaTeX equations: **Partial Eta Squared (η_p^2)**: A measure of effect size for Between-Between Factorial ANOVA.

For main effects:

$$\eta_{p,A}^2 = \frac{SS_A}{SS_A + SSW}$$

$$\eta_{p,B}^2 = \frac{SS_B}{SS_B + SSW}$$

For interaction effect:

$$\eta_{p,AB}^2 = \frac{SS_{AB}}{SS_{AB} + SSW}$$

where SS_A and SS_B are the Sum of Squares for the main effects of independent variables A and B, SS_{AB} is the Sum of Squares for the interaction effect, and SSW is the Sum of Squares Within.

Partial eta squared values range from 0 to 1, with larger values indicating a greater proportion of variance explained by the independent variable or interaction effect.

Cohen (1988) provides guidelines for interpreting effect sizes, including partial eta squared (η_p^2), in the context of ANOVA. These guidelines are not strict rules but can serve as a reference for understanding the magnitude of an effect. For example, the criteria for partial eta squared are as follows:

1. Small effect: $\eta_p^2 \approx 0.01$
2. Medium effect: $\eta_p^2 \approx 0.06$
3. Large effect: $\eta_p^2 \approx 0.14$

It is important to note that these criteria should be considered within your research area's context and the data's nature. For example, effect sizes may vary depending on the field, and what is considered a large effect in one area might be different in another. Additionally, these guidelines should be used with other statistical results, such as p-values and confidence intervals, to draw meaningful conclusions from your analysis.

7 Post-hoc analysis

Post hoc analysis is necessary for a Between-Between Factorial ANOVA when you find a significant main effect or interaction effect. These analyses help identify which specific group differences are driving the significant effects. Several post hoc tests are available; some common ones include Tukey's HSD, Bonferroni, Scheffé, and Least Significant Difference (LSD).

When conducting post hoc analyses, you should choose a method that controls for the family-wise error rate (the probability of making one or more Type I errors when performing multiple comparisons). The familywise error rate increases as the number of comparisons increases, so it is essential to control for it to maintain the overall significance level.

Here is a brief overview of some popular post hoc tests:

1. **Tukey's HSD (Honestly Significant Difference):** This test controls the familywise error rate while comparing all possible pairs of means. It is recommended when you have equal sample sizes and homogeneity of variances.
2. **Bonferroni:** This method involves adjusting the significance level (α) by dividing it by the number of comparisons made. It is a conservative approach that can be used with unequal sample sizes but might have lower statistical power.

3. **Scheffé:** This test is more conservative than Tukey's HSD and applies when you have unequal sample sizes and variances. It can be used for all possible comparisons, including complex contrasts, but may have lower power.
4. **Least Significant Difference (LSD):** This method involves performing pairwise t-tests between groups. Although it is straightforward, it does not control the familywise error rate, so it has recommended using a more conservative test, like the ones mentioned above, to avoid inflated Type I errors.

When selecting a post hoc test, consider factors like sample size, homogeneity of variance, and the desired level of control over the familywise error rate. Choose a method that suits your specific data and research question.

8 Result interpretation

When interpreting the results of a Between-Between Factorial ANOVA, you should follow these steps:

1. **Check the assumptions:** Before interpreting the results, ensure that your data meet the assumptions for Factorial ANOVA, which include normality, homogeneity of variance, and independence of observations. If any assumptions are violated, you might need to transform your data or use non-parametric alternatives.
2. **Examine main effects:** Look at the F-ratios and p-values for each independent variable's main effect. If the p-value is less than your chosen significance level (e.g., $= 0.05$), you can reject the null hypothesis and conclude that there is a significant main effect for that variable.
3. **Examine interaction effects:** Assess the F-ratio and p-value for the interaction between the independent variables. If the p-value is less than your chosen significance level, you can reject the null hypothesis and conclude that there is a significant interaction effect.
4. **Effect size:** Calculate and interpret the partial eta squared (η_p^2) for each significant main effect and interaction. Use Cohen's guidelines to understand the magnitude of the effect sizes. Remember that these are only general guidelines, and effect sizes should be considered within the context of your specific research area.
5. **Post hoc tests:** If you find significant main effects or interaction effects, perform appropriate post hoc tests (e.g., Tukey's HSD, Bonferroni, Scheffé) to determine which specific group differences are driving the significant results. This step is essential for understanding the nature of the effects found in your analysis.

