

# How can we create an interaction plot in R to visualize the relationship between two or more variables?

Authored by  
**stats writer**

April 19, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can we create an interaction plot in R to visualize the relationship between two or more variables?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137177>

An interaction plot in R is a graphical representation that allows us to visualize the relationship between two or more variables. It is created by plotting the mean values of one variable against the levels of another variable. This plot helps us to understand how the relationship between the two variables changes based on the different levels of the second variable. To create an interaction plot in R, we first need to import the necessary packages and then use the "interaction.plot" function, specifying the variables and their levels. This plot is useful for identifying any interactions or dependencies between variables and for gaining insights into their relationship. It is a powerful tool for data analysis and can be easily created in R, making it a popular choice among researchers and statisticians.

## Create an Interaction Plot in R

**A two-way ANOVA is used to determine if there is a difference between the means of three or more independent groups that have been split on two factors.**

**We use a two-way ANOVA when we'd like to know if two specific factors affect a certain response variable.**

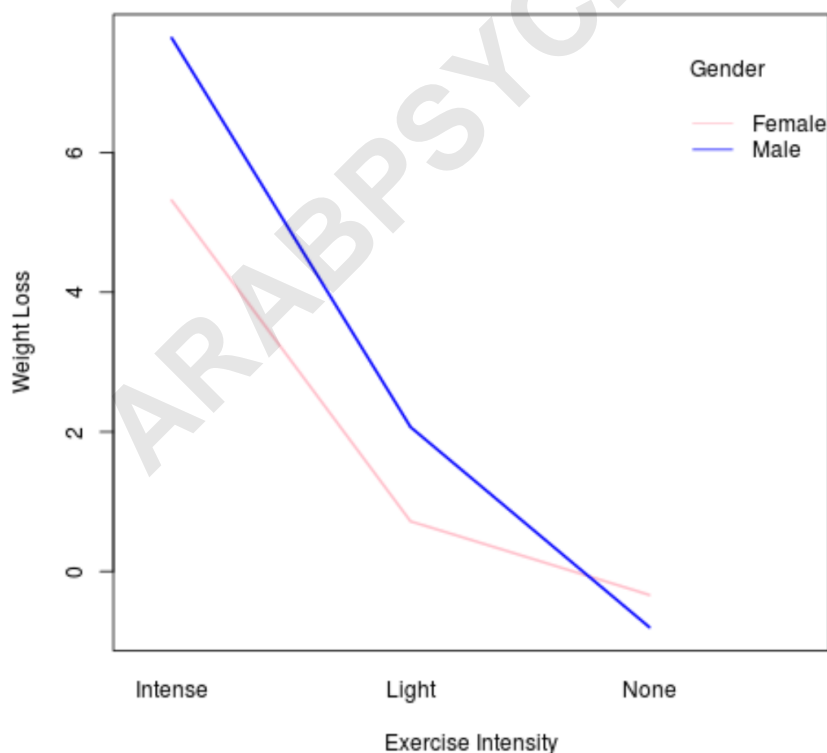
**However, sometimes there is an interaction effect present between the two factors, which can impact the way we interpret the relationship between the factors and the response variable.**

**For example, we might want to know if the factors (1) *exercise* and (2) *gender* affect the response variable *weight loss*. While it's possible that both factors affect weight loss, it's also possible that the two factors interact with each other.**

For example, it's possible that exercise leads to weight loss at different rates for males and females. In this case, there is an interact effect between exercise and gender.

The easiest way to detect and understand interaction effects between two factors is with an interaction plot.

This is a type of plot that displays the fitted values of a response variable on the y-axis and the values of the first factor on the x-axis. Meanwhile, the lines in the plot represent the values of the second factor of interest.



**This tutorial explains how to create and interpret an interaction plot in R.**

**Example: Interaction Plot in R**

**Suppose researchers want to determine if exercise intensity and gender impact weight loss. To test this, they recruit 30 men and 30 women to participate in an experiment in which they randomly assign 10 of each to follow a program of either no exercise, light exercise, or intense exercise for one month.**

**Use the following steps to create a data frame in R, perform a two-way ANOVA, and create an interaction plot to visualize the interaction effect between exercise and gender.**

**Step 1: Create the data.**

**The following code shows how to create a data frame in R:**

```
#make this example reproducible  
set.seed(10)
```

```
#create data frame
```

```
data <- data.frame(gender = rep(c("Male", "Female"),
each = 30),
exercise = rep(c("None", "Light", "Intense"), each = 10,
times = 2),
weight_loss = c(runif(10, -3, 3), runif(10, 0, 5), runif(10, 5,
9),
runif(10, -4, 2), runif(10, 0, 3), runif(10, 3, 8)))
```