6. **Interpret the results in context:** Once you have determined the significant main effects, interaction effects, and specific group differences, interpret your findings within the context of your research question and the existing literature. Consider the practical implications of your results and any potential limitations or biases in your study.
7. **Report your findings:** Report the results of your analysis, including F-ratios, p-values, effect sizes, and post hoc test results. Make sure to provide a comprehensive interpretation of your findings, including your study's practical and theoretical implications.

8.1 Important note!

When conducting a Factorial ANOVA, it is crucial to assess the interaction effect between factors before interpreting the main effects. The reason for this is that a significant interaction effect can change the interpretation of the main effects.

A significant interaction effect implies that the relationship between one factor and the dependent variable is not consistent across the levels of the other factor. In other words, the effect of one factor depends on the levels of the other factor. When an interaction effect is significant, it is essential to explore and understand the nature of the interaction before making conclusions about the main effects.

If you were to interpret the main effects without considering the interaction, you might draw inaccurate conclusions about the relationship between the factors and the dependent variable. Therefore, it is crucial to assess the interaction effect first and, if it is significant, interpret the main effects within the context of that interaction.

Let's consider an example involving a study investigating the effectiveness of two teaching methods (Method A and Method B) on the test scores of male and female students. In this case, we have two factors: teaching method (with two levels: A and B) and gender (with two levels: male and female). We will use a Factorial ANOVA to analyze the data.

Suppose we find a significant interaction effect between teaching method and gender. This means that the effectiveness of the teaching methods is not consistent across genders. For instance, Method A might be more effective for male students, while Method B might be more effective for female students.

If we ignored the interaction effect and only looked at the main effects, we might conclude that one teaching method is generally more effective than the other, without considering the gender differences. However, with the significant interaction effect, we need to interpret the main effects within the context of the interaction. In this case, it would be crucial to report that the effectiveness of the teaching methods depends on the gender of the students, rather than simply stating that one method is superior overall.

By assessing the interaction effect first and interpreting the main effects within the context of the interaction, we can provide a more accurate and nuanced understanding of the relationships between the factors and the dependent variable (test scores in this example).

9 Between-subjects *f*ANOVA example

9.1 Research question

Do the exercise program (A, B, or C) and gender (Male or Female) influence flexibility at Time1?

In this research question, we still have two independent variables: exercise program (A, B, C) and gender (Male, Female). The dependent variable is now flexibility at Time1, which can be analyzed using a 3x2 (three exercise programs by two gender groups) Between-Between Factorial ANOVA.

9.2 Hypothesis Statements

Null Hypothesis (H0A): $\mu_{A1} = \mu_{A2} = \mu_{A3}$

This null hypothesis states that there is no significant difference in the means of flexibility at Time1 across the three exercise programs (i.e., the means are equal).

Null Hypothesis (H0B): $\mu_M = \mu_F$

This null hypothesis states that there is no significant difference in the means of flexibility at Time1 between males and females (i.e., the means are equal).

Null Hypothesis (H0AB)

There is no significant interaction effect between the exercise program and gender on flexibility at Time1.

Alternative Hypothesis (H1A): $\mu_{A1} \neq \mu_{A2} \neq \mu_{A3}$

This alternative hypothesis states that there is a significant difference in the means of flexibility at Time1 across the three exercise programs (i.e., the means are not equal).

Alternative Hypothesis (H1B): $\mu_M \neq \mu_F$

This alternative hypothesis states that there is a significant difference in the means of flexibility at Time1 between males and females (i.e., the means are not equal).

Alternative Hypothesis (H1AB):

There is a significant interaction effect between the exercise program and gender on flexibility at Time1.

9.3 Analyzing with jamovi

[Download](#) jamovi file (data+analysis+output)

[Download](#) jamovi output

Jamovi is an open-source statistical software with a user-friendly interface for conducting various statistical analyses, including Factorial ANOVA. For example, to perform a 3x2 Between-Between Factorial ANOVA using the latest version of Jamovi, follow these steps:

1. **Load the dataset:** Open Jamovi and import the dataset you provided. You can either drag and drop the CSV file into Jamovi or use the “Open” button at the top-left corner to navigate to the file location and open it.
2. **Set up the variables:** After importing the dataset, you will see the variables listed in the left panel. First, make sure that the “Group” and “Gender” variables are set as “Factor,” and the “Flex_Time1” variable is set as “Continuous.” If not, click on the variable type and change it to the correct type.
3. **Conduct the Factorial ANOVA:** Click on the “Analyses” tab at the top-left corner of the screen, then select “ANOVA” > “ANOVA.” This will open the ANOVA module in the right panel.
4. **Specify the dependent variable:** Drag and drop the “Flex_Time1” variable in the “Dependent Variable” box in the ANOVA module.
5. **Specify the fixed factors:** Drag and drop the **Group** and **Gender** variables into the “Fixed Factors” box.
6. **Select Post Hoc tests:** To perform post-hoc tests for significant main effects, click the “Post Hoc Tests”. Move the **Group**, **Gender**, and **Group*Gender** from the left box to the right box, and choose the desired post-hoc test method (e.g., Tukey, Scheffé, or Bonferroni) under “Correction” and choose **Cohen's d** under “Effect Size”².
7. **Choose effect size measure:** Under the “Options” tab, check the box for “Effect size” and select the desired measure of effect size (e.g., partial η^2 or η^2) from the dropdown menu.

²This is the effect size for the post hoc analysis. The effect size for the Overall model test is the partial eta squared discussed earlier.

8. **Interpret the results:** Jamovi will automatically perform the analysis and display the results in the right panel after completing the setup. Look for significant main effects of “Group” and “Gender” as well as their interaction effect by checking the p-values in the “P” column of the ANOVA table. If any of these effects are significant ($p < 0.05$), you can further explore the differences between the levels of the significant factors using the post-hoc test results.

Remember to consider the effect size and any assumptions of the Factorial ANOVA when interpreting the results.

9.4 Analyzing with SPSS

To perform a 3x2 Between-Between Factorial ANOVA using the latest version of IBM SPSS Statistics, follow these steps:

1. **Load the dataset:** Open SPSS and import the dataset you provided. You can either drag and drop the CSV file into SPSS or use the “File” > “Open” > “Data” to navigate to the file location and open it.
2. **Set up the variables:** After importing the dataset, you’ll see the variables listed in the Data View. Make sure that the “Group” and “Gender” variables are set as “Nominal” and the “Flex_Time1” variable is set as “Scale”. If not, switch to the Variable View, click on the “Measure” column for the respective variable, and change it to the correct type.
3. **Conduct the Factorial ANOVA:** In the main menu, click on “Analyze” > “General Linear Model” > “Univariate”. This will open the Univariate General Linear Model dialog box.
4. **Specify the dependent variable:** In the Univariate General Linear Model dialog box, move the “Flex_Time1” variable into the “Dependent Variable” box.
5. **Specify the fixed factors:** Move the “Group” and “Gender” variables into the “Fixed Factor(s)” box.
6. **Choose effect size measure:** Click on the “Options” button in the Univariate General Linear Model dialog box. Check the box for “Eta squared” in the “Display” section to report the effect size measure ². Click “Continue”.
7. **Select post-hoc tests:** To perform post-hoc tests for significant main effects, click on the “Post Hoc” button in the Univariate General Linear Model dialog box. Select the “Group” and “Gender” variables in the “Factors” box, and choose the desired post-hoc test method (e.g., Tukey, Scheffé, or Bonferroni) by checking the respective boxes in the “Post Hoc Tests for:” section. Click “Continue”.

8. **Run the analysis:** Click on the “OK” button in the Univariate General Linear Model dialog box to run the analysis.
9. **Interpret the results:** The output will appear in the SPSS Output Viewer. Look for significant main effects of “Group” and “Gender” as well as their interaction effect by checking the p-values in the “Sig.” column of the “Tests of Between-Subjects Effects” table. If any of these effects are significant ($p < 0.05$), you can further explore the differences between the levels of the significant factors using the post-hoc test results.

Remember to consider the effect size and any assumptions of the Factorial ANOVA when interpreting the results.

SPSS Syntax

Here is the SPSS syntax for the 3x2 Between-Between Factorial ANOVA using your dataset:

```
UNIANOVA Flex_Time1 BY Group Gender
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /POSTHOC=Group(GAMES-HOWELL) Gender(GAMES-HOWELL)
  /CRITERIA=ALPHA(.05)
  /DESIGN=Group Gender Group*Gender.
```

This syntax performs the following:

- **UNIANOVA Flex_Time1 BY Group Gender:** Specifies the dependent variable (Flex_Time1) and the fixed factors (Group and Gender).
- **/METHOD=SSTYPE(3):** Sets the type of sum of squares (Type III) for the analysis, which is the default and most commonly used in SPSS for Factorial ANOVA.
- **/INTERCEPT=INCLUDE:** Includes the intercept in the model.
- **/POSTHOC=Group(GAMES-HOWELL) Gender(GAMES-HOWELL):** Requests post-hoc tests for the main effects of Group and Gender using the Games-Howell method. You can change this to other methods (e.g., TUKEY or BONFERRONI) as needed.
- **/CRITERIA=ALPHA(.05):** Sets the alpha level at 0.05.
- **/DESIGN=Group Gender Group*Gender:** Specifies the design of the Factorial ANOVA, including the main effects (Group and Gender) and the interaction effect (Group*Gender).