**#view first six rows of data frame**

```
head(data)
```

```
gender exercise weight_loss
```

```
1 Male None 0.04486922
```

```
2 Male None -1.15938896
```

```
3 Male None -0.43855400
```

```
4 Male None 1.15861249
```

```
5 Male None -2.48918419
```

```
6 Male None -1.64738030
```

**Step 2: Fit the two-way ANOVA model.**

**The following code shows how to fit a two-way ANOVA to the data:**

```
#fit the two-way ANOVA model
```

```
model <- aov(weight_loss ~ gender * exercise, data = data)
```

```
#view the model output
```

```
summary(model)
```

```
# Df Sum Sq Mean Sq F value Pr(>F)
```

```
#gender 1 15.8 15.80 11.197 0.0015 **
```

```
#exercise 2 505.6 252.78 179.087 <2e-16 ***
```

```
#gender:exercise 2 13.0 6.51 4.615 0.0141 *
```

```
#Residuals 54 76.2 1.41
```

```
#---
```

```
#Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that the p-value (0.0141) for the interaction term between exercise and gender is statistically significant, which indicates that there is a significant interaction effect between the two factors.

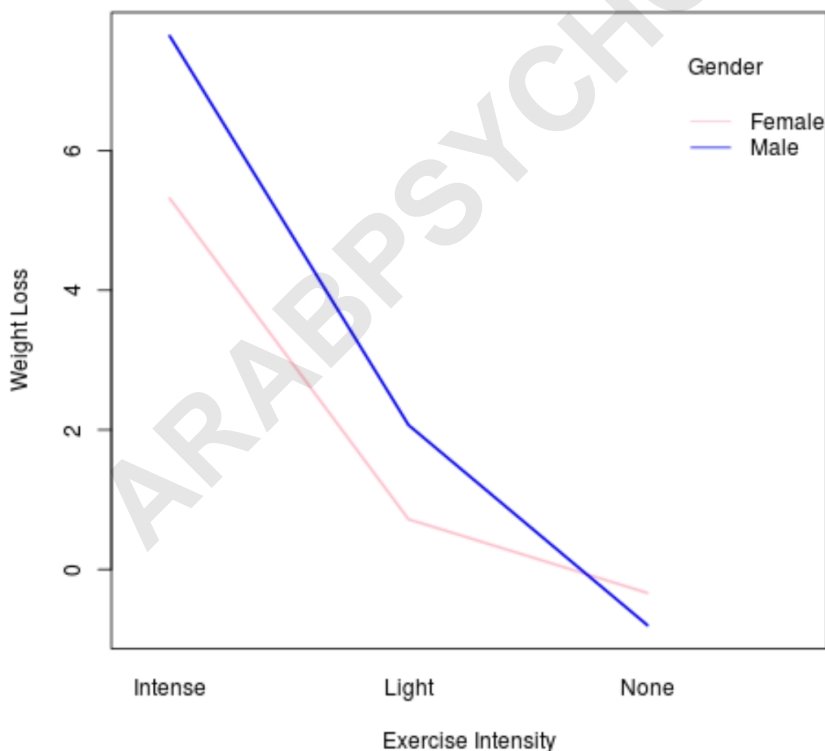
**Step 3: Create the interaction plot.**

The following code shows how to create an interaction plot for exercise and gender:

```
interaction.plot(x.factor = data$exercise, #x-axis
```

## variable

```
trace.factor = data$gender, #variable for lines  
response = data$weight_loss, #y-axis variable  
fun = median, #metric to plot  
ylab = "Weight Loss",  
xlab = "Exercise Intensity",  
col = c("pink", "blue"),  
lty = 1, #line type  
lwd = 2, #line width  
trace.label = "Gender")
```



**In general, if the two lines on the interaction plot are**

parallel then there is no interaction effect. However, if the lines intersect then there is likely an interaction effect.

We can see in this plot that the lines for males and females do intersect, which indicates that there is likely an interaction effect between the variables of exercise intensity and gender.

This matches the fact that the p-value in the output of the ANOVA table was statistically significant for the interaction term in the ANOVA model.

**How to Conduct a One-Way ANOVA in R**

**How to Conduct a Two-Way ANOVA in R**