To run the analysis using this syntax, copy and paste it into the SPSS Syntax Editor and click the “Run” button or press Ctrl+R. The results will appear in the SPSS Output Viewer.

Table 1: Factorial ANOVA Results

term	df	sumsq	meansq	statistic	p.value
Group	2	847.511111	423.755556	51.3563290	0.0000000
Gender	1	1.111111	1.111111	0.1346592	0.7156329
Group:Gender	2	8.822222	4.411111	0.5345971	0.5901372
Residuals	39	321.800000	8.251282	NA	NA

9.5 Analyzing with R

```
# Load the necessary packages
suppressPackageStartupMessages({
  library(readr)
  library(car)
  library(emmeans)
  library(broom)
  library(knitr)
})

# Import the dataset
data <- suppressWarnings(read_csv("../data/flexibility.csv", show_col_types = FALSE))

# Convert Group and Gender to factors
data$Group <- as.factor(data$Group)
data$Gender <- as.factor(data$Gender)

# Conduct the Factorial ANOVA
model <- aov(Flex_Time1 ~ Group * Gender, data = data)

# Print the summary of the model
anova_results <- tidy(model)
kable(anova_results, caption = "Factorial ANOVA Results")
```

```
# Post-hoc tests
emm <- emmeans(model, ~ Group | Gender)
pairwise_tukey <- pairs(emm, adjust = "tukey")
posthoc_results <- suppressWarnings(data.frame(pairwise_tukey))
kable(posthoc_results, caption = "Post-hoc Test Results")
```

Table 2: Post-hoc Test Results

contrast	Gender	estimate	SE	df	t.ratio	p.value
A - B	Female	-9.2	1.816731	39	-5.0640399	0.0000300
A - C	Female	0.3	1.573335	39	0.1906777	0.9801660
B - C	Female	9.5	1.573335	39	6.0381273	0.0000014
A - B	Male	-8.1	1.284623	39	-6.3053515	0.0000006
A - C	Male	2.6	1.573335	39	1.6525401	0.2362838
B - C	Male	10.7	1.573335	39	6.8008381	0.0000001

9.6 Interpreting the results

Based on the results from the Factorial ANOVA and Post-hoc test tables, we can interpret the findings as follows:

1. **Main effect of Group:** The p-value for the Group factor is statistically significant ($p < 0.001$). This indicates that there is a significant difference in the flexibility between the three exercise programs (A, B, and C).
2. **Main effect of Gender:** The p-value for the Gender factor is not statistically significant ($p = 0.716$). This suggests that there is no significant difference in flexibility between males and females overall.
3. **Interaction effect between Group and Gender:** The p-value for the Group:Gender interaction term is not statistically significant ($p = 0.590$). This indicates that the effect of the exercise programs on flexibility does not depend on the gender of the participants.

The post-hoc test results provide further information on the differences between the exercise programs:

- **Female participants:** There is a significant difference between Group A and Group B ($p < 0.001$), with Group B showing greater flexibility improvements. There is also a significant difference between Group B and Group C ($p < 0.001$), with Group B showing greater flexibility improvements. However, there is no significant difference between Group A and Group C ($p = 0.980$).
- **Male participants:** There is a significant difference between Group A and Group B ($p < 0.001$), with Group B showing greater flexibility improvements. There is also a significant difference between Group B and Group C ($p < 0.001$), with Group B showing greater flexibility improvements. However, there is no significant difference between Group A and Group C ($p = 0.236$).

In summary, exercise program B leads to significantly greater improvements in flexibility compared to programs A and C, irrespective of the participants' gender. There is no significant

difference in flexibility improvements between exercise programs A and C. Furthermore, the effect of exercise programs on flexibility does not depend on the gender of the participants.

9.7 APA Style

The results for this analysis can be written following the APA Style as shown below.

According to the results of a two-way factorial ANOVA, there was a significant main effect of exercise program on flexibility, $F(2, 39) = 51.36$, $p < .001$, $\eta^2 = .725$. However, there was no significant main effect of gender on flexibility, $F(1, 39) = 0.13$, $p = .716$, $\eta^2 = .003$. Additionally, no significant interaction was found between exercise program and gender, $F(2, 39) = 0.53$, $p = .590$, $\eta^2 = .027$. Post hoc comparisons using the Tukey HSD test revealed that, for both female and male participants, exercise program B led to significantly greater improvements in flexibility compared to programs A and C. Among female participants, Group B demonstrated significantly greater flexibility improvements compared to Group A ($M_{\text{diff}} = -9.2$, $SE = 1.82$, $p < .001$) and Group C ($M_{\text{diff}} = 9.5$, $SE = 1.57$, $p < .001$). No significant difference was observed between Groups A and C ($M_{\text{diff}} = 0.3$, $SE = 1.57$, $p = .980$) for female participants. Similarly, among male participants, Group B demonstrated significantly greater flexibility improvements compared to Group A ($M_{\text{diff}} = -8.1$, $SE = 1.28$, $p < .001$) and Group C ($M_{\text{diff}} = 10.7$, $SE = 1.57$, $p < .001$). No significant difference was observed between Groups A and C ($M_{\text{diff}} = 2.6$, $SE = 1.57$, $p = .236$) for male participants. In summary, exercise program B led to significantly greater improvements in flexibility compared to programs A and C, irrespective of the participants' gender. There was no significant difference in flexibility improvements between exercise programs A and C. Moreover, the effect of exercise programs on flexibility did not depend on the gender of the participants.

10 Nonparametric

10.1 R

If the data were not normally distributed, you would need to consider non-parametric alternatives to the factorial ANOVA. Non-parametric tests make fewer assumptions about the underlying distribution of the data and are more robust to violations of the normality assumption.

For a two-way factorial design with non-normal data, you can use the aligned rank transform (ART) test, which is specifically designed for non-parametric factorial analysis.

To conduct the ART test, you can use the ARTool package in R, which will enable you to analyze two-way factorial designs with non-normal data.

Keep in mind that non-parametric tests usually have less power than their parametric counterparts, so they might not detect significant effects that would be found with a parametric test. It's important to consider the distribution of your data, the assumptions of the statistical tests, and the research question when deciding which test to use.

10.2 jamovi and SPSS

You should consider the Kruskal-Wallis test as a non-parametric alternative to one-way ANOVA for each of the factors separately. However, there isn't a direct non-parametric equivalent to the two-way factorial ANOVA for between-between designs. To analyze interactions, you may need to perform post-hoc pairwise comparisons using the Mann-Whitney U test, but this approach has limitations and may increase the risk of type I errors due to multiple testing.

10.2.1 jamovi:

To perform the Kruskal-Wallis test in jamovi:

1. Open your dataset in jamovi.
2. Click on the "Analyses" tab at the top-left corner.
3. Go to "Frequencies" > "Kruskal-Wallis H test."
4. Drag the dependent variable into the "Dependent Variable" box.
5. Drag the independent variable (factor) into the "Grouping Variable" box.
6. The results will be displayed in the output window.

10.2.2 SPSS:

To perform the Kruskal-Wallis test in SPSS:

1. Open your dataset in SPSS.
2. Go to "Analyze" > "Nonparametric Tests" > "Independent Samples."
3. In the "Independent Samples" window, select "Kruskal-Wallis H" as the test type.
4. Move the dependent variable into the "Test Variable List" box.
5. Move the independent variable (factor) into the "Grouping Variable" box.
6. Click "OK" to run the test.

To analyze interactions in SPSS, you can perform pairwise comparisons using the Mann-Whitney U test. However, this approach has limitations and may increase the risk of type I errors due to multiple testing. In such cases, it's important to apply a correction for multiple comparisons (e.g., Bonferroni or Šidák correction) to control the family-wise error rate.

11 Other resources

ANOVA in jamovi (*ANOVA* — *Jamovi*, 2019)

Image credit

ANOVA — *Jamovi*. (2019, March 7). <https://www.youtube.com/watch?v=TJoJTVgDyqY>
Furtado, O. (2023, April 8). *RandomStats - One-Way ANOVA* [Blog]. RandomStats. <https://drfurtado.github.io/randomstats/posts/04082023-one-way-anova/>